Micro Electronic
and
Mechanical Systems

Micro Electronic
and
Mechanical Systems

Edited by
Kenichi Takahata

I-Tech

Preface

The miniaturization and performance improvement in semiconductor devices and integrated circuits (ICs) are expected to continue through leveraging of nanotechnologies and nanomaterials. This evolution should accelerate the System-on-a-Chip (SoC) trend, i.e., single-chip integration of multifunctional, mixed-signal electronic components, toward realizing embedded nanoelectronic systems. In parallel with advances in electronics, we are witnessing the rise of micro-electro-mechanical systems (MEMS), with rapidly growing commercial opportunities and markets extending to a broader range of industrial sectors on a global scale.

The emergence of MEMS is primarily attributed to the establishment of sophisticated IC manufacturing techniques and processes that served as a foundation for realizing many innovative silicon-based micromachining technologies. Advances in this area have brought about a revolution in mechanical engineering, enabling the miniaturization and system-level integration of mechanical structures and devices with ICs on a chip for MEMS fabrication. With miniaturized sensors and actuators, MEMS provide us with the ability to interact with micro-scale environments with non-electrical/-electronic parameters, found in the mechanical, optical, chemical, biological, and other domains. This exceptional ability has led to their application in fields ranging from implantable medical sensors to video game controllers. There is no doubt that continued development of MEMS and microsystems with electromechanical functionalities will extend their contribution to society, in parallel with the evolution of IC technologies.

This book discusses key aspects of these technology areas, organized in twenty-seven chapters that present the latest research developments in micro electronic and mechanical systems. The book addresses a wide range of fundamental and practical issues related to MEMS, advanced metal-oxide-semiconductor (MOS) and complementary MOS (CMOS) devices, SoC technology, integrated circuit testing and verification, and other important topics in the field. Several chapters cover state-of-the-art microfabrication techniques and materials as enabling technologies for the microsystems. Reliability issues concerning both electronic and mechanical aspects of these devices and systems are also addressed in various chapters.

This book is the result of contributions from many researchers worldwide. I would like to thank the authors for their kind cooperation and efforts to provide their most up-to-date research results. A special thanks goes to the IN-TECH team for their dedicated work in making this book possible.

November 2009

Editor

Kenichi Takahata
Canada Research Chair
University of British Columbia, Vancouver,
Canada

Contents

Membrane Micro Emboss (MeME) Process for 3-D Membrane Microdevice

Masashi Ikeuchi and Koji Ikuta
Graduate School of Engineering, Nagoya University
Japan

1. Introduction

Recent advances in micro- and nanofabrication technologies have enabled the development of miniaturized accelerometers, gyroscopes, µTAS chips, etc. These microdevices are made of substrates having thicknesses relatively greater (~100µm) than the feature scale of the microfabricated components (1~10 µm, Fig.1a). Conversely, the microscale organelles or tissues of natural creatures are made of substrates, or membranes, that are relatively thin compared to their feature size. For example, a human blood capillary, which is 10~100 µm in diameter, has vessel walls with thicknesses of ~1 µm. To give another example, a cell with a diameter of ~10 µm is composed of lipid bilayer membranes with thicknesses of ~10 nm. This fundamental characteristic of the architecture of biological microstructures, which is totally different from that of artificial microdevices, makes life a highly adaptable system from both chemical and physical perspectives. The small thickness of the membrane enhances transport of heat and substances between the body and its surroundings, and it provides softness to the body, enabling passive and active morphological changes for adapting to the environment. These characteristics of biological microstructures should greatly encourage us to develop new types of MEMS and µTAS devices. However, in reality, little research has been conducted on the development of 3-D microdevices composed of thin membranes, which we call "3-D membrane microdevices" (Fig.1b).
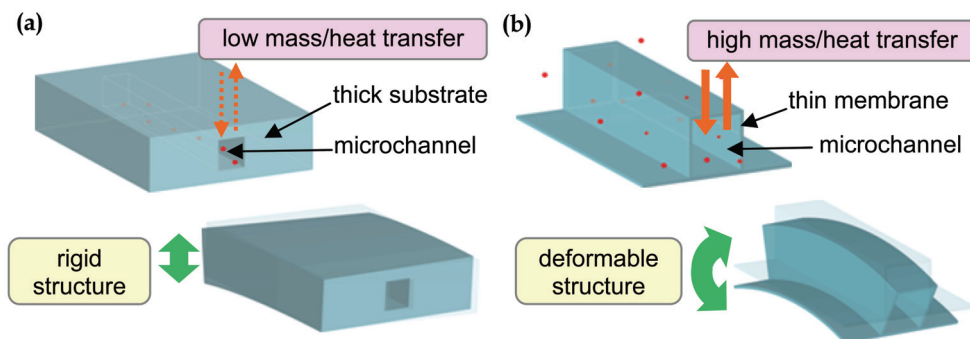
Fig. 1. Schematics of (a) conventional "bulk" microdevice and (b) "3-D membrane microdevice"

The purpose of this chapter is to introduce the concept of 3-D membrane microdevices and highlight some advances being made in our laboratory. The chapter starts with a section describing a novel microfabrication technique, namely, the membrane micro emboss (MeME) process, which was developed to realize 3-D membrane microstructures. In the following sections, several applications of 3-D membrane microdevices in μTAS and MEMS fields are presented. First, a microfluidic device composed of thin porous biodegradable membranes is described. This device was developed for tissue engineering purposes. Next, a novel micropneumatic actuator composed of folded 3-D membrane chambers is described. The actuator was intended for use as a microactive catheter for safer intravascular treatment. Finally, we conclude the chapter and present our perspectives on 3-D membrane microdevices.

2. Membrane Micro Emboss (MeME) process

Various micro fabrication processes can be used to fabricate MEMS or μTAS devices. However, few processes are useful for the fabrication of 3-D membrane microstructure, especially for polymer materials. Among conventional microfabrication processes, the chemical vapor deposition (CVD) process using parylene and the microthermoforming process can be employed. Although the CVD process using parylene is used to fabricate 3-D membrane microstructures (Zhenga et al., 2007; Liua et al., 2008), the limitations caused by the unavailability of suitable materials and low production rates present significant problems. The microthermoforming process (Truckenmüller et al., 2002; Giselbrecht et al., 2006) can be applied to a wide variety of thermoplastic materials and is suitable for mass production; however, it cannot be applied to highly porous membranes because the pressurized fluid leaks through the pores.
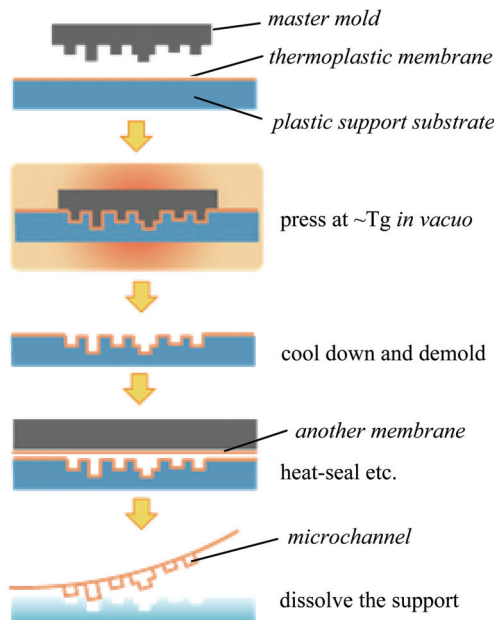
Fig. 2. Flowchart of the MeME process

The MeME process (Fig.2) was developed to realize 3-D membrane microstructures from a wide variety of materials including porous materials (Ikeuchi & Ikuta, 2005; Ikeuchi & Ikuta, 2006, a). This process needs a master mold, a thermoplastic polymer membrane, and a deformable plastic support substrate. First, the polymer membrane is set between the master mold and the support substrate. Then, this assemblage is heated to temperatures around the glass transition point (Tg) of the polymer membrane. Next, the master mold is pressurized against the membrane *in vacuo*. During pressurization, the membrane is deformed along with the support substrate to match the surface of the master mold. After cooling to the initial temperature, the master mold is separated from the deformed membrane. To fabricate sealed microchannels, another planar membrane is placed on the deformed membrane and sealed using heat-sealing, solvent gas bonding, or other sealing techniques. The fabrication of the membrane microfluidic device is completed by dissolving the support substrate in a selective solvent. The process is applicable to various materials, since it only requires the membrane material to be thermoplastic. When polylactic acid (PLA) (Tg, 57°C; thickness, 5 μm) was used as a membrane material and paraffin (melting point, 70°C) was used as a support substrate, the lateral and vertical resolutions of the process were at least 10 μm and 5 μm, respectively (Fig. 3) (Ikeuchi & Ikuta, 2006, a). The resolutions can be further improved by using thinner membranes and harder support substrates.

In the following two sections, several applications of the MeME process are described.
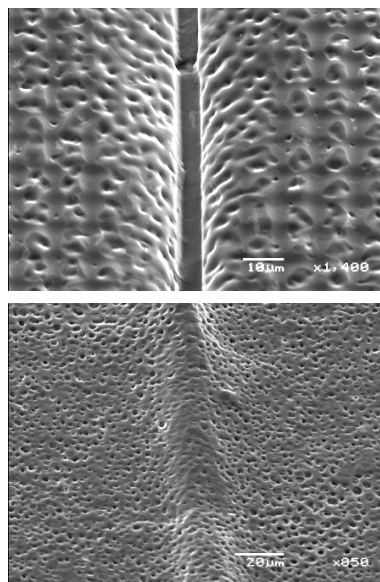
Fig. 3. SEM images of the topside (upper) and backside (lower) of the deformed porous PLA membrane.

3. Membrane microfluidic device for tissue engineering

3.1 Background

Throughout the history of biology, cell culture has been carried out on planar glass or in polymer dishes. The cells cultured on a planar substrate proliferate laterally to form a thin

layer of cells. Biologists have studied cellular dynamics using these two-dimensional cellular constructs. In the natural environment, however, cells proliferate three-dimensionally, and thus, show behaviours and functions different from those of cells in 2-D *in vitro* cultures.

Recently, cell culture in 3-D conditions has attracted considerable attention for studying natural cell behaviours and, from a more practical perspective, for regenerating fully functional large tissues and organs for transplantation. Some biologists culture cells under 3-D conditions by using soft hydrogel materials (collagen, Matrigel™, etc.) or stacking cell sheets (Liu & Bhatia, 2002; Bryant & Anseth, 2002; Sekiya et al., 2006.). There is a big difference, however, between artificial 3-D conditions and *in vivo* conditions because of the thickness of the cultured cellular constructs. Thick tissues *in vivo* can survive on nutrients supplied from surrounding blood capillary networks. In contrast, we can stack only a few layers of cells *in vitro* due to limitations with regard to the diffusion distance of nutrients, which can be supplied only through the outer surface of the construct.

To solve this problem, King et al. (2004.) attempted to construct microfluidic chips made of biodegradable polymers. They fabricated microchannels in biodegradable polymer substrates using µTAS or lab-on-chip technologies, and they cultured cells on the chip by supplying the culture medium through microchannels (Fig. 4a). They were unable to culture thick tissues, however, because the cells cultured on the chip tended to be distributed at a low density with poor homogeneity. These problems arise due to the thickness of the chip. Cells seeded on the thick microchannel chips proliferate on the surface of the chip rather than growing within the chip substrate.

In this section, we describe the MeME process as applied to fabricate 3-D thin membrane microstructure, which solves the problems associated with conventional methods for tissue engineering.

3.2 Artificial capillary network chip

To realize both the nutrients supply and homogeneous cell distribution in 3-D constructs, we propose the artificial capillary network chip as a novel 3-D cell culture device (Fig. 4b). This chip has a microchannel network made of a thin biocompatible polymer membrane with penetrating micropores.

Cells seeded on this chip with soft hydrogel materials, or cells stacked on this chip as cell sheets, can maintain a thick 3-D construct because of nutrients supplied from the porous microchannel network. Unlike the thick conventional microchannel chip, the membrane composing the microchannel wall is thin enough for cells to distribute homogeneously in the 3-D constructs. Biodegradable polymers can be used, instead of conventional polymers, as the membrane material to regenerate tissues for transplantation. Larger tissues can be fabricated by stacking these chips (Fig. 4c).

A prototype of the chip with highly branched microchannels was fabricated from a porous PLA membrane. The porous PLA membrane was formed by spin-coating following phase separation technique (Ikeuchi & Ikuta, 2006, b). The diameter and density of the pores can be controlled independently by adjusting the water content and PLA content of the coating solution, respectively. Here, the pore diameter was adjusted to $\phi \sim 1$ µm to prevent the cells ($\phi \sim 5$ µm) from entering the microchannel, and the thickness was adjusted to 5 µm.

The master mold was made by microstereolithography developed in our laboratory (Ikuta & Hirowatari, 1993). The surface of the mold was coated with a fluorocarbon polymer for easy
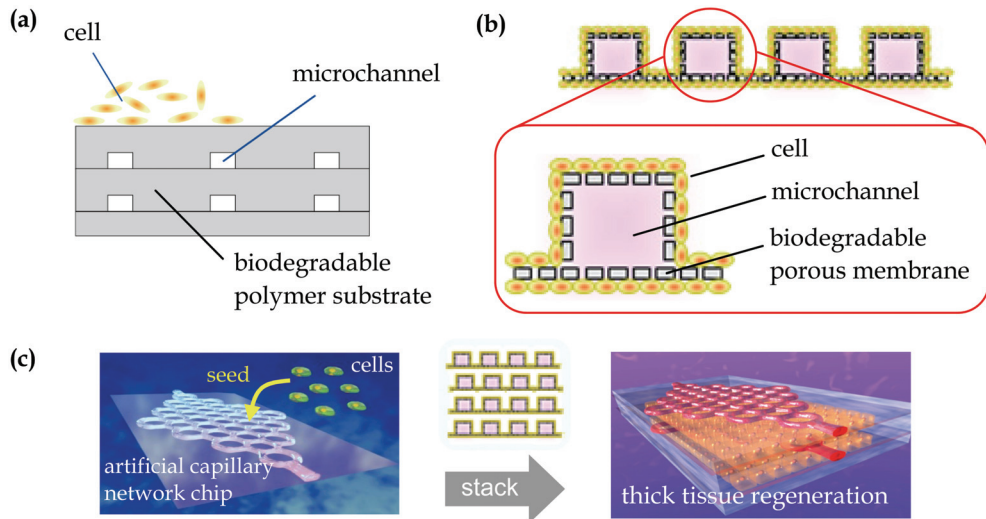
Fig. 4. (a) Schematic cross-section of the conventional microchannel chip for cell culture. (b) Schematic cross-section of artificial capillary network chip. (c) Conceptual scheme of *in vitro* 3-D thick tissue regeneration using artificial capillary network chip

removal of the mold. The master mold was pressurized onto the membrane at 0.5 μm/s for 500 s. at 55°C *in vacuo*. After cooling to 25°C, the mold was removed. The embossed membrane was heat-sealed with another membrane of the same material at 70°C for 30 s. A red solution was filled into the microchannels by capillary force. No leaking or blockage of the microchannel was observed (Fig. 5a). Figs. 5b and c show the topside and the backside of a microchannel in the prototype chip before sealing, respectively. Most of the micropores can be preserved on both sides of the microchannel wall even after the MeME process by fine tuning the process parameters (speed, temperature, and material of support substrate).

3.3 Validation of the chip

To check the size-selective permeability of the microchannel wall of the chip, a suspension of microbeads with diameters varying from $\phi100$ nm to $\phi15$ μm was poured on the chip (Fig. 6a). Beads smaller than $\phi1$ μm penetrated the wall but larger beads were trapped on the wall (Fig. 6b). This result means that nutrients and gases flowing through the microchannels can diffuse out into the cellular constructs on the chip, while at the same time, the microchannel walls support the thick 3-D cellular constructs.

The biocompatibility of the chip was also tested by culturing human endothelial cells (HUVEC) using another prototype chip. Fig. 7a shows a fluorescent image of the cells on the chip after culturing for 120 h. The cells spread as usual and showed no damage. The time course of the cell density on the chip was also equivalent to that for conventional tissue culture polystyrene flasks (Fig. 7b). These results prove that the chip was biocompatible with HUVEC. The success of HUVEC culture on the microchannel offers interesting possibilities for co-culture with other parenchyma cells to fabricate functional tissues.
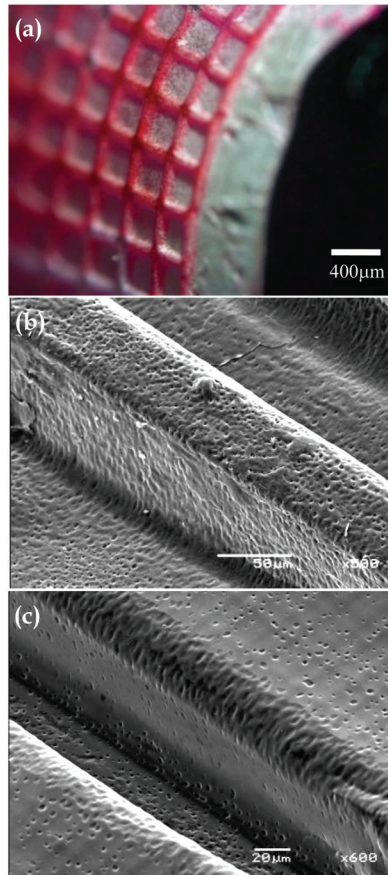
Fig. 5. A prototype chip made of a porous PLA membrane. (a) Optical microscopy image. (b, c) SEM images of the topside and the backside of the chip, respectively.
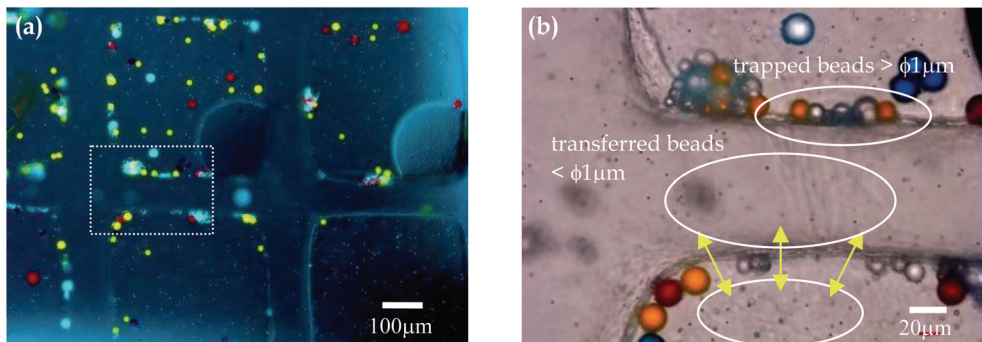
Fig. 6. (a) Fluorescent microscopy image of the chip after pouring a microbead suspension (b) Magnified view of the white-rectangle area in (a).
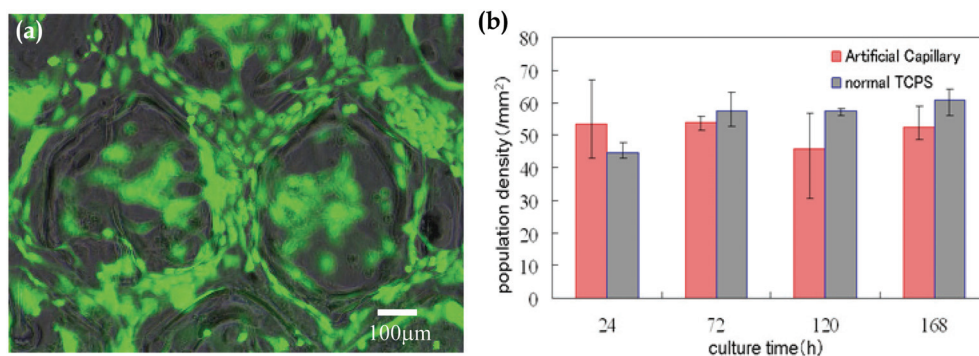
Fig. 7. HUVEC culture on the chip. (a) Fluorescent view of the cell on the chip after 120 h. (b) Transition of cell population density with culture time.

3.4 Summary

In this section, the artificial capillary network chip with a 3-D membrane microstructure was proposed and its development from the viewpoint of realizing thick 3-D tissue culture *in vitro* was described. Prototype chips were successfully fabricated using the MeME process, and their size-selective permeability and biocompatibility were verified. This chip could potentially become a key technology in the study of cellular dynamics under 3-D conditions; moreover, it could be used to regenerate large tissues or organs for transplantation in the near future.

4. Pressure-driven microactive catheter

4.1 Background

Recently, catheterization has been widely applied in intravascular surgery as an alternative to conventional surgical techniques, which are highly invasive. In catheterization, a thin flexible tube called a catheter is inserted into a blood vessel from the leg or arm. The catheter can be advanced into the patient's heart or brain for treatment or inspection. The operation leaves just a tiny puncture on the arm or leg where the catheter has been inserted, and therefore, causes less damage and fewer scars on the patient than conventional open surgery.

A major problem with catheterization, however, is the difficulty of manipulation in narrow and branched blood vessels. Since conventional catheters have no active bending capability at the tip, the doctor can control the direction of the tip only by pushing and rotating the catheter at the inlet which is far away from the tip. Thus, catheterization in narrow and complicated blood vessels is extremely difficult.

To solve this problem, several types of active catheters have been proposed (Mineta et al., 2002; Ikuta et al., 2003; Fang et al., 2007). They are classified into two types depending on the bending mechanism. The first type consists of electrically driven active catheters. These catheters have actuators that use shape memory alloys or polymer gels at the tip and can be bent from outside the body by applying a current to the actuators. Even though electrical actuators are suitable for miniaturization, the use of electricity inside the heart or brain

poses the risk of fatal damage due to microshock or heat in the case of an accident (Manecke et al., 2002; Bunch et al., 2005).

The second type consists of a pressure-driven active catheter, as proposed by Ikuta et al. (2003). It has a hollow bellows made of soft silicone rubber at the tip, and the tip can be bent by supplying saline water into the bellows through a tube connected to the bellows. Since no electricity is necessary for actuation, it is superior in safety compared to electrically driven active catheters. In addition, it can be applied to MRI monitoring, which is a fundamental tool in catheterization, because no metal parts are used in this catheter. In spite of its superiority, the minimum size of this type of catheter that can be attained with conventional injection molding processes using a pair of a male and a female mold is $\phi \sim 1$ mm, whereas the catheter must be smaller than $\phi \sim 300$ μm for complex intravascular surgery. This limitation arises due to the difficulties involved in 3-D fabrication of a pair of male and female molds with micrometer accuracy.

Although a pressure-driven balloon-type microactuator made from a polydimethylsiloxane (PDMS) molding technique was reported for use in MEMS applications (Konishi et al., 2006), it cannot be applied to catheterization due to the risk of damage to blood vessels caused by large expansion of the actuator during bending. In short, there is no process available to fabricate microscale pressure-driven active catheters.

In this section, we describe how the MeME process can be combined with an excimer laser ablation technique to realize a pressure-driven microactive catheter with a 3-D thin membrane microstructure.

4.2 Bellows composed of folded membrane microchambers

We designed a pressure-driven microactive catheter composed of hollow bellows; the catheter was made of a biocompatible polymer membrane (thickness, 5 μm), a motorized syringe, and a Teflon microtube (Fig. 8a). A pressure gauge was attached to the microtube at the base to monitor the pressure and provide the pressure value as feedback to the motorized syringe. The diameter of the catheter was set at ϕ 300 μm, because that is the minimum size used in clinical practice.

The bellows are composed of a series of folded microchambers and microchannels connecting the chambers. Since the bottom of each folded chamber is fixed to another membrane, only the upper part of the chamber can be expanded by increasing the inner pressure of the chamber (Fig. 8b). Thus, the bellows in their entirety can be bent in one direction by supplying saline water from the syringe through the microtube, because only one side of the bellows extends. Furthermore, the alternating arrangement of microchannels and microchambers prevents the bellows from expanding in diameter during bending, since the microchannels work as rigid frames to connect the topside and backside membranes of the bellows.

The catheter was fabricated using the membrane micro emboss following excimer laser ablation (MeME-X) process (Fig. 9) (Ikeuchi & Ikuta, 2008). In the MeME-X process, at first, the hollow microbellows were formed from PLA membranes (thickness, 5 μm) using the MeME process. By using excimer laser ablation (ArF, 193 nm), the outline of the bellows was cut out from the sealed membranes, and an opening was made at one end. After the bellows were connected to a microtube by an adhesive under an optical microscope, the

(a)

Bend

$50\mu m$

$300\mu m$

$200\mu m$

$150\mu m$

$150\mu m$

Syringe   Saline water

(b)

Folded bag
structure

Micro
channel

Expand
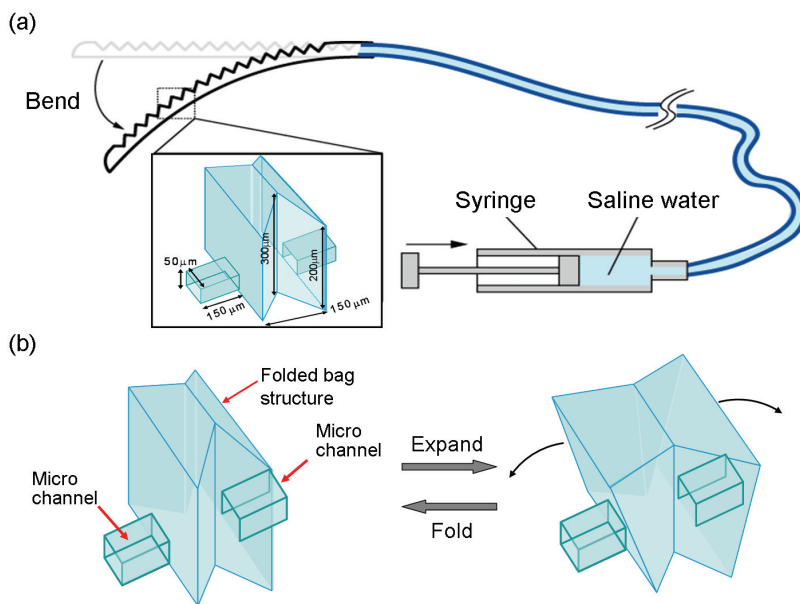
Fold

Micro
channel

Fig. 8. (a) Schematic of the pressure-driven microactive catheter system with bendable bellows made of a thin membrane at the tip (b) Bending of the bellows through expansion of each folded microchamber

mold

membrane
(poly lacticacid)

support substrate
(Paraffin)

MeME process (See Fig.2)

excimer laser ablation

Ar:F excimer
laser (193nm)

assembly

microtube
$\phi300\ \mu m$

adhesive

*Remove the support in
selective solvents*

lift-off

pressure-driven micro active catheter
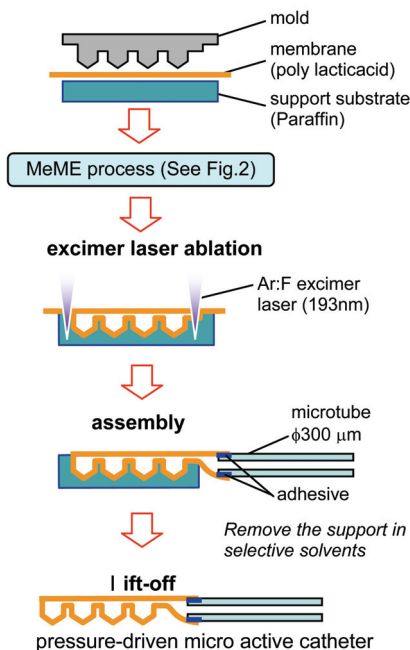
Fig. 9. Flowchart of the MeME-X process

support substrate was selectively dissolved by immersion in hexane. Finally, the catheter was successfully fabricated (Fig. 10a). The entire process was completed in 10–15min. To show the cross-section of the hollow bellows, the bellows were cut in the middle using the excimer laser. The bellows composed of a series of folded microchambers and microchannels were precisely fabricated on both the outside and the inside, and the thin membrane was uniformly deformed to yield a hollow microstructure (Fig. 10b).
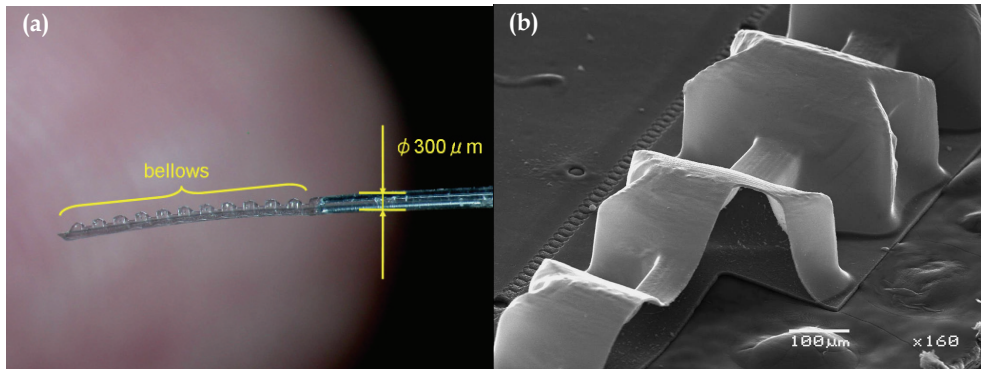
Fig. 10. (a) Completed pressure-driven micro active catheter ϕ 300 μm. (b) SEM image of the bellows cut at the middle to show the cross-section and the inner structure.

4.3 Validation of the catheter

The bellows were bent at an arbitrary angle between 0° and 180° through water pressure applied by a motorized syringe (Fig. 11a). The range of the bending angle is sufficient for intravascular operation, and it can be extended by increasing the folding angle of each microchamber of the bellows or by increasing the number of microchambers, if necessary.
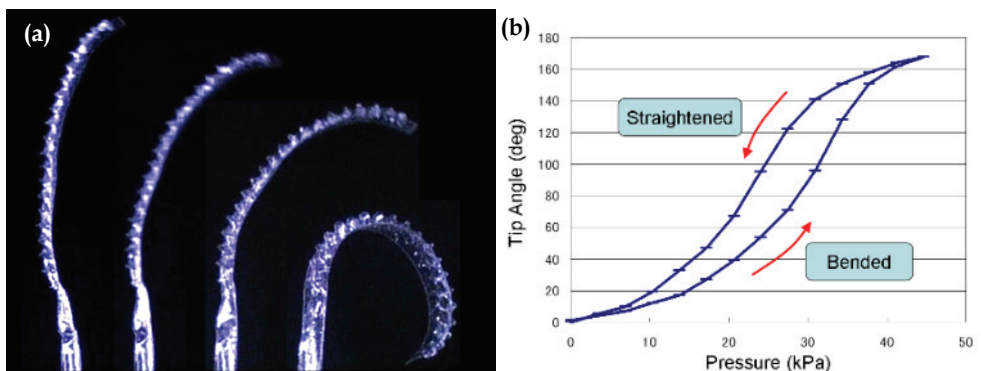
Fig. 11. (a) Bending demonstration of the pressure-driven micro active catheter from 0 to 180 degrees. (b) Relation between applied pressure (P) and bending angle (θ) of the tip

The hysteresis of the P-θ curve is apparently caused by the buckling behavior of the folded chambers and air trapped in the microtube (Fig. 11b). The buckling behavior can be improved by modifying the folding angle and pattern of the chambers, and the trapping of air in the system can be prevented by assembling the catheter *in vacuo*. Most importantly, little increase in the diameter of the bellows was observed during bending due to the microchannels inserted between the microchambers. This leads to a safer and smoother insertion of the catheter at bifurcations.

For *in vitro* demonstration of the active catheter, a small blood vessel model made of silicone was fabricated using the lost-wax method. The model consists of narrow blood vessels of φ 1 ~ 3 mm into which conventional active catheters could not be inserted. The pressure-driven microactive catheter was actuated and inserted into the narrow vessels (Fig. 12a). At the bifurcation, the catheter was bent slightly to the left from the straight position (Fig. 12b,c) by supplying saline water from the syringe, turned to the desired direction, and then successfully introduced into the target aneurysm (Fig.12d).
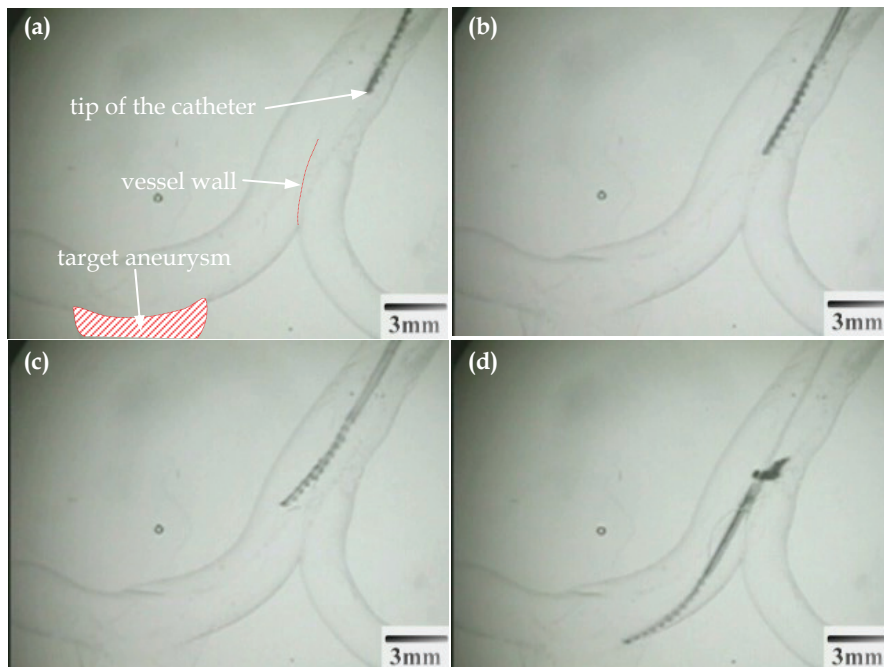
Fig. 12. Video frames showing insertion of the catheter into a 3-D vascular model

4.4 Summary

In this section, the pressure-driven microactive catheter was proposed and its development by the MeME-X process was described. The pressure-driven microactive catheter, with its extremely small size and high safety, should promote the application of catheterization in

complex intravascular surgery, which is at present not possible with conventional surgical tools. For further improvements, microchannels for drug delivery and/or blood sampling could be attached to the bellows. This can be achieved by simply adding microchannel templates on the master mold of the bellows. Furthermore, the nonelectrical actuation mechanism of this catheter, which has a 3-D membrane microstructure, can be widely extended to safe medical tools and microactuators in the microrobotics field.

5. Conclusions and perspectives

In this chapter, the concept of 3-D membrane microdevices was introduced and the development of the MeME process was described. To utilize its characteristics, the concept was applied to actual devices in two different fields. First, focusing on the efficient transfer of substances and heat in 3-D membrane microchannels, an artificial capillary network chip was developed for tissue engineering purposes. Second, utilizing the high elastic deformability of 3-D membrane microstructures, hollow bellows composed of folded microchambers and microchannels were developed to realize a pressure-driven microactive catheter.

Biological organisms are fundamentally characterized by a 3-D membrane microstructure. From intracellular organelles to vascular networks, from plant leaves to insect wings, the exquisite architectures prevalent in nature greatly inspires us to develop novel micro/nanodevices. The study of 3-D membrane microdevices has just emerged out of the proof-of-concept stage. To further expand the scope of applications of 3-D membrane microdevices, our laboratory is actively engaged in the exploration of a variety of materials applicable to the MeME process and improvement of the resolutions of the MeME process toward the nanometer scale. With its unique advantages, the 3-D membrane microdevice technology should contribute to drug delivery, tissue engineering, electric power generation, smart skin development and many other fields in the near future.

6. References

Bryant, S. & Anseth, K. (2002). Hydrogel properties influence ECM production by chondrocytes photoencapsulated in poly(ethylene glycol) hydrogels. *Journal of Biomedical Materials Research*, Vol. 59, Issue 1, 63-72.

Bunch, T.J., Bruce, G.K., Mahapatra S., Johnson S.B., Miller D.V., Sarabanda A.V., Milton M.A. & Packer D.L. (2005). Mechanisms of phrenic nerve injury during radiofrequency ablation at the pulmonary vein orifice. *Journal of Cardiovascular Electrophysiology*, Vol. 16, Issue 12, 1318-1325.

Fang, B.K., Ju, M.S. & Lin, C.C.K. (2007). A new approach to develop ionic polymer–metal composites (IPMC) actuator: Fabrication and control for active catheter systems. *Sensors and Actuators A*, Vol. 137, Issue 2, 321-329

Giselbrecht, S., Gietzelt, T., Gottwald, E., Trautmann, C., Truckenmüller, R., Weibezahn, K.F. & Welle, A. (2006). 3D tissue culture substrates produced by microthermoforming of pre-processed polymer films. *Biomedical Microdevices*, Vol. 8, Issue 3, 191-199.

Ikeuchi, M. & Ikuta, K. (2005), Fabrication of biodegradable membrane micro-channels for artificial blood capillary networks using membrane micro embossing (MeME). *Transactions of JSMBE*. Vol. 43, Issue 4, 646-652

Ikeuchi, M. & Ikuta, K. (2006,a). The membrane micro emboss (MeME) process for fabricating 3-D microfluidic device formed from thin polymer membrane. *Proc. μTAS'06*, pp. 693-695, ISBN4-9903269-0-3-C3043, Tokyo, Nov. 2006.

Ikeuchi, M. & Ikuta, K. (2006,b). On-site size-selective particle sampling using mesoporous polymer membrane microfluidic device. *Proc. μTAS'06*, pp. 1169-1171, ISBN4-9903269-0-3-C3043, Tokyo, Nov. 2006.

Ikeuchi, M. & Ikuta, K. (2008). Membrane micro emboss following excimer laser ablation (MeME-X) process for pressure-driven micro active catheter. *Proc. MEMS'08*, pp. 62-65, ISBN978-1-4244-1792-6, Tucson, Jan. 2008.

Ikuta, K. & Hirowatari, K. (1993). Real three dimensional micro fabrication using stereo lithography and metal molding. *Proc. MEMS'93*, pp. 42-47, ISBN0-7803-0957-X, Fort Lauderdale, Feb. 1993.

Ikuta, K., Ichikawa, H., Suzuki, K. & Yamamoto, T. (2003). Safety active catheter with multi-segments driven by innovative hydro-pressure micro actuators. *Proc. MEMS'03*, pp. 130-135, ISBN0-7803-7744-3, Kyoto, Jan. 2003.

King, K., Wang, C., Mofrad, M., Vacanti, J.P. & Borenstein, J. (2004). Biodegradable microfluidics. *Advanced Materials*, Vol. 16, 2007-2012

Konishi, S., Nokata, M., Jeong, O.C., Kusuda, S., Sakakibara, T., Kuwayama, M. & Tsutsumi, H. Pneumatic micro hand and miniaturized parallel link robot for micro manipulation robot system, *Proc. ICRA'06*, pp. 1036-1041, ISBN0-7803-9505-0, Orlando, May 2006.

Liu, V. & Bhatia, S. (2002). Three-dimensional photopatterning of hydrogels containing living cells. *Biomedical Microdevices*, Vol. 4, Issue 4, 257-266.

Liua, M.C., Hob, D. & Tai, Y.C. (2008). Monolithic fabrication of three-dimensional microfluidic networks for constructing cell culture array with an "integrated combinatorial mixer. *Sensors and Actuators B*, Vol. 129, Issue 2, 826-833

Manecke, G.R., Brown, J.C. Landau, A.A., Kapelanski, D.P., St. Laurent, C.M. & Auger, W.R. (2002). An unusual case of pulmonary artery catheter malfunction. *Anesthesia Analgesia*, Vol. 95, 302-304.

Mineta, T., Mitsui, T., Watanabe, Y., Kobayashi, S., Haga, Y. & Esashi, M. (2002). An active guide wire with shape memory alloy bending actuator fabricated by room temperature process. *Sensors and Actuators A*, Vol. 97-98, 632-637

Sekiya, S., Shimizu, T., Yamato, M., Kikuchi, A. & Okano, T. (2006). Bioengineered cardiac cell sheet grafts have intrinsic angiogenic potential. *Biochemical and Biophysical Research Communications*, Vol. 341, Issue 2, 573-82.

Truckenmuller, R., Rummler, Z., Schaller, T. & Schomburg, W.K. (2002). Low-cost thermoforming of micro fluidic analysis chips. *Journal of Micromechanics and Microengineering*, Vol. 12, 375–379

Zhenga, J., Webstera, J.R., Mastrangelob, C.H., Ugazc, V.M., Burnsd M.A. & Burkee, D.T. (2007). Integrated plastic microfluidic device for ssDNA separation. *Sensors and Actuators B*, Vol. 125, Issue 1, 343-351

2

A Review of Thermoelectric MEMS Devices for Micro-power Generation, Heating and Cooling Applications

Chris Gould and Noel Shammas
Staffordshire University
UK

1. Introduction

Thermoelectric technology can be used to generate a small amount of electrical power, typically in the µW or mW range, if a temperature difference is maintained between two terminals of a thermoelectric module. Alternatively, a thermoelectric module can operate as a heat pump, providing heating or cooling of an object connected to one side of a thermoelectric module if a DC current is applied to the module's input terminals. This chapter reviews the development of microelectromechanical systems (MEMS) based thermoelectric devices suitable for micro-power generation, heating and cooling applications. The chapter begins with a brief overview of thermoelectric technology, macro-thermoelectric module construction and operation. Micro-thermoelectric modules are introduced, and a review of recent developments in research, commercial development, and typical application of MEMS based micro-thermoelectric devices is made. The chapter draws conclusions on the development and potential application of MEMS based thermoelectric devices suitable for thermoelectric cooling, heating and micro-power generation.

2. Overview of thermoelectric technology, module construction and operation

2.1 Overview of thermoelectric technology

Thermoelectricity utilises the Seebeck, Peltier and Thomson effects that were first observed between 1821 and 1851 (Nolas et al, 2001). Practical thermoelectric devices emerged in the 1960's and have developed significantly since then with a number of manufacturers now marketing thermoelectric modules for cooling, heating and power generation applications. Thermoelectric power generation is mainly influenced by the Seebeck effect, with thermoelectric cooling and heating influenced predominantly by the Peltier effect. The Thomson effect does not have a major influence although it should always be included in detailed calculations (Rowe, 2006). For power generation applications, a small amount of electrical power, typically in the µW or mW range, can be generated by a thermoelectric module if a temperature difference is maintained between two terminals of a thermoelectric module. Alternatively, a thermoelectric module can operate as a heat pump, providing

heating or cooling of an object connected to one side of a thermoelectric module if a DC current is applied to the module's input terminals. The technology has achieved commercial success in mini-refrigeration, cooling and space-craft power applications, with the consumer market for mini-refrigerators and coolers currently the most successful commercial application (Hachiuma and Fukuda, 2007). Future developments in thermoelectric technology will include the need to reduce the size, and improve the performance, of current thermoelectric devices in order to address thermal problems in microelectronics, and create localised low-power energy sources for electronic systems.

2.2 Standard thermoelectric module construction

Standard thermoelectric modules are constructed from P-type and N-type thermo-elements, often referred to as thermoelectric couples, connected electrically in series and thermally in parallel. Each couple is constructed from two 'pellets' of semiconductor material usually made from Bismuth Telluride. One of these pellets is doped to create a P-type pellet, the other is doped to produce an N-type pellet. The two pellets are physically linked together on one side, usually with a small strip of copper, and placed between two ceramic plates. The ceramic plates perform two functions; they serve as a foundation on which to mount the thermo-element; and also electrically insulate the thermo-element (Riffat and Ma, 2003). A single couple of a thermoelectric module is shown below in Fig. 1.
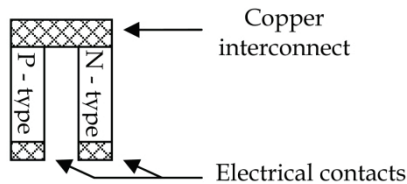
Fig. 1. A single couple of a thermoelectric module

The thermo-element, or couple, is then connected electrically in series and thermally in parallel to other couples. Standard thermoelectric modules typically contain a minimum of 3 couples, rising to 127 couples for larger devices. A schematic diagram of a thermoelectric module is shown in Fig. 2.
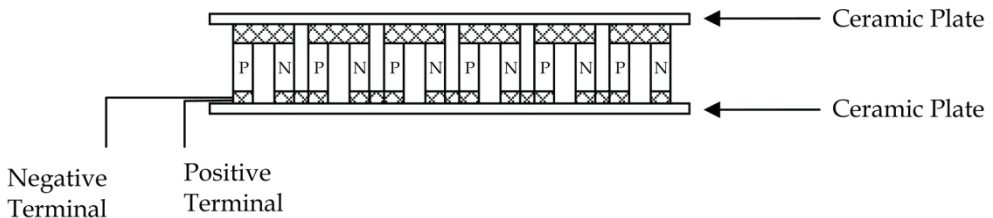
Fig. 2. A schematic diagram of a thermoelectric module

2.3 Thermoelectric module configuration

A thermoelectric module can cool or lower the temperature of an object, if the object is attached to the 'cold' side of the module, often referred to as 'TC', and DC electrical power is applied to the module's terminals. Heat from the object will be absorbed by the 'cold' side of the thermoelectric module, and transferred or 'pumped' through to the 'hot' side of the

module 'TH' due to the Peltier effect. Normally, the hot side of the module will be attached to a heat sink in order to reject this heat into the atmosphere. A thermoelectric module operating as a thermoelectric cooler or heat-pump is shown below in Fig. 3. If the polarity of the DC current applied to the thermoelectric module terminals is now reversed, the module will heat the object connected to the cold side of the module, with the other side of the module now cooling down. In this condition, the thermoelectric module is referred to as a thermoelectric heater.
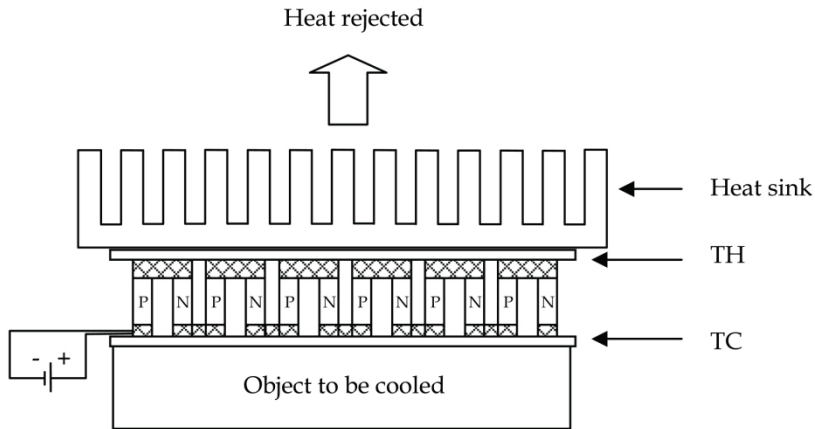
Fig. 3. A thermoelectric module operating as a thermoelectric cooler or heat-pump

A thermoelectric module can also be used to generate a small amount of electrical power, typically in the µW or mW range, if a temperature difference is maintained between both sides of the module. Normally, one side of the module is attached to a heat source and is referred to as the 'hot' side or 'TH'. The other side of the module is usually attached to a heat sink and is called the 'cold' side or 'TC'. The heat sink is used to create a temperature difference between the cold and hot sides of the module. If a resistive load (RL) is connected across the module's output terminals, electrical power will be generated in the resistive load when a temperature difference exists between the hot and cold sides of the module, due to the Seebeck effect. A thermoelectric module, operating as a thermoelectric power generator, is shown below in Fig. 4.
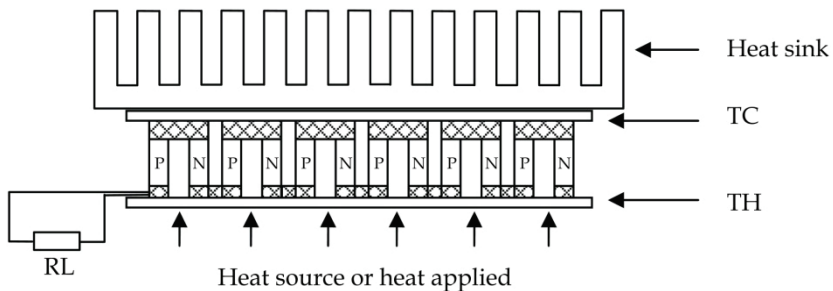
Fig. 4. A thermoelectric module operating as a thermoelectric power generator

2.4 Operation of standard thermoelectric modules

Semiconductor theory can be used to describe the operation of thermoelectric devices. In Fig. 5, a single thermoelectric couple is connected to operate as a heat pump.
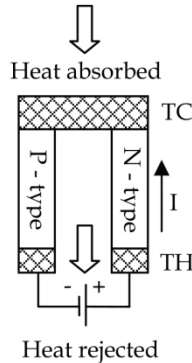
Fig. 5. A single thermoelectric couple connected as a heat pump

When a DC voltage is applied to the module terminals, electrical current flows from the positive terminal of the supply voltage to the negative terminal. This is shown as an anti-clockwise current flow in the configuration shown in Fig. 5. The negative charge carriers, i.e. the electrons, in the n-type bismuth telluride pellet are attracted by the positive pole of the supply voltage, and repelled by the negative potential. Similarly, the positive charge carriers, i.e. the holes, in the p-type material are attracted by the negative potential of the supply voltage, and repelled by the positive potential, and move in an opposite direction to the electron flow. It is these charge carriers that actually transfer the heat from one side of the thermoelectric couple to the other side in the direction of charge carrier movement. In the n-type pellet, the negatively charged electrons are the charge carriers and absorb heat from the 'cold' side of the thermoelectric couple and transfer or 'pump' this heat to the 'hot' side of the couple in a clock-wise direction. Similarly, the positively charged carriers in the p-type pellet, the holes, absorb heat from the cold side of the couple and transfer this heat to the hot side of the couple in an anti-clockwise direction. Practical thermoelectric modules are manufactured with several of these thermoelectric couples connected electrically in series and thermally in parallel. Arranging the thermoelectric couples in this way allows the heat to be pumped in the same direction.

According to (Rowe, 2006), the energy efficiency of a thermoelectric device, operating in a cooling or refrigeration mode, is measured by its Coefficient of Performance (COP), found by:

$$\text{COP} = \frac{\text{Heat absorbed}}{\text{Electrical power input}} \tag{1}$$

For thermoelectric power generation, if a temperature difference is maintained between two sides of the module, thermal energy is moving through the n-type and p-type pellets. As these pellets are electrically conductive, charge carries are transported by this heat. This movement of heat and charge carriers creates an electrical voltage, called the Seebeck voltage. If a resistive load is connected across the module's output terminals, current will flow in the load and an electrical voltage will be generated. A thermoelectric couple connected as a thermoelectric power generator is shown in Fig. 6.

Heat input
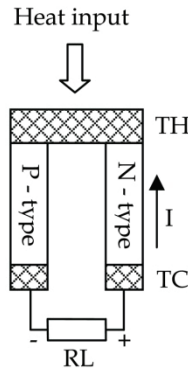
Fig. 6. A single thermoelectric couple connected as a thermoelectric power generator

The efficiency of a thermoelectric module, operating as a power generator, can be found by:

$$\eta \ = \frac{\text{Energy supplied to the load}}{\text{Heat energy absorbed at the hot junction}} \qquad (2)$$

In thermoelectricity, efficiency is normally expressed as a function of the temperature over which the device is operated, referred to as the dimensionless thermoelectric figure-of-merit ZT.
The thermoelectric figure of merit ZT can be found by:

$$\text{ZT} \ = \frac{\alpha^2 \sigma}{\lambda} \qquad (3)$$

where α is the Seebeck coefficient, σ is the electrical conductivity, and λ is the total thermal conductivity (Sales, 2007).
Thermoelectric phenomena are exhibited in almost all conducting materials, with the exception of superconductors below specific temperatures. Materials which possess a ZT > 0.5 are usually regarded as thermoelectric materials (Rowe, 2006). The best thermoelectric materials used in commercial macro-thermoelectric devices, Bi_2Te_3-Sb_2Te_3 alloys, operating around room temperature, have typical values of α=225μV/K, $\sigma = 10^5/\Omega m$, and $\lambda = 1.5$ W/mK, which results in ZT $\approx$ 1 (Sales, 2007). Bismuth Telluride is the most common material used in standard thermoelectric modules, as it exhibits the most pronounced thermoelectric effect around room temperature. Other material combinations are also used including; Alloys based on bismuth in combination with antimony, tellurium and selenium; lead telluride; and silicon germanium alloys (Rowe, 2006).

2.5 Development of micro-thermoelectric modules
Standard thermoelectric modules range in size from 4 x 4 x 3 mm^3 to around 50 x 50 x 50 mm^3. Although, in principle, the dimensions can be reduced further, the fabrication of conventional thermoelectric modules for power generation or heating and cooling applications is a bulk technology, and is incompatible with microelectronic fabrication processes (Volklein & Meier, 2006). The development of micro-thermoelectric devices that are compatible with standard microelectronic technology and manufacturing processes have the potential to enhance the performance of microelectronic systems, achieve significant

reductions in size, improve the performance of thermoelectric devices, and open up new areas of research and commercial application.

Until recently, thermoelectric devices have been confined to niche applications because of their relatively low conversion efficiency and thermoelectric figure-of-merit ZT when compared with other technologies (Riffat & Ma, 2003). For thermoelectric power generation, current thermoelectric efficiencies are between 5% to 10% (Nuwayhid et al, 2005), with a practical thermoelectric figure-of-merit ZT ~ 1. For thermoelectric cooling and refrigeration, a COP of 0.5 is typical, which is lower than that achieved by conventional refrigeration techniques (Bass et al, 2004). According to (Stabler, 2006), since the early 1990's, materials with ZT > 1 have been discovered, and reports of ZT ~ 2 are widely known today with evidence that higher values of ZT are possible (Vining, 2007). Improving the efficiency and thermoelectric figure-of-merit ZT, reducing the cost of thermoelectric devices, and the use of alternative materials that are more widely available are focus areas for current research activity. However, thermoelectric technology does have several advantages over other technologies; For cooling or refrigeration applications, thermoelectric modules do not use any chlorofluorocarbons or other materials that require periodic replenishment; they can achieve precise temperature control to within +/- 0.1°C; the same thermoelectric device can be used for heating or cooling and can cool to temperatures below 0°C (Riffat & Ma, 2003); the modules are electrically quite in operation and are relatively small in size and weight (Alaoui & Salameh, 2001); and do not import dust or any other particles that could cause an electrical short circuit.

3. Thermoelectric MEMS devices

3.1 Overview
There is an increasing amount of published research in support of developing MEMS based thermoelectric devices. MEMS technology, combined with microelectronics and micromachining techniques, has been successfully and widely utilised in micro-sensor and micro-actuator applications, and there is significant commercial value in developing next generation thermoelectric devices for applications in power generation and integrated circuit cooling (Huang et al, 2007). Current micro-sensors and micro-actuators may also be based on thermal and thermoelectric principles, and use thin-film technology to achieve sensing and actuator functionality, with micromachining techniques to achieve device optimisation (Volklein & Meier, 2006). According to (Min, 2006), the development of thermoelectric devices compatible with standard semiconductor manufacturing processes has the potential to address many applications in microelectronics, with MEMS technology, along with nanotechnology, of significant interest to thermoelectric manufacturers and researchers. It is anticipated that these technologies can be used to reduce the size, and improve the performance, of thermoelectric devices suitable for micro-power generation, heating and cooling applications. Current MEMS based devices will also benefit from incorporating thermoelectric technology, for example where a MEMS based device has an electrical power consumption in the micro-watt range, this could potentially be supplied by thermoelectric devices (Huesgen et al, 2008), or where there is a need for temperature stabilisation of MEMS based microelectronic components and circuits (Li et al, 2003).

3.2 Emerging thermoelectric MEMS based devices
Research into manufacturing a thermoelectric MEMS based device, using thin-film technology, has resulted in the proposal of different device structures; a vertical device

structure; and a horizontal device structure (Min, 2006). Commercially available micro-thermoelectric devices, based on thin-film technology, have also recently started to emerge. According to (Vining, 2007), two start-up companies have started to market thermoelectric devices based on thin-film technology. One company has developed thermoelectric devices based on a MEMS like process that use a sputtering deposition method and Bi_2Te_3 related materials. Another company has developed thermoelectric devices based on Bi_2Te_3-Sb_2Te_3 superlattice technology. (Bottner et al, 2007; 2005; 2004; 2002) describe in some detail the development of thin-film MEMS like thermoelectric devices using a sputtering deposition technique. Similarly, (Venkatsubramanian et al, 2007) and (Koester et al, 2006) describe the development of commercial thermoelectric devices using superlattice nanoscale materials.

The concept of MEMS like thermoelectric devices for cooling and micro-power generation applications, using a thin-film sputtering deposition technique, is to have a common vertical architecture of thermoelectric devices that use standard silicon/silicondioxide wafers as a substrate. One of these wafers is used to create an n-type semiconductor using Bi_2Te_3 related materials, and another, separate wafer is used to create a p-type semiconductor. The Bi_2Te_3 related material is deposited using a sputtering method, and after dry etching to create the device structure, the wafers are then sawn in order to create a single n-type and p-type die. The n-type and p-type die are then soldered together to create a thermoelectric couple (Bottner, 2005).

Another approach to creating micro-thermoelectric devices, that are compatible with modern semiconductor processing techniques, is the development of thin-film thermoelectric devices using nanoscale materials. According to (Venkatsubramanian et al, 2007), significant developments have occurred in the last few years in the area of nanoscale thermoelectric materials using superlattices and self-assembled quantum dots. Thin-film thermoelectric superlattices can be manufactured using Planar semiconductor device technology and are compatible with standard microelectronic processing and packaging tools.

There are a number of other examples of recently published work into MEMS based thermoelectric devices. Although not an exhaustive list, a basic literature search will highlight activity by (Liu et al, 2007) on the integration of micro-thermoelectric devices into a silicon based light-emitting diode (LED) in order to stabilise the LED's temperature; a planar multi-stage micro-thermoelectric device for cooling applications is presented by (Hwang et al, 2008); and the development of two micro-thermoelectric cooling devices, one based on a column-type telluride material, and another using a bridge-type polysilicon material and fabricated using MEMS based techniques by (Huang et al, 2008).

3.3 Future application of MEMS based thermoelectric devices

MEMS based thermoelectric devices can be used in thermoelectric cooling, heating and micro-power generation applications. The miniaturisation of thermoelectric modules, and the potentially higher thermoelectric performance that can be obtained, will also allow the development of new applications to emerge.

Micro-thermoelectric devices, fabricated in thin-film technology, have achieved sufficient miniaturisation to be integrated inside semiconductor packaged devices, rather than having to be mounted onto the outside of a semiconductor device, as is normal with a macro-thermoelectric module. As the semiconductor industry further reduces the size of transistors in integrated circuits, a trend is to fabricate more of the external circuitry inside the semiconductor packaging. Removing the heat within these integrated circuits is becoming

more of a design challenge, and the miniaturisation of cooling devices can be used to solve these problems (Baliga, 2005). Historically, the motivation for using thermoelectric technology to cool microelectronic integrated circuits in the computer industry has been to increase their clock speed below ambient temperatures. Increasing microprocessor performance has usually been accompanied by an increase in power and on-chip power density. Both of these present a challenge in cooling microelectric devices (Mahajan et al, 2006). The computer industry may begin to approach the limit of forced-air cooled systems and will need to find alternative solutions (Sharp et al, 2006).

Localised areas of high heat flux on microprocessors can produce 'hot spots' that limit their reliability and performance, and are becoming more severe as local power density and overall die power consumption increase. Although a macro-thermoelectric module can be used in this application to provide cooling of the entire integrated circuit, micro-thermoelectric cooling of these localised regions of higher temperature or 'hot spots' may provide a better alternative. According to (Snyder et al, 2006), embedded thin-film micro-thermoelectric devices is a promising approach to reduce the temperature of localised, high heat flux hot spots generated by modern microprocessors. Micro-thermoelectric devices are also suitable for addressing other thermal management problems in microelectronics, and could be used to cool or stabilise the temperature of laser diodes, and provide a faster response time than conventional cooling techniques. It may also be possible to integrate a micro-thermoelectric device inside the laser diode packaging (Baliga, 2005). Infra-red detectors, charge coupled devices (CCD), light-emitting diodes (LED) and other opto-electronic devices may also benefit from micro-thermoelectric cooling.

Thermoelectric micro-power generation and energy harvesting is also a target market for micro-thermoelectric devices. (Bottner et al, 2007) believes that self-powered electronic sensor systems will require MEMS like manufacturing of micro-thermoelectric devices to meet the high volume requirements of this market. Energy harvesting or scavenging systems can be designed to replace batteries in autonomous sensor and wireless systems, and it has been shown that body heat can be used as an energy source to power low-energy devices, including a wrist watch or hearing-aid (Weber at al, 2006). Micro-thermoelectric power generators could also be used to supply power to electronic devices for wearable electronics applications (Bottner, 2002).

4. Conclusion

Thermoelectric technology can be used in cooling, heating and micro-power generation applications. Macro-thermoelectric devices have developed significantly since their introduction in the 1960's, and have achieved commercial success in mini-refrigeration, cooling and space-craft power applications. There is a requirement to reduce the size, and improve the performance, of current thermoelectric devices in order to address the need to solve thermal problems in microelectronics, and create localised low-power energy sources for electronic systems.

The miniaturisation and development of MEMS based thermoelectric devices has the potential to improve the performance of thermoelectric devices, and create new applications for the technology. Thermoelectric MEMS based devices, based on thin-film technology, that are compatible with modern semiconductor processing techniques have now started to enter the market place. Thermoelectric devices based on a MEMS like process that use a sputtering deposition method and Bi_2Te_3 related materials, and thermoelectric devices manufactured using Bi_2Te_3-Sb_2Te_3 superlattice technology are two recent entries into the thermoelectric market place.

It is anticipated that MEMS based thermoelectric devices can address the need to solve thermal problems in microelectronics, including the cooling of integrated circuits in the computer industry, and the cooling of optoelectronic and telecommunication devices. Micro-thermoelectric power generation is also expected to supply low-level localised power to other electronic components and systems, and provide a power source for energy harvesting systems.

5. References

Alaoui, C.; Salameh, Z.M. (2001). Solid State Heater Cooler: Design and Evaluation, *Proceedings LESCOPE*, pp. 139-145, 2001

Baliga, J. (2005). Thermoelectric Cooling Prepares for the Small Stage. *Semiconductor International*, October 2005, pp. 42

Bass, J.C.; Allen, D.T. ; Ghamaty, S. ; Elsner, N.B. (2004). New Technology for Thermoelectric Cooling, *Proceedings of 20th IEEE Semiconductor Thermal Measurement and Management Symposium*, pp. 18-20, ISBN 0-7803-8363-X, March 2004

Bottner, H.; Nurnus, J.; Schubert, A.; Volkert, F. (2007). New high density micro structured thermogenerators for stand alone sensor systems, *Proceedings of 25th International Conference on Thermoelectrics (ICT2007)*, pp. 306-309, ISBN 978-1-4244-2262-3, Jeju Island, June 2007

Bottner, H. (2005). Micropelt Miniaturized Thermoelectric Devices: Small Size, High Cooling Power Densities, Short Response Time, *Proceedings of 24th International Conference on Thermoelectrics (ICT2005)*, pp. 1-8, ISBN 0-7803-9552-2, June 2005

Bottner, H.; Nurnus, J.; Gavrikov, A.; Kuhner, G.; Jagle, M.; Kunzel, C.; Eberhard, D.; Plescher, G.; Schubert, A.; Schlereth, K. (2004). New Thermoelectric Components Using Microsystem Technologies. *Journal of Microelectromechanical Systems*, Vol. 13, No. 3, June 2004, pp. 414-420, ISSN 1057-7157

Bottner, H. (2002). Thermoelectric Micro Devices: Current State, Recent Developments and Future Aspects for Technological Progress and Applications, *Proceedings of 22nd International Conference on Thermoelectrics (ICT2002)*, pp. 511-518, ISBN 0-7803-7683-8, August 2002

Hachiuma, H.; Fukuda, K. (2007). Activities and Future Vision of Komatsu Thermo modules. *Proceedings of ECT2007*, Available: http//ect2007.thermion-company.com/proc.contents

Huang, I.; Lin, J.; She, K.; Li, M.; Chen, J.; Kuo, J. (2008). Development of low-cost micro-thermoelectric coolers utilizing MEMS technology, *Sensors and Actuators A*, Vol. 148, No. 1, November 2008, pp. 176-185

Huang, I.; Li, M.; Chen, K.; Zeng, G.; She, K. (2007). Design and Fabrication of a Column-type Microthermoelectric Cooler with Bismuth Telluride and Antimony Telluride Pillars by Using Electroplating and MEMS Technology, *Proceedings of 2nd IEEE International Conf. on Nano/Micro Engineered and Molecular Systems*, pp. 749-752, ISBN 1-4244-0610-2, Bangkok, Thailand, January 2007

Huesgen, T.; Woias, P.; Kockmann, N. (2008). Design and fabrication of MEMS thermoelectric generators with high temperature efficiency. *Sensors and Actuators A*, Vol. 145-146, 2008, pp. 423-429

Hwang, G.S.; Gross, A.J.; Kim, H.; Lee, S.W.; Ghafouri, N.; Huang, B.L.; Lawrence, C.; Uher, C.; Najafi, K.; Kaviany, M. (2008). Micro thermoelectric cooler: Planar multistage. *International Journal of Heat and Mass Transfer*, Vol. 52, No. 7-8, March 2009, pp. 1843-1852

Koester, D.; Venkatasubramanian, R.; Conner, B.; Snyder, G.J. (2006). Embedded Thermoelectric Coolers For Semiconductor Hot Spot Cooling, *Proceedings 10th Conference on Thermal and Thermomechanical Phenomena in Electronics Systems*, pp. 491-496, ISBN 0-7803-9524-7, San Diego, CA, June 2006

Li, J.; Tanaka, S.; Umeki, T.; Sugimoto, S.; Esashi, M.; Watanabe, R. (2003). Microfabrication of thermoelectric materials by silicon molding process, *Sensors and Actuators A*, Vol. 108, No. 1-3, November 2003, pp. 97-102

Liu, C.K.; Dai, M.; Yu, C.; Kuo, S. (2007). High Efficiency Silicon-based High Power LED Package Integrated with Micro-thermoelectric Device, *Proceedings IEEE International Conference Microsystems, Packaging, Assembly and Circuits Technology (IMPACT2007)*, pp. 29-33, ISBN 978-1-4244-1636-3, Taipei, October 2007

Mahajan, R. ; Chiu, C. ; Chrysler, G. (2006), Cooling a Microprocessor Chip, *Proceedings of the IEEE*, Vol. 94, No. 8, August 2006, pp. 1476-1486, ISSN 0018-9219

Min, G. (2006). Thermoelectric Module Design Theories, In: *Thermoelectrics Handbook – Macro to Nano*, D.M. Rowe (Ed.), pp. 11-1 – 11-15, CRC Taylor & Francis Group, ISBN 0-8493-2264-2, Boca Raton, Florida

Nolas, G.S.; Sharp, J. ; Goldsmid, H.J. (2001). *Thermoelectrics – Basic Principles and New Materials Developments*, Springer-Verlag, ISBN 3-540-41245-X

Nuwayhid, R.Y.; Shihadeh, A.; Ghaddar, N. (2005). Development and testing of a domestic woodstove thermoelectric generator with natural convection cooling. *Energy Conversion and Management*, Vol. 46, No. 9-10, June 2005, pp. 1631-1643

Riffat, S.B.; Ma, X. (2003). Thermoelectrics: a review of present and potential applications, *Applied Thermal Engineering*, Vol. 23, No. 8, June 2003, pp. 913-935

Rowe, D.M. (2006). General Principles and Basic Considerations, In: *Thermoelectrics Handbook – Macro to Nano*, D.M. Rowe (Ed.), pp. 1–14, CRC Taylor & Francis Group, ISBN 0-8493-2264-2, Boca Raton, Florida

Sales, B.C. (2007). Critical overview of Recent Approaches to Improved Thermoelectric Materials, *Int. Journal of Applied Ceramic Technology*, Vol. 4, No. 4, August 2007, pp. 291-296

Sharp, J.; Bierschenk, J.; Lyon, Jr. H.B. (2006), Overview of Solid-State Thermoelectric Refrigerators and Possible Applications to On-Chip Thermal Management, *Proceedings of the IEEE*, Vol. 94, No. 8, August 2006, pp. 1602-1612, ISSN 0018-9219

Snyder, G.J.; Soto, M.; Alley, R. ; Koester, D. ; Conner, B. (2006). Hot Spot cooling using embedded thermoelectric coolers, *Proceedings 22nd IEEE Semiconductor Thermal Measurement and Management Symposium*, pp. 135-143, ISBN 1-4244-0153-4, Dallas, TX, March 2006

Stabler, F.R. (2006). Commercialization of Thermoelectric Technology, *Proceedings of Materials Research Society Symposium*, pp. 13-21, 2006

Venkatasubramanian, R.; Watkins, C.; Stokes, D.; Posthill, J.; Caylor, C. (2007). Energy Harvesting for Electronics with Thermoelectric Devices using Nanoscale Materials, *Proceedings IEEE International Electron Devices Meeting (IEDM2007), pp. 367-370, ISBN 978-1-4244-1507-6, Washington DC, December 2007*

Vining, C.B. (2007). ZT ~ 3.5: Fifteen years of progress and things to come. (2007), *Proceedings of European Conference on Thermoelectrics (ECT2007)*, Odessa, Ukraine, September 2007

Volklein, F.; Meier, A. (2006). Thermoelectric Micromechanical Systems, In: *Thermoelectrics Handbook – Macro to Nano*, D.M. Rowe (Ed.), pp. 47-1 - 47–16, CRC Taylor & Francis Group, ISBN 0-8493-2264-2, Boca Raton, Florida

Weber, J.; Potje-Kamloth, K.; Haase, F.; Detemple, F.; Volklein, F.; Doll, T. (2006). Coin-sized coiled-up polymer foil thermoelectric power generator for wearable electronics. *Sensors and Actuators A*, Vol. 132, 2006, pp. 325-330

Micro Power Generation from Micro Fuel Cell Combined with Micro Methanol Reformer

Taegyu Kim
Chosun University
Republic of Korea

1. Introduction

1.1 Background

Thanks to the breakthroughs in microfabrication technologies, numerous concepts of microsystems including micro aerial vehicles, microbots, and nanosatellites have been proposed. Contrary to ordinary electronic devices, these microsystems perform mechanical work and require the extended operation. As their functions are getting complex and advanced, their energy consumption is also increasing exponentially. In order to activate these microsystems, a high density power source in a small scale is required. However, present portable devices still extract power from existing batteries. The energy density of the current batteries is too low to support these microsystems (Holladay et al., 2004). Therefore, a new micro power source is essential for the successful development of new microsystems.

Various concepts for micro power generations have been introduced such as micro engines, micro gas turbines, thermoelectric generators combined with a micro combustor, and micro fuel cells. All of these concepts extract energy from a chemical fuel that have energy density much greater than that of the existing batteries. The first challenge to micro power source was the miniaturization of conventional heat engines. However, the development of micro heat engine reached a deadlock due to the difficulties of microfabrication and realization of miniaturized fast moving components and kinematics' mechanism to generate power in micro scale. Micro fuel cells have drawn attention as a primary candidate for a micro power source due to its distinctive merits that are absence of moving parts and high efficiencies. The fuel cell is an electrochemical device that directly converts chemical energy to electric energy. Due to its different energy conversion path, the fuel cell has high thermal efficiency compared to the heat engines. The energy density of the fuel cell is higher than that of the existing batteries because it uses a chemical fuel such as hydrogen (Nguyen & Chan, 2006).

There are several types of fuel cell as summarized in Table 1 (O'Hayre et al., 2006), such as polymer electrolyte membrane fuel cell (PEMFC), phosphoric acid fuel cell (PAFC), alkaline fuel cell (AFC), molten carbon fuel cell (MCFC), and solid oxide fuel cell (SOFC). Of these fuel cells, PEMFC is suitable to a micro power device due to its low operating temperature and solid phase of electrolyte. Direct methanol fuel cell (DMFC) is a kind of PEMFC except that it directly uses methanol instead of hydrogen as a fuel. Formic acid, chemical hydrides, and other alcohols can be used as a direct fuel.

	PEMFC	PAFC	AFC	MCFC	SOFC
Electrolyte	Polymer	H_3PO_4	KOH	Molten carbonate	Ceramic
Charge carrier	H^+	H^+	OH^-	CO_3^{2-}	O^{2-}
Temperature	80 ˚C	200 ˚C	60-220 ˚C	650 ˚C	600-1000 ˚C
Catalyst	Platinum	Platinum	Platinum	Nickel	Perovskite
Cell components	Carbon	Carbon	Carbon	Stainless	Ceramic
Fuel compatibility	H_2, CH_3OH	H_2	H_2	H_2, CH_4	H_2, CH_4, CO

Table 1. Descriptions of major fuel cell types

In the beginning of research, DMFC has been widely investigated as a possible candidate for micro power generation due to the use of liquid fuel and its simple structure (Lua et al., 2004). However, the fuel crossover phenomena is an inherent problem of DMFC, which severely limits its power output. It is known that the power output of PEMFC is much greater than that of DMFC, and there is no fuel crossover in PEMFC. Major obstacle in the successful development of PEMFC is the difficulties of the hydrogen storage with high density. Although possible to use hydrogen in either compressed gas or liquid form, it gives significant hazards due to its explosive nature. Metal hydride suffers from high weight per unit hydrogen storage and low response for a sudden increase in hydrogen demand. Chemical storage in the form of liquid fuel such as methanol has significantly higher energy density compared to the suggested technologies. It can be reformed to generate hydrogen gas when needed. The fuel reformer is a device that extract hydrogen from a chemical fuel including methanol, methane, propane, octane, gasoline, diesel, kerosene, and so on. The fuel choice is more flexible than the direct fuel cells. Although a fuel cell combined with the reformer is more attractive, it is complex and bulky compared to the DMFC due to the fuel reformer. Therefore, the miniaturization of the reformer has been a major research activity for the successful development of PEMFC system in recent years (Pattekar & Kothare, 2004). MEMS technology is a useful tool to reduce the size of reformer and fuel cell (Yamazaki, 2004). The use of MEMS technology in a thermo-chemical system is relatively new concept. It allows the miniaturization of conventional reactors while keeping its throughput and yield. The microreactor has a relatively large specific surface area, which provides the increased rate of heat and mass transport, and short response time. In addition, MEMS-compatible materials are suitable to various chemical reaction applications due to their high thermal and chemical resistances.

1.2 Literature survey

Catalytic steam reforming of methanol for hydrogen production using conventional reactors has been already carried out in the literature. However, the use of microreactors is a relatively new challenge and other approaches are required for the development of micro reformers using MEMS technologies. Nevertheless, the study on the methanol reforming reaction in the conventional reactors give a good background for the development of micro methanol reformer.

Various research groups have successfully developed micro fuel reformers using MEMS technologies. Pattekar & Kothare, 2004 developed a micro-packed bed microreactor for hydrogen production, which is fabricated by deep reactive ion etching (DRIE). The width of

microchannels was 1000 μm and the depth ranged from 200 to 400 μm. The microchannels were grooved on 1000 μm thick silicon substrate using photolithography followed by DRIE. A 10 μm thick photoresist (Shipley 1045, single/dual coat) was used as a etch mask for DRIE. Commercial $Cu/ZnO/Al_2O_3$ catalyst was load by passing the water-based suspension of catalyst particles ranging from 50 to 70 μm via microchannels. The microfilter was fabricated at the end of microchannels, and the catalyst particles larger than 20 μm were trapped in the microchannels. The platinum resistance temperature detector was used as a temperature sensor with a linear temperature versus resistance characteristic. The platinum microheater was deposited along the microchannels. The methanol conversion was 88% at the steam-to-carbon ratio (S/C) of 1.5 and the methanol feed rate of 5 ml/h. The hydrogen production rate was 0.1794 mol/h that is the sufficient flow to generate 9.48 W in a typical PEMFC. Pattekar & Kothare, 2005 also developed a radial flow reactor that has less pressure drop compared to conventional one due to the increased flow cross section area along the reaction path.

Kundu et al., 2006 fabricated a microchannel reformer on a silicon wafer using silicon DRIE process. The split type channels were made in the micro vaporizer region to reduce the back pressure at the inlet port and to get a more uniform flow of fluid. The dimensions of the micro reformer were 30 mm in length and 30 mm in width, and each channel was 28 mm in length. The width of each channel was 1 mm and the depth was 300 μm. The commercial $CuO/ZnO/Al_2O_3$ catalyst (Johnson Matthey) was packed inside the channels by injecting the water-based catalyst slurry. The catalyst particles were trapped in the microchannels by filters that were in the form of 90 μm thick parallel walls spaced 10 μm apart oriented parallel to the direction of the fluid flow. The catalyst deactivation was observed after operating continuously for 8 hours using catalyst characterization. It can be seen that the performance with the serpentine channel was higher than with the parallel channel due to the longer residence time. The hydrogen production rate was 0.0445 mol/h which can produce 2.4 W assuming an 80% fuel cell operation efficiency.

Kazushi et al., 2006 developed a micro fuel reformer integrated with a combustor and a microchannel evaporator. Two fuel reforming reactors were placed on either side of a combustor to make the system compact and to use combustion heat efficiently. The silicon and Pyrex® glass wafer that are used as a substrate were stacked by anodic bonding. A commercially available reforming catalyst made of $CuO/ZnO/Al_2O_3$ (MDC-3, Süd-Chemie Catalysts Japan, Inc.) was filled into a microchamber fabricated on glass substrates after being powdered and hardened by polyvinylalcohol (PVA). The Pt loaded on TiO_2 support made by sol-gel method was used as a catalyst of the combustor. Thin film resistive temperature sensors made of Pt/Ti (100 nm/50 nm) to measure temperature inside the fuel reformer was fabricated on the wall of the combustion chamber by the lift-off process. The six kinds of microchannel evaporators were fabricated on the silicon substrates; as a result, it was found that the design of the microchannel evaporator is critical to obtain larger hydrogen output. The 32.9 ml/min of hydrogen, which is equivalent to 5.9 W in lower heating value, was produced when input combustion power was 11 W. The maximum efficiency of 36.3% was obtained and the power density of the reformer was 2.1 W/cm^3.

Though the work on the MEMS-based reformer has been continuously reported in the recent literature, there is no novel change and significant improvement. The literature could be classified into two standpoints. In terms of substrate materials, silicon wafers has been mostly used as a substrate of microreactors. Different materials have been also used such as

glass wafer, polydimethylsiloxane (PDMS), and low temperature co-fired ceramic (LTCC). In terms of a method of catalyst loading in the reactor bed, either catalyst coating or packing has been used. In almost results, the heat to sustain the methanol steam reforming reaction was provided by an external heater, while some results presented the use of a catalytic combustor as a heat source.

1.3 Fuel reforming process and system

Fuel reforming is a chemical process that extracts hydrogen from a liquid fuel. Fuel reformer is a device that produces hydrogen from the reforming reaction. Liquid fuel is used as a feed of the reformer due to its higher density than gaseous fuels. Considering hydrogen content and ease of reforming, methanol was chosen as the primary fuel in hydrogen sources such as alcohols and hydrocarbons (Schuessler et al., 2003).

There are a number of fuel reforming techniques available, including steam reforming (Lindström & Pettersson, 2001), partial oxidation (Wang et al., 2003), and autothermal reforming (Lindström et al., 2003). Of all considered techniques, the steam reforming process provides the highest attainable hydrogen concentration in the reformate gas. This reaction takes place at relatively low temperature in the range of 200-300 ˚C. The chemical reaction of the methanol steam reforming process is expressed below:

$$CH_3OH + H_2O \rightarrow 3H_2 + CO_2 \tag{1}$$

Equation 1 is a primary reforming process that is the stoichiometric conversion of methanol to hydrogen. It can be regarded as the overall reaction of the methanol decomposition and the water-gas shift reaction. First, the methanol decomposes to generate carbon monoxide.

$$CH_3OH \rightarrow 2H_2 + CO \tag{2}$$
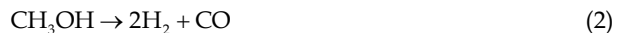
The presence of water can convert carbon monoxide to carbon dioxide through the water-gas shift reaction.

$$CO + H_2O \rightarrow H_2 + CO_2 \tag{3}$$

The formation of carbon monoxide lowers the hydrogen production rate and the carbon monoxide also acts as a poison for the fuel cell catalyst. Typically, carbon monoxide is converted to carbon dioxide either in a separate water-gas shift reactor or a preferential oxidation called PROX (Delsman et al., 2004). Palladium/silver alloy membrane is also used to separate selectively the carbon monoxide. Other byproducts such as carbon dioxide and excess water vapor can be safely discharged to atmosphere.

Cu-based catalysts are used for the steam reforming of methanol, and the well-known one is $Cu/ZnO/Al_2O_3$. Generally, it has been claimed that Cu^0 provides catalytic activity and ZnO acts as a stabilizer of Cu surface area. Addition of Al_2O_3 to the binary mixture enhances Cu dispersion and catalyst stability (Agrell et al., 2003).

The steam reforming of methanol is endothermic reaction. An external electric heaters or catalytic combustors can be used as a heat sources to sustain the reforming reaction. The amount of the endothermic heat per a mole of methanol is 48.96 kJ/mol at 298 K. The electric microheater is the simplest method to supply heat to the reformer because its control is relatively easy and the fabrication can be simply integrated into MEMS process. However, the electric heater is usually used for startup period only due to its low thermal efficiency.

The catalytic combustors are an ideal alternative heat source to the electric heater because its high thermal efficiency. Methanol can be directly used in the combustor to facilitate methanol reforming reaction. Part of the hydrogen produced out of the reformer can be fed to the combustor. While it is possible that the catalytic hydrogen combustion with Pt as the catalyst even at room temperature, the methanol combustion requires preheaters to initiate the reaction. In the present study, the catalytic combustion of hydrogen and the catalytic decomposition of hydrogen peroxide were used as heat sources of the methanol steam reformer. Hydrogen peroxide as a heat source is the first attempt in the world.

Figure 1 shows the schematic of a typical reformer-combined fuel cell system, which consists of a fuel reformer and a fuel cell. The fuel reformer is classified into four units; fuel vaporizer/preheater, steam reformer, combustor/heat-exchanger, and PROX reactor. First, methanol is fed with water and is heated by the vaporizer. The methanol is reformed by the reforming catalyst to generate hydrogen in the steam reformer. To supply heat to the steam reformer, part of hydrogen from the anode off-gas of fuel cell can be fed to the combustor. The combustor generates the sufficient amount of heat to sustain the methanol reforming reaction. As mentioned before, the extremely small amount of carbon monoxide deactivates the fuel cell catalyst, which should be reduced to below 10 ppm by PROX.
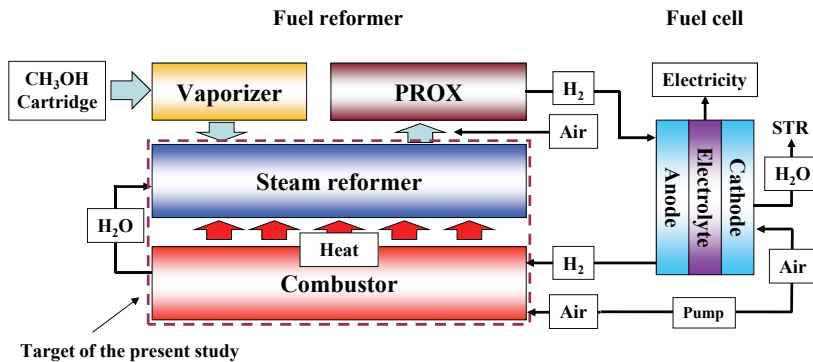
Fig. 1. Schematic of the fuel cell system combined with the fuel reformer

1.4 Outline of chapter

This chapter presents design, fabrication and evaluation of MEMS methanol reformer. First, a methanol reformer was fabricated and integrated with a catalytic combustor. Cu/ZnO was selected as a catalyst for the methanol steam reforming reaction and Pt for the hydrogen catalytic combustion. Wet impregnation method was used to load the catalysts on a porous support. The catalyst-loaded supports were inserted in the cavity made on the glass wafer. The performance of the micro methanol reformer was measured at various test conditions and the optimum operation condition was sought. Next, new concept of micro methanol reformer was proposed in the present study. The micro reformer consists of the methanol reforming reactor, the catalytic decomposition reactor of hydrogen peroxide, and a heat-exchanger between the two reactors. In this system, the catalytic decomposition of hydrogen peroxide is used as a process to supply heat to the reforming reactor. The decomposition process of hydrogen peroxide produces water vapor and oxygen as a product, which can be used efficiently to operate the reformer/PEMFC system. Microreactor was fabricated for

preferential oxidation of carbon monoxide using a photosensitive glass process integrated with a catalyst coating process. A γ-Al_2O_3 layer was coated as a catalyst support on the surface of microchannels using sol-gel method. The wet impregnation method was used to load Pt/Ru in the support. The conversion of carbon monoxide was measured with varying the ratio of oxygen to carbon (O_2/C) and the catalyst loading amount. Micro fuel cell was fabricated and the integrated test with the MEMS methanol reformer was performed to validate the micro power generation from the micro fuel cell system.

2. Micro reformer integrated with catalytic combustor

2.1 Design

Figure 2 depicts the construction of the integrated micro methanol reformer. The mixture of methanol and water enters the steam reformer at the top and the reformate gas leaves the reactor. The mixture of hydrogen and air flows into the catalytic combustor at the bottom with counter flow stream against the reforming stream. The heat generated from the catalytic combustor is transferred to the steam reformer through the heat-exchanger layer that has micro-fins to increase the surface area and the suspended membrane to enhance the heat transfer rate. The porous catalyst supports were inserted in the cavity made on the glass wafer as shown in Fig. 2. The micro reformer structure was made of five glass wafers; two for top and bottom, one for the steam reformer, one for the catalytic combustor, and the reminder for the heat-exchanger in-between.
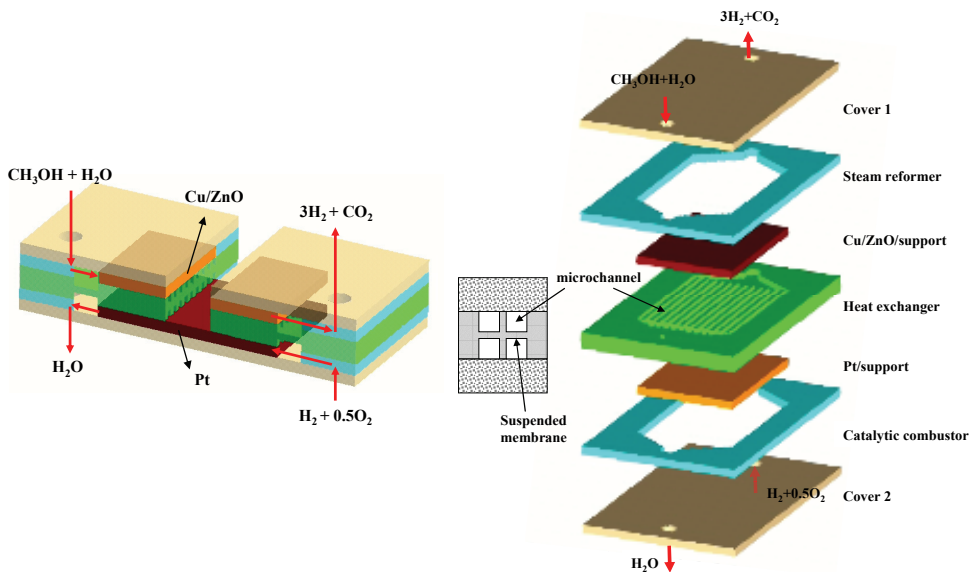
Fig. 2. Construction of the integrated micro methanol reformer

The porous ceramic material (ISOLITE®) was used as a catalyst support due to its large surface area and thermal stability (Kim et al., 2007). The typical ceramic support is composed of 40% Al_2O_3 and 55% SiO_2 with traces of the other metal oxides, and the porosity is approximately 71%. Figure 3 shows SEM images of the support material. The scale of the

bulk pores was between 100 and 300 µm, while smaller scale pores were a few microns. This structure of the porous support can enhance the heat and mass transport between catalyst active sites and reactants.
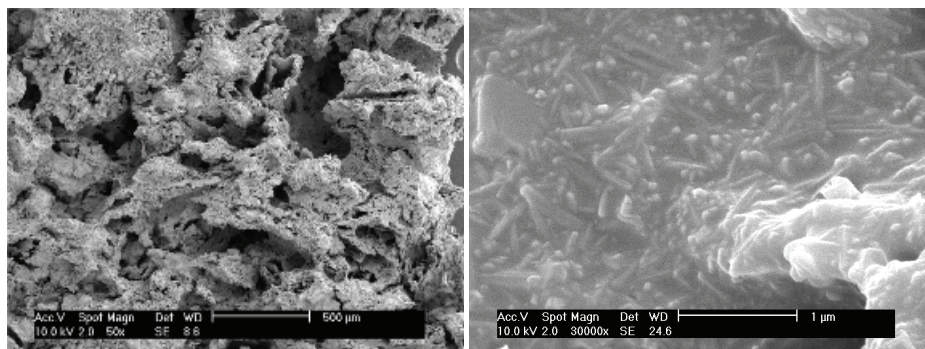
Fig. 3. SEM images of the porous ceramic material used as a catalyst support

2.2 Fabrication

The overall fabrication process was integrated with a catalyst loading step as shown in Fig. 4. The fabrication process for an individual glass wafer is as follows: (1) exposure to ultraviolet (UV) light under a mask at the intensity of 2 J/cm^2; (2) heat treatment at 585 ˚C for 1 hour to crystallize portion of the glass that was exposed to UV; and (3) etching the crystallized portion of the glass in the 10% hydrofluoric (HF) solution to result in the desired shape. The etching rate was 1 mm per hour. With step 1-3 in Fig. 4, two covers, a reformer layer, and a combustor layer were fabricated. To obtain the membrane heat-exchanger, the glass wafer was exposed by UV light on both sides of the wafer. After the heat treatment, the wafer was etched standing in the etching bath. The tooth shape cross-section of the membrane heat-exchanger layer was fabricated by controlling etching time as shown in the step 4-6 of Fig. 4. The complete micro methanol reformer was constructed by fusion-bonding the fabricated glass layers, where the porous catalyst supports were inserted in the reformer layer and the combustor layer, respectively. The best fusion-bonding between glass wafers was obtained by pressing the wafers against each other at 1000 N/m^2 in a furnace held at 500 ˚C (Kim & Kwon, 2006a).

As a final step, the catalysts were loaded on the porous catalyst supports. The Cu/ZnO was selected as a catalyst for methanol reforming reaction, considering its proven reactivity and selectivity (Kim & Kwon, 2006b). The Pt was chosen as a catalyst for the hydrogen catalytic combustion. The wet impregnation method was used to load both catalysts on the porous supports. A mixture of a 0.7 M aqueous solution of $Cu(NO_3)_2$ and a 0.3 M aqueous solution of $Zn(NO_3)_2$ was prepared. The mixture was injected in the catalyst support inserted in the reformer layer using a syringe pump. The moisture was removed by drying the catalyst-loaded support in a convection oven at 70 ˚C for 12 hours. Calcination procedure followed in a furnace at 350 ˚C for 3 hours. The similar procedures were used for Pt coating with 1 M aqueous solution of H_2PtCl_6. The amount of the loaded Cu/ZnO was 7.0 wt % while Pt was 5.0 wt % of the total weight of the catalyst support. The catalysts were reduced for 4 hours in an environment of mixture of 4% H_2 in N_2, which is steadily flowing into the reformer at a rate of 10 ml/min in a furnace of 280 ˚C.

Figure 5 shows the fabrication results, including etched glass wafers, a complete micro methanol reformer, a cross-section view of the reformer and SEM image of the membrane heat-exchanger. The total volume of the reformer was 3.6 cm^3 (20 mm×30 mm×6mm) and the weight was approximately 13.4 g.
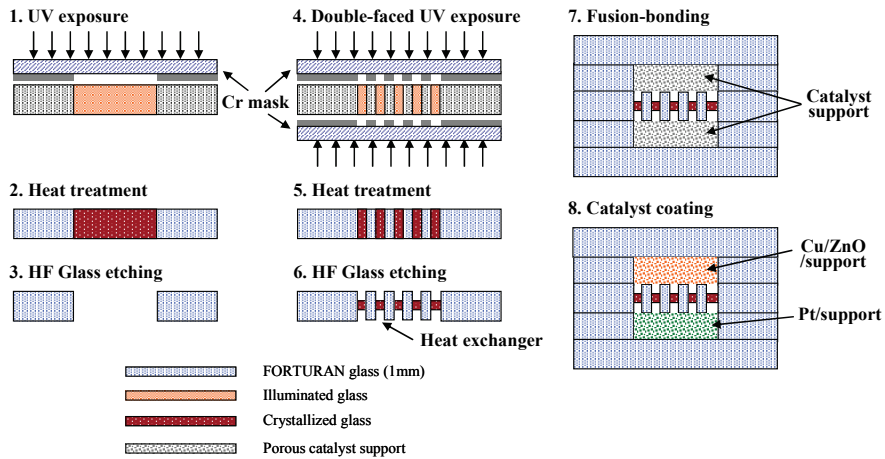
Fig. 4. Overall fabrication procedure of the micro methanol reformer.
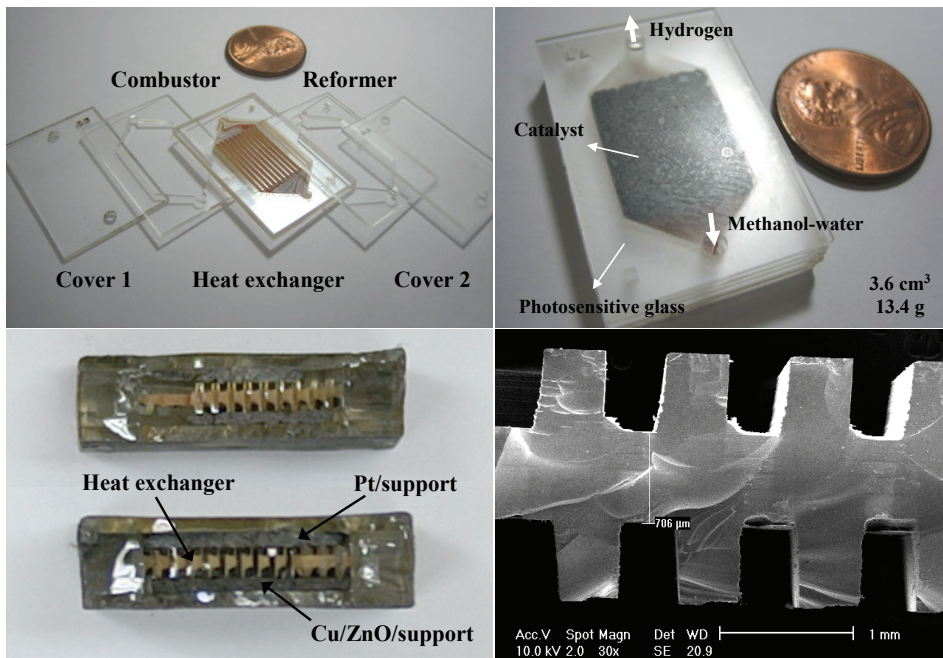
Fig. 5. Fabricated results of the micro methanol reformer

2.3 Performance measurement

Experimental setup was equipped to measure the performance of the micro methanol reformer. A syringe pump (KDS200, KD Scientific) supplied a mixture of methanol and water to the reformer at a controlled rate. The flow rate of hydrogen and air was controlled by mass flow controllers (EL-FLOW, Bronkhorst). After mixed them in a mixing chamber, the mixture gas was supplied to the combustor. The temperature of each reactor was recorded by thermocouples. The product gas of the reformer was cooled and the condensable portion was removed in a cold trap. The non-condensable product gas was analyzed by a gas chromatography (Agilent HP6890). The flow rate of dry gas was measured by a bubble meter. The column in the gas chromatography was Carboxen-1000 (60/80 mesh, 1/8", 18 ft) that can separate H_2, N_2, CO, CO_2, CH_4 and others. Nitrogen carrier gas at known flow rate was mixed with the product gases before entering the gas chromatography. The exact hydrogen production rate can be calculated by comparing the ratio of hydrogen to nitrogen because the flow rate of the carrier gas is known. The gas composition was detected by a TCD (thermal conductivity detector) with Ar as a reference gas. The product gas of the catalytic combustor was analyzed, after moisture was removed in a cold trap.

The energy balance between the methanol reformer and the catalytic combustor was calculated as shown in Table 2. The total heating energy consists of the energy to raise the reformer temperature and the heat of reaction. The heat of reaction is the sum of the reforming heat, the evaporation heat and the heat to raise mixture to reforming temperature (sensible heating). The energy to reform 1 mole methanol with 1 mole water is 158.3 kJ, which can be provided by burning 0.66 mole hydrogen by the catalytic combustor. The hydrogen can be provided by recycling the off-gas of the fuel cell. The reformer produces 2.7 moles hydrogen from 1 mole methanol when methanol conversion is 95% and hydrogen selectivity is 95%. Assuming that hydrogen utilization of the fuel cell is 72%, the amount of the hydrogen off-gas is 0.756 mole, which is greater than the hydrogen requiremnt for the combustor to sustain the reformer. Based on this calculation, the expected production of hydrogen is 54.5 ml/min when the methanol feed rate is 2 ml/h. The fuel cell consumes 72% portion (39.2 ml/min) in the reformed hydrogen and the remainder (15.3 ml/min) can be used to operate catalytic combustor.

	Calculation	Flow rate
Methanol input	1 mol	2 ml/h
Energy requirement for the reformer[*]	153.8 kJ	
Evaporation and sensible heating of methanol	48.4 kJ	
Evaporation and sensible heating of water	51.5 kJ	
Heat of reaction	58.4 kJ	
Expected production of hydrogen[**]	2.7 mol	54.5 ml/min
Hydrogen requirement for the combustor	0.66 mol	13.3 ml/min
Anode off-gas of fuel cell[***]	0.756 mol	15.3 ml/min

[*]Reforming temperature: 250 ˚C, [**]95% methanol conversion, 95% hydrogen selectivity, [***]Fuel cell utilization: 72%

Table 2. Energy balance calculation between the methanol reformer and the combustor

2.4 Results and discussion

The performance of the reformer was measured at various test conditions and an optimum operation condition was sought. The measured performance of the reformer was expressed in terms of the methanol conversion, which is defined as follows:

$$CH_3OH \text{ conversion [mol\%]} = \frac{\text{mol } (CH_3OH)_{in} - \text{mol } (CH_3OH)_{out}}{\text{mol } (CH_3OH)_{in}} \times 100 \qquad (4)$$

Figure 6 shows the methanol conversion as a function of the reformer temperature at each methanol feed rate with the steam-to-carbon ratio of 1.1. The methanol conversion decreased as the methanol feed rate increased, while the methanol conversion increased as the reformer temperature increased. The maximum methanol feed rate was 2 ml/h to obtain the methanol conversion higher than 90% at temperature lower than 250 ˚C. At the feed rate of 2 ml/h and the reformer temperature of 250 ˚C, the hydrogen production rate was 53.9 ml/min and the composition of carbon monoxide in the reformate gas was 0.49%.
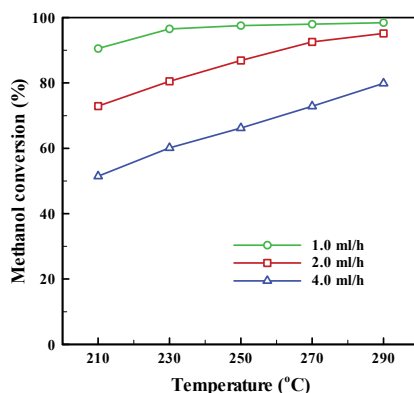
Fig. 6. Methanol conversion as a function of the reformer temperature

The performance of the catalytic combustor was measured at various conditions. Figure 7 shows the temperature variation of the catalytic combustor as a function of the reaction time at an equivalence ratio of 1.0. This plot includes the change of reformer temperature, which has to reach 250 ˚C to obtain the optimal methanol conversion. The temperatures of reformer and catalytic combustor were measured as varying the hydrogen feed rate. The air was mixed with hydrogen in the mixing chamber at the equivalent ratio of 1.0 and the gas mixture was fed into the combustor. In the energy balance calculation, the hydrogen requirement of the combustor was 15.3 ml/min to sustain the methanol reforming reaction at the methanol feed rate of 2 ml/h. At the feed rate of 15.3 ml/min, the temperature of the catalytic combustor reached 148.7 ˚C when 18 min elapsed after the initiation of the reaction. The hydrogen feed rate increased to reduce the time for the startup of the reformer. At the hydrogen feed rate of 41.3 ml/min, the combustor temperature reached 271 ˚C within 8.6 min after the start of operation and the reformer temperature was 250 ˚C. As the hydrogen feed rate increased, the combustion heat increased and the time for startup decreased. However, the hydrogen conversion decreased at the increase of the hydrogen feed rate due to the short residence time that is proportional to the inverse of the feed rate. Furthermore,

the hot-spot appeared in the fore part of the combustor, which can damage the catalyst and the reactor substrate. The temperature difference between the reformer and the combustor increased with the hydrogen feed rate. At the feed rate of 41.3 ml/min, the temperature difference was 21 ˚C when the reformer temperature reached 250 ˚C.
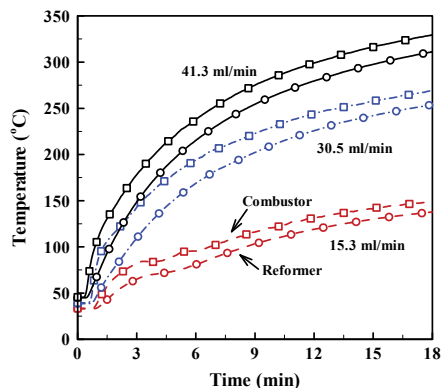
Fig. 7. Temperature variation of the catalytic combustor as a function of the reaction time.

Figure 8 represents the result of simultaneous operation of the methanol steam reformer and the catalytic combustor. The reformer was heated up to 250 ˚C by an external preheater with the increasing rate of temperature of 11.4 ˚C/min. The combustor was operated when the reformer temperature reached 250 ˚C. The hydrogen feed rate was 15.3 ml/min, which can be supplied from the anode off-gas of fuel cell when the methanol feed rate is 2 ml/h. The air was mixed with hydrogen to fix the equivalent ratio at 1.0. The methanol was fed into the reformer with the feed rate of 2 ml/h. The water feed rate was 0.98 ml/h to satisfy the steam-to-carbon ratio of 1.1. The reformer temperature was maintained constantly after the methanol reforming reaction was initiated. After 8 minutes into the simultaneous operation, steady reforming reaction was attained and the methanol conversion was higher than 90%. The maximum conversion of methanol was 95.7%. The temperature difference between the reformer and the combustor was approximately 4 ˚C.
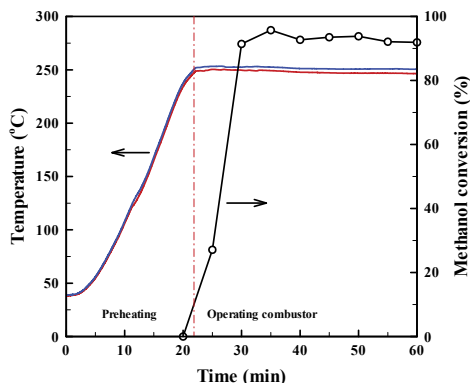
Fig. 8. Simultaneous operation of the methanol steam reformer and the catalytic combustor
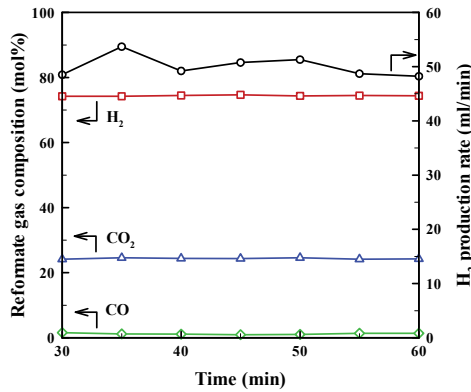
Fig. 9. The composition of reformate gas and the production rate of hydrogen

Figure 9 shows the composition of reformate gas and the hydrogen production rate after the start of complete operation. As the steady reforming reaction lasted, the composition of reformate gas remained constant. The reformate gas composition was 74.4% H_2, 24.36% CO_2, and 1.24% CO, and its flow rate was 67.2 ml/min. The hydrogen production rate was approximately 50 ml/min, which can generate 4.5 W electric power on a typical PEMFC. The concentration of carbon monoxide at the integrated test was higher than that at the separate test of the reformer. Although the catalytic combustor gave the sufficient amount of heat to operate the reformer, it could not form uniform temperature distribution within the reformer. As a result, the high temperature gradient occurred in the reformer, increasing the selectivity of carbon monoxide. The thermal efficiency of the conventional reformer combined with the combustor is defined by:

$$\eta_T = \frac{LHV_{H_2\_produced}}{LHV_{CH_3OH\_reformer} - LHV_{H_2\_combustor}} \times 100 \qquad (5)$$

where the LHV means the lower heating value. The thermal efficiency of the integrated micro methanol reformer was 76.6%. The operating conditions and the performance of the micro methanol reformer is summarized in Table 3.

Operating condition	Reformer	Combustor
Feed flow rate	2 ml/h CH_3OH	15.3 ml/min H_2
S/C (steam-to-carbon ratio)	1.1	
Equivalence ratio		1.0
Temperature	250 ˚C	251 ˚C
Performance	Reformer only	Integrated operation
Temperature		247 ˚C (reformer)
Conversion	96.2%	95.7%
H_2 production rate	53.9 ml/min	50 ml/min
CO composition	0.49%	1.24%
Thermal efficiency		76.6%

Table 3. The operating conditions and the performance of the micro methanol reformer

3. Micro reformer heated by hydrogen peroxide decomposition

3.1 Hydrogen peroxide as a heat source

In the previous section, the catalytic combustor is used as a heat source of the methanol steam reformer. However, it is still problematic that non-uniform distribution of reaction and hot spot formations in the fore region of the combustor. In the present study, the catalytic decomposition of hydrogen peroxide is used as a process to supply heat to the reformer. The decomposition reaction of hydrogen peroxide is expressed below:

$$H_2O_2 \rightarrow H_2O + 0.5O_2, \quad \Delta H_f^o = -54.24\,kJ/mol \tag{6}$$
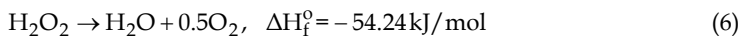
The construction of the micro methanol reformer complete with a heat source is presented in Fig. 10, in which the catalytic reactor for the hydrogen peroxide decomposition is included. The hydrogen peroxide decomposition is a highly exothermic reaction and generates the sufficient amount of heat to sustain the methanol steam reforming reaction. The catalytic decomposition of hydrogen peroxide has great reactivity and selectivity on various metal elements, such as Fe, Cu, Ni, Cr, Pt, Pd, Ir, and Mn (Teshima et al., 2004). The hydrogen peroxide decomposition generates steam and oxygen as products. The steam can be recycled into the reformer for the steam reforming reaction. The oxygen can be used as an oxidizer at the fuel cell cathode and to remove carbon monoxide in the preferential oxidation. The present concept renders the system far more compact than the existing reformer/combustor model because hydrogen peroxide is stored and used in condensed phase and oxygen enrichment enhances the system efficiency.

In the present study, the performance evaluation of the methanol steam reformer with hydrogen peroxide heat source was carried out at various test conditions and an optimum operation condition was sought.
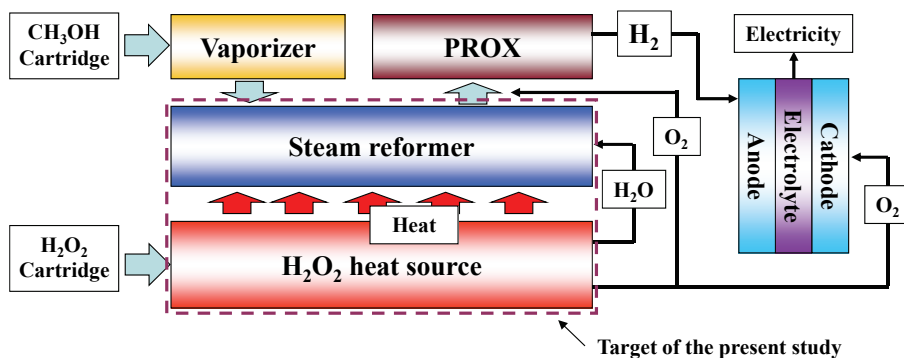
Fig. 10. Concept of methanol steam reformer integrated with hydrogen peroxide heat source

3.2 Experimental

Experimental apparatus for the performance measurement of the reformer system is similar with the combustor experiment. Two syringe pumps supplied reactants to the reactor at a controlled rate; one for the mixture of methanol and water, and the other for hydrogen peroxide. The temperature of each reactor was recorded by thermocouples. The analysis of the product gas composition was the same with the section 2.3. The concentration of

hydrogen peroxide was measured using a refractometer (PR-50HO, ATAGO) with a small quantity of sample. The product gas of hydrogen peroxide decomposition was analyzed, after moisture removed in a cold trap.

The measured performance of the reformer was expressed in terms of the methanol conversion, hydrogen selectivity and hydrogen peroxide conversion, which are defined as follows:

$$\text{CH}_3\text{OH conversion [mol\%]} = \frac{\text{mol (CH}_3\text{OH)}_{in} - \text{mol (CH}_3\text{OH)}_{out}}{\text{mol (CH}_3\text{OH)}_{in}} \times 100 \qquad (4)$$

$$\text{H}_2 \text{ selectivity [\%]} = \frac{\text{mol (H}_2) \times 1/3}{\text{mol (CH}_3\text{OH)}_{in} - \text{mol (CH}_3\text{OH)}_{out}} \times 100 \qquad (7)$$

$$\text{H}_2\text{O}_2 \text{ conversion [mol\%]} = \frac{\text{mol (H}_2\text{O}_2)_{in} - \text{mol (H}_2\text{O}_2)_{out}}{\text{mol (H}_2\text{O}_2)_{in}} \times 100 \qquad (8)$$

3.3 Operation parameter

The chemical equation of methanol steam reforming reaction is expressed below:

$$\text{CH}_3\text{OH} + s\text{H}_2\text{O} \rightarrow 3\text{H}_2 + \text{CO}_2 + (1\text{-}s)\text{H}_2\text{O} \qquad (9)$$
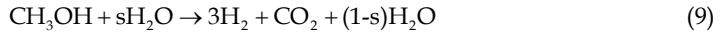
where symbol s is the molal ratio of water to methanol ($\text{H}_2\text{O}/\text{CH}_3\text{OH}$), which is the same with the steam-to-carbon ratio. Decomposition reaction of hydrogen peroxide is expressed below:

$$a(x\text{H}_2\text{O}_2 + (1\text{-}x)\text{H}_2\text{O})) \rightarrow 0.5ax\text{O}_2 + (1 - x + ax)\text{H}_2\text{O} \qquad (10)$$
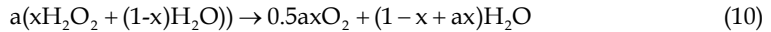
where symbol a and x are the molal ratio of hydrogen peroxide to methanol ($\text{H}_2\text{O}_2/\text{CH}_3\text{OH}$) and the molal concentration of hydrogen peroxide, respectively. The performance of the reformer system depends on these parameters. In order to determine the reaction condition, the concentration of hydrogen peroxide and the weight hourly space velocity (WHSV) were used as control parameters. The weight hourly space velocity indicates the ratio of the reactant flow rate to the catalyst mass as follows:

$$\text{WHSV} = \frac{\text{Molal flow rate of reactants (mol/h)}}{\text{Catalyst mass (g)}} \text{ [mol/g-h]} \qquad (11)$$

Overall heat output of the integrated reformer system was calculated as shown in Fig. 11. Figure 11 (a) shows the variation in the decomposition reaction heat of hydrogen peroxide as a function of the weight concentration of hydrogen peroxide. It can be seen that the hydrogen peroxide concentration has to be higher than 73.9 wt % to generate the sufficient heat to complete the reforming reaction of methanol at s = 1.0 and a = 9.0, respectively. Hydrogen peroxide with even higher concentration is needed when the steam-to-carbon ratio is higher or the hydrogen peroxide-to-methanol ratio is lower.

Figure 11 (b) illustrates the net heat output that amounts to the difference between the decomposition heat of hydrogen peroxide and the heat required to maintain the reformer at the optimum operation condition. The decomposition heat of 5.3 moles hydrogen peroxide

at 81.5 wt % concentration releases the sufficient amount of heat to reform the mixture of 1 mole methanol and 1 mole water. The required amount of hydrogen peroxide will decrease when the hydrogen peroxide concentration increases or the steam-to-carbon ratio decreases. In the calculation that leaded to Fig. 11, the heat loss to the surrounding was ignored. Considering the heat loss of the reformer, higher concentration of hydrogen peroxide or higher hydrogen peroxide-to-methanol ratio is required. In the present study, hydrogen peroxide of 82 wt % concentration was used and the steam-to-carbon ratio was fixed at 1.1 for convenience in the experiment. The performance characteristics of the reformer was investigated with three control parameters; methanol space velocity, hydrogen peroxide space velocity, and hydrogen peroxiode-to-methanol ratio.

(a)   (b)
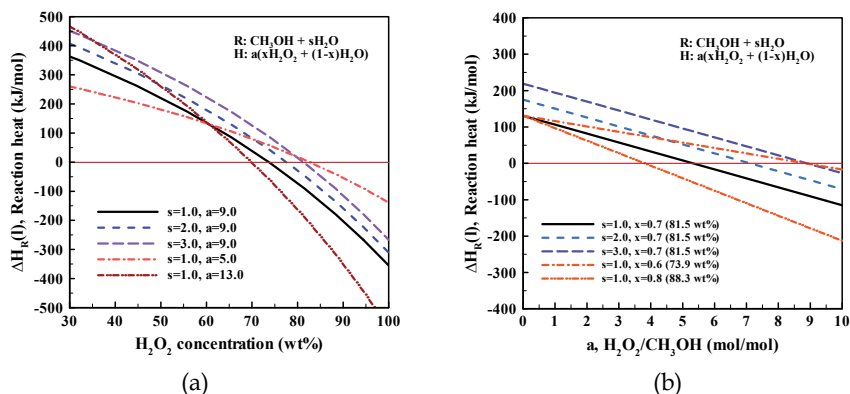
Fig. 11. Overall heat output of the integrated reformer system

3.4 Results and discussion

The temperature of the hydrogen peroxide decomposition reactor was measured as varying the hydrogen peroxide space velocity. Figure 12 (a) shows the temperature of the hydrogen peroxide decomposition reactor as a function of reaction time at each space velocity, in which the hydrogen peroxide conversion is included. At the space velocity of 6.32 mol/g-h, the hydrogen peroxide conversion was 98.2% and the reactor temperature reached 150 °C when 200 seconds elapsed after the initiation of reaction. At the space velocity of 37.3 mol/g-h, the reactor temperature reached 250 °C, which is the optimal temperature for the methanol reforming reaction, within a minute after the start of operation. The amount of reaction heat increases with the feed rate of hydrogen peroxide, reducing the time to obtain the optimal reformer temperature. At high space velocity, however, reactants does not take the residence time enough to react on the catalyst, resulting in the decrease of hydrogen peroxide conversion. At the low space velocity, the temperature difference between the reformer and the decomposition reactor was within 5 °C. At the space velocity of 37.3 mol/g-h, however, the temperature difference increased with the time after the start-up as shown in Fig. 12 (b). When the temperature of decomposition reactor reached 250 °C, the reformer temperature was less than 200 °C.

Figure 13 represents the simultaneous operation result of the methanol steam reformer and the hydrogen peroxide decomposition reactor. The reformer was heated up to 250 °C by the decomposition reactor with 82 wt% hydrogen peroxide at the space velocity of 9.48 mol/g-

h. The mixture of methanol and water was fed into the reformer with the steam-to-carbon ratio at 1.1. The space velocity of methanol was 0.68 mol/g-h. The temperature increased steadily after the methanol reforming reaction was initiated. It implies that the hydrogen peroxide feed rate exceeds the minimum to sustain the methanol reforming reaction. By reducing the feed rate down to the space velocity of 6.32 mol/g-h after 5 minutes into the operation, an ideal reaction condition was obtained as shown in Fig. 13. After 8 minutes into the operation, steady methanol reforming reaction was obtained and the methanol conversion was higher than 91.2%. The temperature inside the reformer and the decomposition reactor were 253 °C and 278 °C, respectively.

	(a)	(b)
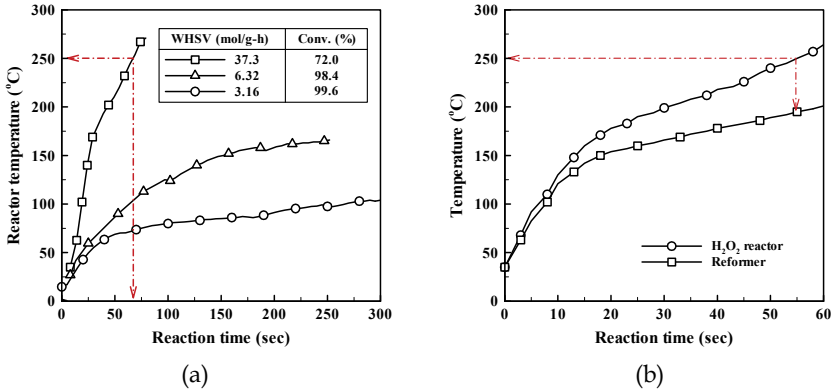
Fig. 12. The performance of hydrogen peroxide decomposition reactor
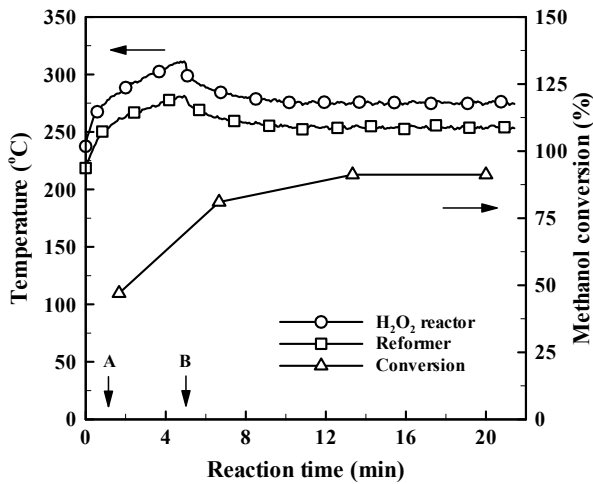
Fig. 13. Simultaneous operation of the micro reformer with hydrogen peroxide heat source

The performance characteristics of the micro reformer with hydrogen peroxide heat source was investigated at various conditions. Figure 14 (a) shows the effect of the methanol space velocity on the methanol conversion and the reformer temperature with the conditions of

the decomposition reactor fixed (S/C = 1.1, 82 wt% H_2O_2, H_2O_2 WHSV 6.32 mol/g-h). As the methanol space velocity increased, the reformer temperature decreased gradually because the hydrogen peroxide decomposition heat was consumed to vaporize the methanol supplied in liquid phase. As a result, the reformer decreased in temperature and did not sustain the methanol reforming reaction. Figure 14 (b) shows the effect of the reformer temperature on the methanol conversion. The feed rate of the methanol was fixed while the reformer temperature was determined by varying the feed rate of hydrogen peroxide (CH_3OH WHSV 0.68 mol/g-h, S/C = 1.1, 82 wt% H_2O_2). The reformer temperature increased with the space velocity of hydrogen peroxide because the decomposition heat of hydrogen peroxide increased. The methanol conversion increased with the reformer temperature, when the temperature was below 250 ˚C. For the reformer temperature higher than 250 ˚C, the methanol conversion maintained its value at 250 ˚C.

(a)   (b)
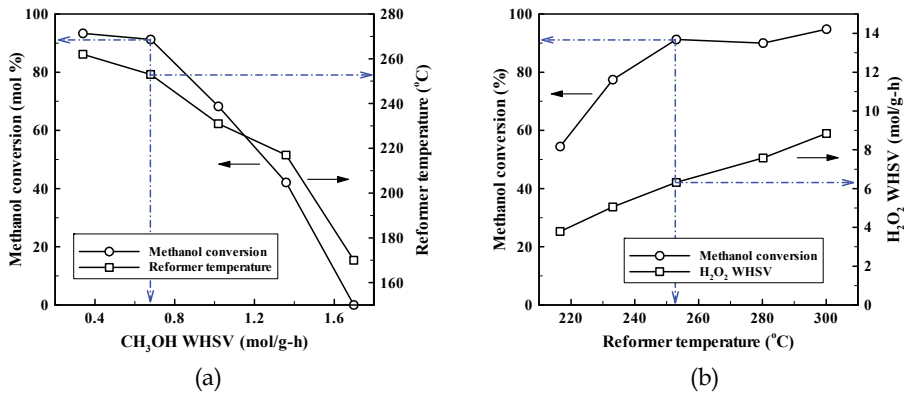
Fig. 14. Performance characteristics of micro reformer with hydrogen peroxide heat source
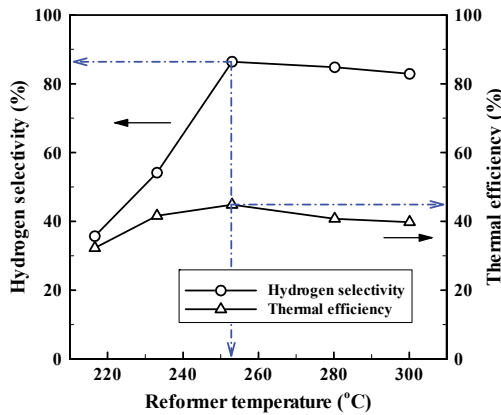
Fig. 15. Hydrogen selectivity and thermal efficiency as a function of reformer temperature

Figure 15 shows the hydrogen selectivity and the thermal efficiency of the system as a function of reformer temperature with the conditions of the reformer fixed. The thermal efficiency of the conventional reformer/combustor model is defined by:

$$\eta_T = \frac{LHV_{H_2\_produced}}{LHV_{CH_3OH\_reformer} - LHV_{H_2\_combustor}} \times 100 \qquad (5)$$

This formula could not be applied to the methanol reformer integrated with the hydrogen peroxide decomposition reactor, because the LHV of hydrogen peroxide is not defined. In the present study, the thermal efficiency for the reformer system is defined as follows:

$$\eta_T' = \frac{\Delta H^R_{H_2\_produced}}{\Delta H^R_{CH_3OH\_reformer} - \Delta H^R_{H_2O_2}} \times 100 \qquad (12)$$

The LHV was replaced with the heat of reaction. The LHV of hydrogen provided to the combustor in Eq. 5 was replaced with the decomposition heat of hydrogen peroxide. The hydrogen selectivity increased with the thermal efficiency as the reformer temperature increased. At the reformer temperature higher than 250 ˚C, however, the hydrogen selectivity decreased as the reformer temperature increased, because the production of carbon monoxide increased. The maximum hydrogen selectivity and the thermal efficiency were 86.4% and 44.8%, respectively. The product gas included 74.1% H_2, 24.5% CO_2 and 1.4% CO, and the total volume production rate was 23.5 ml/min. The hydrogen production rate is the sufficient amount to generate 1.5 W electrical power on a typical PEMFC. The optimum condition and the performance of the methanol reformer with hydrogen peroxide heat source are shown in Table 4.

The overall efficiency of typical PEMFC system using a methanol reformer is approximately 40% (Ishihara et al., 2004). In present study, the exergy loss can be reduced by the use of hydrogen peroxide decomposition reaction. The use of oxygen generated by the decomposition reaction raises the cell voltage, resulting in the increase of the fuel cell efficiency. It is understood that the overall efficiency of fuel cell system presented in present study is higher than that of the existing fuel cell model.

	H_2O_2 reactor	Reformer
Temperature	278 ˚C	253 ˚C
S/C (steam-to-carbon ratio)		1.1
H_2O_2 concentration	82 wt%	
Feed flow rate	2 ml/h	10 ml/h
WHSV	0.68 mol/g-h	6.32 mol/g-h
Conversion	98.4 %	91.2 %
H_2 production rate		23.5 ml/min
CO composition		1.4 %
Hydrogen selectivity		86.4%
Thermal efficiency		44.8%

Table 4. The optimum operation conditions and the performance of the integrated reformer

4. Integrated test with micro fuel cell

4.1 Removal of carbon monoxide

Removal of carbon monoxide from the reformate gas mixture is of paramount importance for development of a reformer in fuel cell applications because carbon monoxide deactivates

the anode catalyst of PEMFC. There are several processes for the carbon monoxide removal including pressure/temperature swing adsorption (PSA/TSA), methanation, membrane separation, and preferential oxidation. PSA/TSA are energy-intensive and expensive. Methanation consumes three moles of hydrogen to convert 1 mole CO into 1 mole methane as given below:

$$CO + 3H_2 \rightarrow CH_4 + H_2O \tag{13}$$

It is therefore not recommended. The membrane separation is attractive method because high purity hydrogen can be obtained. PROX also is the preferred method because the small amount of oxygen is required to oxidize CO into CO_2 as expressed below:

$$CO + 1/2O_2 \rightarrow CO_2 \tag{14}$$

4.2 Microreactor for preferential oxidation

Microreactor for PROX was prepared as shown in Fig. 16. Pt/Ru was selected as a catalyst of PROX. Microchannels were fabricated on a photosensitive glass.
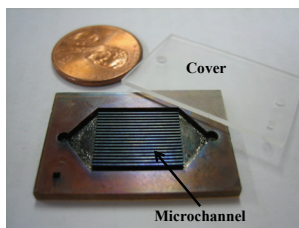
Fig. 16. Microreactor for preferential oxidation

As a washcoat layer, γ-Al_2O_3 was coated on the microchannels using sol-gel method and the catalyst was loaded by wet impregnation method. First, aluminum isopropoxide was hydrolyzed in deionized water with vigorous stirring for 1 hour at 80 ˚C. The sol was peptized by adding nitric acid (HNO_3) with adjusting the pH. Polyvinyl alcohol (PVA) solution was prepared by dissolving the PAV in deionized water at 75 ˚C. The presence of PVA can reduce crack formations of the washcoat layer at the drying time. The peptized sol and the PVA solution were mixed with adding the γ-Al_2O_3 powder to increase the concentration of γ-Al_2O_3 in the slurry. The mixture slurry was ball-milled for 72 hours. The glass substrate was then dipped into the prepared γ-Al_2O_3 slurry and dried for 2 hours at 120 ˚C after blowing off the excess slurry. This procedure was repeated to obtain the desired weight of the γ-Al_2O_3 washcoat layer. The washcoated microchannels were then calcined at 350 ˚C for 4 hours. A mixture of a 0.5 M aqueous solution of H_2PtCl_6 and a 0.5 M aqueous solution of $RuCl_3$ were prepared. The substrate was immersed in the mixture for 12 hours. The moisture was removed by drying the catalyst-loaded substrate in a convection oven at 70 ˚C for 12 hours. The calcination followed in a furnace at 350 ˚C for 3 hours. The catalyst was activated by reduction in a steady flowing hydrogen environment at 350 ˚C for 5 hours. The carbon monoxide conversion of PROX reactor as a function of the reaction temperature with varying the ratio of oxygen to carbon is shown in Fig. 17. Mixture gas including 69.91% H_2, 3.06% CO, 2.03% CH_4, and 25% CO_2 was used in the test of PROX reactor. The carbon monoxide conversion increased with the oxygen-to-carbon ratio and the reactor

temperature. In the case of 5 wt% Pt/Ru/γ-Al$_2$O$_3$ catalyst, the carbon monoxide was removed completely with oxygen-to-carbon ratio of 4 at 200 ˚C.
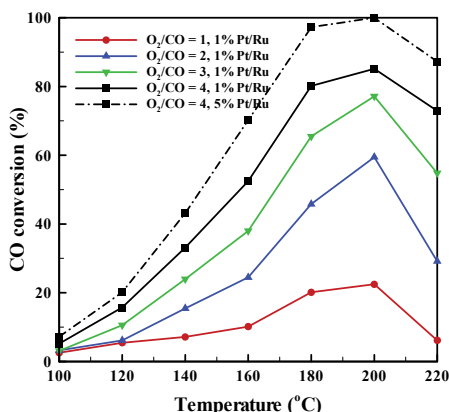
Fig. 17. Conversion of carbon monoxide of PROX microreactor

4.3 Integrated test with micro fuel cell

MEMS fuel cell was fabricated for integrated tests with the micro reformer. The structure of the micro fuel cell is shown in Fig. 18. Membrane electrode assembly (MEA) was prepared by coating 0.3 mg/cm^2 Pt-Ru/C for an anode catalyst and 0.3 mg/cm^2 Pt/C for a cathode catalyst on a Nafion-112 membrane. The reason to select Pt-Ru/C as an anode catalyst is because Pt/C is poisoned by carbon monoxide in the reformate gas even if removed via PROX reaction. Carbon paper (TGP-H-090, 260 μm) was used as a gas diffusion layer (GDL). Flow channels were fabricated by etching the photosensitive glass wafer, on which the current collectors, Ag/Ti layer, were sputtered. Overall fabrication process is presented in Fig. 18 and the fabricated micro fuel cell is shown in Fig. 19.
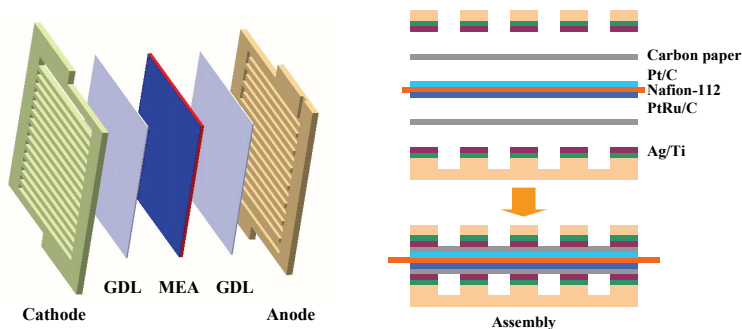
Fig. 18. Structure and fabrication process of MEMS fuel cell

Experimental layout for integrated tests of the reformer with the micro fuel cell is shown in Fig. 20. The micro fuel cell was tested with pure hydrogen to compare with the result with the reformate gas. Simultaneous operation of the micro reformer, PROX reactor, and micro fuel cell was carried out.

< Anode side >   < Cathode side >
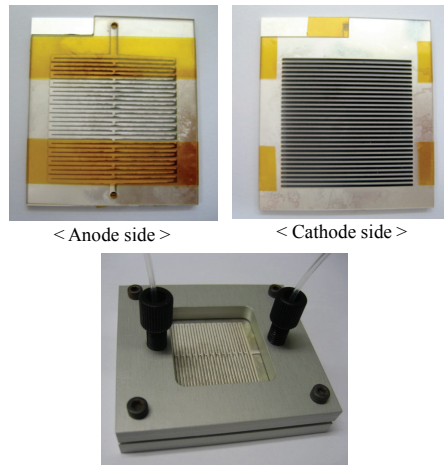
Fig. 19. Fabricated results of the micro fuel cell
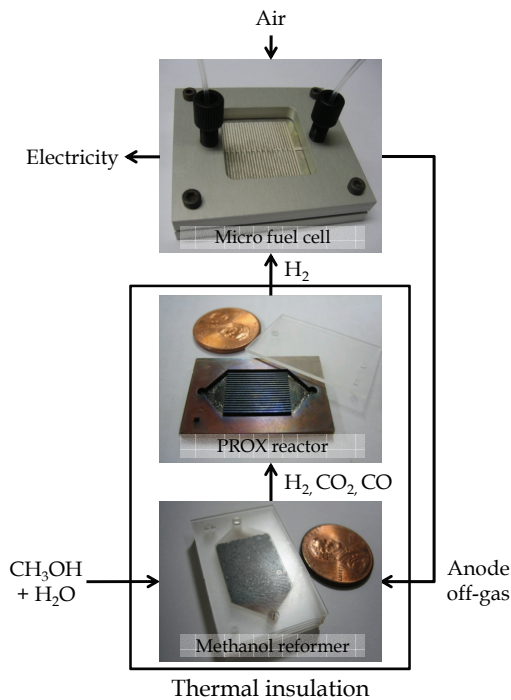
Fig. 20. Schematic of the integrated test of micro reformer-PROX reactor-micro fuel cell

4.4 Results and discussion

Performance of MEMS fuel cell system with pure hydrogen and the reformate gas is shown in Fig. 21. Pure hydrogen gas feed rate was set in 50 ml/min. When methanol feed rate was

2 ml/h, the flow rate of reformate gas was 71.96 ml/min. The reformate gas included 74.4% hydrogen, thus the hydrogen flow was 53.5 ml/min. The power density was 184 mW/cm^2 when the potential was 0.64 V. The performance was low compared with the result for pure hydrogen due to the feed at the fuel cell that included undesired CO, CO_2, and N_2.
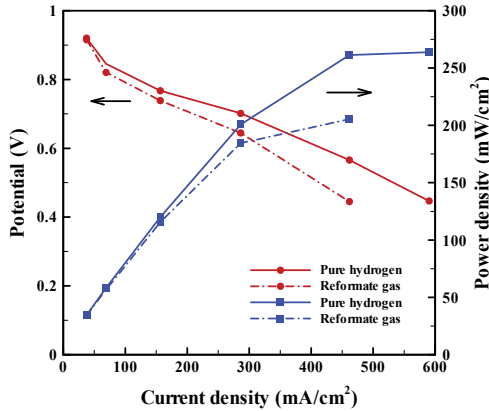
Fig. 21. Performance curve of MEMS fuel cell system

Specific energy density of the micro fuel cell system was calculated to compare with the state-of-art batteries. First, the overall energy budget for operation of the fuel cell system was calculated. Figure 22 presents the energy specification of each reaction step.
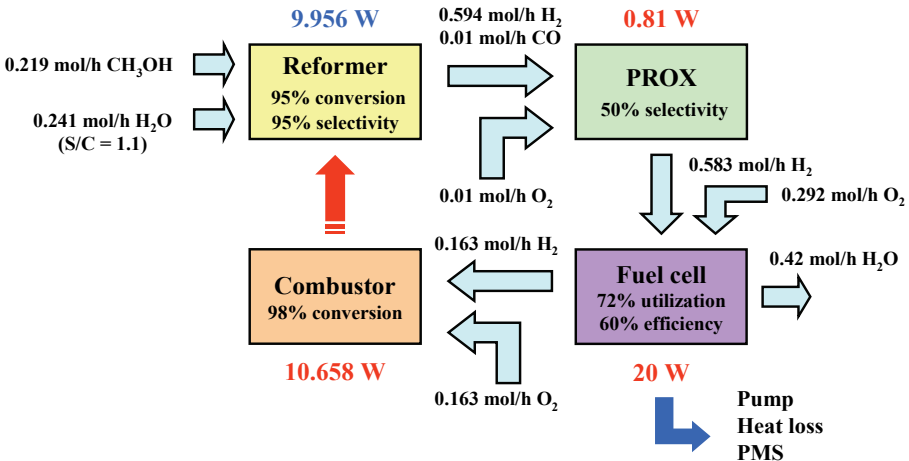
Fig. 22. Energy budget for a fuel cell system

The 20 W fuel cell system requires the hydrogen of 0.42 mol/hr. Thus, methanol feed rate of 0.219 mol/hr is required, assuming 95% methanol conversion and 95% hydrogen selectivity of the reformer. The energy requirement of the reformer consists of sensible heat, vaporization heat, and endothermic reforming reaction heat as given below:

$$
\int_{298}^{338} C_{p,CH_3OH(l)}dT + \int_{338}^{523} C_{p,CH_3OH(g)}dT + \Delta H_{CH_3OH}^v
$$

$$
+ \int_{298}^{373} C_{p,H_2O(l)}dT + \int_{373}^{523} C_{p,H_2O(g)}dT + \Delta H_{H_2O}^v + \Delta H_{523}^R
$$

(15)

The total energy input for the methanol reformer is 9.956 W. The catalytic combustor generates 10.658 W heat energy with the fuel cell off-gas of 0.163 mol/hr, which is greater than the reformer energy requirement. It means that the fuel cell system can be operated without the additional heat supply to sustain the methanol reforming reaction.

The methanol storage of 4.386 moles is required for the duration of 20 hours (0.219 mol/hr × 20 hr). The water feed requirement is 0.241 mol/hr at the steam-to-carbon ratio of 1.1, thus the water storage is 4.825 moles (0.241 mol/hr × 20 hr). These translate into 140.49 g (178.97 cc) methanol, and 87.093 g (87.25 cc) water, respectively. Therefore, the net fuel mixture storage requirement would be 227.58 g or 266.22 cc.

The specifications of the fabricated fuel cell are: mass of 0.5 g, volume of 2.7 cc, active area of 4 cm², and power density of 180 mW/cm². Thus 20 W fuel cell would have a mass of 13.89 g (0.5 g × 20 W / (0.18 W/cm² × 4 cm²)) and a volume of 75 cc. The specific power density of the micro reformer was 0.34 W/g or 1.25 W/cc. The reformer would have a mass of 59.62 g and a volume of 16 cc for 20 W fuel cell to be operated in the sufficient hydrogen supply. Therefore, the mass and volume of the total system were 301 g and 357 cc, respectively.

The energy storage capacity was 400 W·hr (20 W × 20 hr). So, the fuel cell system would have a weight specific energy density of 1329 W·hr/kg and a volume specific energy density of 1120 W·hr/L, which are values 10 times higher than the state-of-art of rechargeable batteries. The system energy density as the duration is shown in Fig. 23. The water production rate in the fuel cell was 0.42 mol/hr, which is greater than the water supply of the reformer (0.241 mol/hr) as shown in Fig. 22. Thus, the water from the fuel cell can be recycled into the reformer, improving the system energy densities. The specific energy densities for 10 days duration would be 2728 W·hr/kg and 2144 W·hr/L, respectively. It means that the micro fuel cell system can be an ideal alternative solution for portable micro power sources in the future.
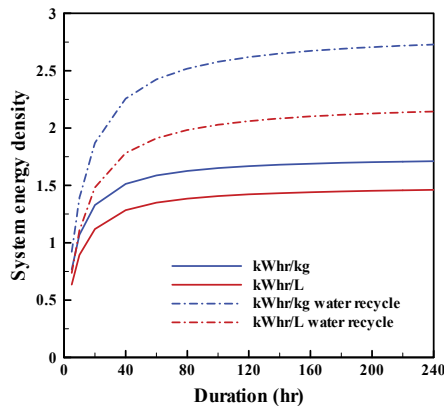
Fig. 23. System energy density as a function of the duration

5. Conclusion and future research

5.1 Conclusion

The design, fabrication and performance evaluation of micro methanol reformer integrated with a heat source were described in this chapter. The micro methanol reformer consists of the steam reformer, the catalytic combustor, and the heat exchanger in-between. The two heat sources for the reformer were used; one is the hydrogen catalytic combustion and the other is the hydrogen peroxide decomposition.

All reactions, the methanol reforming reaction, the hydrogen combustion, and the hydrogen peroxide decomposition, are the catalytic process. Cu/ZnO was used for the reformer and Pt for the catalytic combustor. The porous ceramic material was used as the catalyst support to enhance the catalytic surface area. The catalytic microreactor was fabricated on five photosensitive glass wafers; top and bottom covers, a reformer layer with Cu/ZnO/support insert, a combustor layer with Pt/support insert, and a heat exchanger layer in-between.

The performance of the reformer complete with the catalytic combustor was measured. The methanol conversion was 95.7%, and the thermal efficiency was 76.6%. The reformate gas flow including three major elements, 74.4% H_2, 24.36% CO_2, and 1.24% CO was 67.2 ml/min. The hydrogen flow in the reformate gas was the sufficient amount to run 4.5 W PEMFC.

The performance characteristics of the methanol reformer with the hydrogen peroxide heat source was investigated. The methanol conversion over 91.2% and the hydrogen selectivity over 86.4% were obtained. A modified thermal efficiency using the reaction heat of hydrogen peroxide instead of the LHV was defined and the thermal efficiency of the system was 44.8%. The reformate gas flow including 74.1% H_2, 24.5% CO_2 and 1.4% CO was 23.5 ml/min. This hydrogen was the sufficient amount to run 1.5 W PEMFC. The performance of the present methanol reformer can be further enhanced by using hydrogen peroxide with higher concentration.

The microreactor for the PROX reaction was fabricated using the photosensitive glass process integrated with the Pt/Ru/γ-Al$_2$O$_3$ sol-gel coating process. The carbon monoxide in the reformate gas was removed to use directly in the micro fuel cell.

The micro fuel cell was fabricated and connected with the micro reformer and PROX reactor.

The power density of the micro fuel cell system was 184 mW/cm² at the potential of 0.64 V and is lower than that in the case of pure hydrogen test, because the reformate gas included the undesired CO, CO_2, and N_2.

The system energy density of the micro fuel cell system integrated with the methanol reformer was calculated. The overall energy budget was calculated to operate the reformer-combined fuel cell system. The system energy storage density of the micro fuel cell system was obtained in the range of 1329 W·hr/kg to 2728 W·hr/kg. It is estimated that the micro fuel cell combined with the micro reformer has the energy density of up to 10 times higher than existing batteries, thus expecting to appear in the mobile energy market of the future.

5.2 Future research

Although the integrated methanol reformer developed in the present study can be used directly to operate the micro fuel cell, several works may be continued such as a fully integrated microfabrication, thermal packing, and optimization.

The micro reformer should be insulted thermally to obtain the high thermal efficiency and the low package temperature of the micro fuel cell system. The excess heat loss of the

reformer makes the catalytic combustor difficult to sustain the methanol reforming reaction. The thermal insulation of the reformer facilitates the integration of the reformer with the micro fuel cell at the low package temperature. The heat loss through conduction and convention can be prevented by the vacuum packaging technology using an anodic bonding process. The thermal design of the micro reformer through the extensive modeling of the heat transfer will be preceded to improve the overall thermal efficiency of the micro fuel cell system.

The fully integrated microfabrication of the micro fuel cell system is the next challenge to improve the system packaging efficiency. The batch fabrication of all elements including the micro reformer, PROX reactor, and micro fuel cell can reduce the fabrication cost. The overall integrated design of the micro fuel cell system should be optimized in consideration of the thermal balance and fluidic interconnections between the reactors. The micropump, microvalve, and control circuitry will be integrated with the micro reformer and micro fuel cell in the future.

6. Notation

a   Molal ratio of hydrogen peroxide to methanol
C_p Constant pressure specific heat, kJ/mol-K
LHV   Lower heating value, kJ/mol
O_2/C Oxygen-to-carbon ratio
S/C   Steam-to-carbon ratio
s   Molal ratio of water to methanol
WHSV   Weight hourly space velocity, mol/g-h
x   Molal concentration of hydrogen peroxide
η_T Thermal efficiency
ΔH^R Heat of reaction, kJ/mol
ΔH^V Vaporization heat, kJ/mol

7. References

Agrell, J.; Boutonnet, M. & Fierro, J. (2003). Production of hydrogen from methanol over binary Cu/ZnO catalysts Part II. Catalytic activity and reaction pathways, *Applied Catalysis A: General*, Vol. 253, pp. 213–223, 0926-860X

Delsman, E.; De Croon, M.; Pierik, A.; Kramer, G.; Cobden, P.; Hofmann, C.; Cominos, V. & Schouten, J. (2004). Design and operation of a preferential oxidation microdevice for a portable fuel processor, *Chemical Engineering Science*, Vol. 59, pp. 4795-4802, 0009-2509

Holladay, D.; Wainright, S.; Jones, O. & Gano, R. (2004). Power generation using a mesoscale fuel cell integrated with a microscale fuel processor, *Journal of Power Sources*, Vol. 130, pp. 111–118, 0378-7753

Ishihara, A.; Mitsushima, S.; Kamiya, N. & Ota, K. (2004). Exergy analysis of polymer electrolyte fuel cell systems using methanol, *Journal of Power Sources*, Vol. 126, pp. 34–40, 0378-7753

Kim, T. & Kwon, S. (2006a). Design, fabrication and testing of a catalytic microreactor for hydrogen production, *Journal of Micromechenics and Microengineering*, Vol. 16, pp. 1752–1760, 0960-1317

Kim, T. & Kwon, S. (2006b). Preparation of Cu/ZnO for Fabrication of a Micro Methanol Reformer, *Chemical Engineering Journal*, Vol. 123, No. 3, pp. 93-102, 1369-703X

Kim, T.; Hwang, J. & Kwon, S. (2007). A MEMS methanol reformer heated by decomposition of hydrogen peroxide, *Lab on a Chip*, Vol. 7, No. 7, pp. 836–847, 1473-0197

Kundu, A.; Jang, J.; Lee, H.; Kim, S.; Gil, J.; Jung, C. & Oh, Y. (2006). MEMS-based micro-fuel processor for application in a cell phone, *Journal of Power Sources*, Vol. 162, pp. 572–578, 0378-7753

Lindstrom, B. & Pettersson, L. (2001). Hydrogen generation by steam reforming of methanol over copper-based catalysts for fuel cell applications, *International Journal of Hydrogen Energy*, Vol. 26, pp. 923–933, 0360-3199

Lindström, B.; Agrell, J. & Pettersson, L. (2003). Combined methanol reforming for hydrogen generation over monolithic catalysts, *Chemical Engineering Journal*, Vol. 93, pp. 91–101, 1369-703X

Lua, G.; Wang, C.; Yen, T. & Zhang, X. (2004). Development and characterization of a silicon-based micro direct methanol fuel cell, *Electrochimica Acta*, Vol. 49, pp. 821–828, 0013-4686

Nguyen, N. & Chan S. (2006). Micromachined polymer electrolyte membrane and direct methanol fuel cells—a review, *Journal of Micromechanics and Microengineering*, Vol. 16, pp. R1–R12, 0960-1317

O'Hayre, R.; Cha, S.; Colella, W. & Prinz, F. (2006). *Fuel Cell Fundamentals*, pp. 10-11, John Wiley & Sons, Inc., 978-0-471-74148-0, New York

Pattekar, A. & Kothare, M. (2004). A Microreactor for Hydrogen Production in Micro Fuel Cell Applications, *Journal of Microelectromechical Systems*, Vol. 13, No. 1, pp. 7-18, 1057-7157

Pattekar, A. & Kothare, M. (2005). A radial microfluidic fuel processor, *Journal of Power Sources*, Vol. 147, pp. 116–127, 0378-7753

Schuessler, M.; Portscher, M. & Limbeck, U. (2003). Monolithic integrated fuel processor for the conversion of liquid methanol, *Catalysis Today*, Vol. 79–80, pp. 511–520, 0920-5861

Teshima, N.; Genfa, Z. & Dasgupta, P. Catalytic decomposition of hydrogen peroxide by a flow-through self-regulating platinum black heater, *Analytica Chimica Acta*, Vol. 510, pp. 9–13, 0003-2670

Wang, Z.; Xi, J.; Wang, W. & Lu, G. (2003). Selective production of hydrogen by partial oxidation of methanol over Cu/Cr catalysts, *Journal of Molecular Catalysis A: Chemical*, Vol. 191, pp. 123–134, 1381-1169

Yamazaki, Y. (2004). Application of MEMS technology to micro fuel cells, *Electrochimica Acta*, Vol. 50, pp. 663–666, 0013-4686

Yoshida, K.; Tanaka, S.; Hiraki, H. & Esashi, M. (2006). A micro fuel reformer integrated with a combustor and a microchannel evaporator, *Journal of Micromechanics and Microengineering*, Vol. 16, pp. S191–S197, 0960-1317

4

Non-contact Measurement of Thickness Uniformity of Chemically Etched Si Membranes by Fiber-Optic Low-Coherence Interferometry

Zoran Djinovic[1,3], Milos Tomic[2], Lazo Manojlovic[3],
Zarko Lazic[4] and Milce Smiljanic[4]

[1]*Vienna University of Technology, ISAS, Floragasse 7, Vienna,*
[2]*Institut bezbednosti, Kraljice Ane BB, Belgrade,*
[3]*Integrated Microsystems Austria, Viktor Kaplan str. 2, Wr. Neustadt,*
[4]*Institute of Microelectronics and Single Crystals, IHTM, Njegoseva 12, Belgrade,*
[1,3]*Austria*
[2,4]*Serbia*

1. Introduction

Micromachining of bulk Si is, nowadays, a matured technology in production of microelectromechanical (MEMS) devices such as freestanding mechanical structures like beams and membranes (Kovacs et al., 1998). There are two main techniques currently in use; wet and dry etching. The first one, particularly the anisotropic wet etching, is very often in standard production chain of piezoresistive pressure sensor (Sugiyama et al., 1983).

However, it is recognized very fast that etch uniformity across a wafer strongly depends on the crystal orientation of Si and type of etchant. This usually results with non-uniformity of the membrane thickness all around the Si wafer or within the membrane itself. (Dibi et al., 2000) report on the strong influence of the Si membrane flatness on the piezoresistive pressure sensor response. The main reason for the sensitivity loss of pressure sensor is the lack of parallelism of the two membrane sides. They found that flatness non-uniformity less than 1% on 30 μm membrane causes electrical response loss of about 3%. Also, they found that irregularity of the etched surface could be an important reason for the voltage offset of the final sensor. Because of that it is of paramount significance to know the final thickness of the membrane as well as the thickness uniformity across the wafer and membrane itself.

There is a list of papers (Bernstein et al., 1988; Tosaka et al., 1995; Mesheder & Koetter, 1999) dealing with different measuring techniques. The most interesting are optical techniques being fast, nondestructive and offer in situ measurement of Si membrane thickness during the etching. (Bernstein et al., 1988) adopted a commercial reflectance spectrometer to measure Si membrane thickness. The main drawback was that the instrument could work well in the range of 0.1-5 μm only. (Tosaka et al., 1995) developed a method for in situ monitoring of the Si diaphragm thickness based on multiple–beam interference spectroscopy. Again, the main limitation of the method was the measuring range of 2-20 μm of the membrane thickness. Additionally, the light spot on the diaphragm was 200 μm in

diameter that was the reason for the reduced space resolution. (Mesheder & Koetter, 1999) proposed transmission spectroscopy technique for in situ measurement of membrane thickness based on fiber-optic bundle for illuminating of the target. The technique works well but only in the range of 10-500 μm of thickness. The overall accuracy of the method is determined by the wafer homogeneity and accuracy of the initially measured wafer thickness.

In this paper we propose a contactless fiber-optic interferometric technique applicable for the fast and accurate measurement of the membrane thickness with accuracy of about 100nm. The method is based on low coherence interferometry, performed by "all-in-fiber" sensing configuration that is described in more details in (Djinovic et al. 2005; Tomic et al. 2002).

2. Principle of operation

Our sensing system is based on low-coherence interferometry performed in "all-in-fiber" Michelson interferometer shown in Fig. 1.

The core part of the sensing configuration is a fused 2x2 single mode (9/125 μm) optical coupler. The input arms of the coupler are connected with a low-coherence light source (LCS) and photodetector (PD), an InGaAs photodiode. As a low-coherence light source (LCS) we used the a superluminescent diode, SUPERLUM SLD-56-M1, that emmits at 1300 nm with spectral width of about 40 nm at FWHM. The outlet arms are directed to the Si membrane (sensing arm) and to the mirror (reference arm).
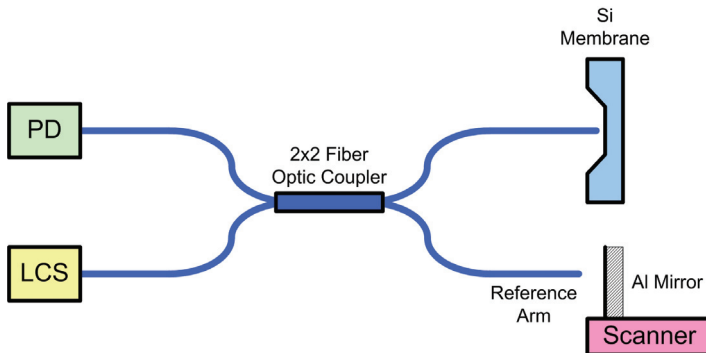
Fig. 1. Schematic presentation of "all-in-fiber" Michelson interferometer, LCS–white-light source, PD-photodiode

2.1 Algorithm

In Fig. 2 we depict a typical interferometric raw signal that we captured by photodiode. There are several characteristic low-coherence interferometric patterns that reflect the change of the transmitting media of low-coherence light. The four patterns, designed by 1, 2, 3 and 4 are important for the measurement.

The first pattern is the result of interference of the back reflected signal from the end of the sensing fiber and mirror in the reference arm. The second large pattern comes up due to interference between the light beams back-reflected from the front Si membrane surface and mirror in the reference arm. The third one is the result of interference of the back reflected

light beam from the rear Si membrane surface. The fourth one comes due to the multiple reflections inside the membrane and it is not important for the further analyze.
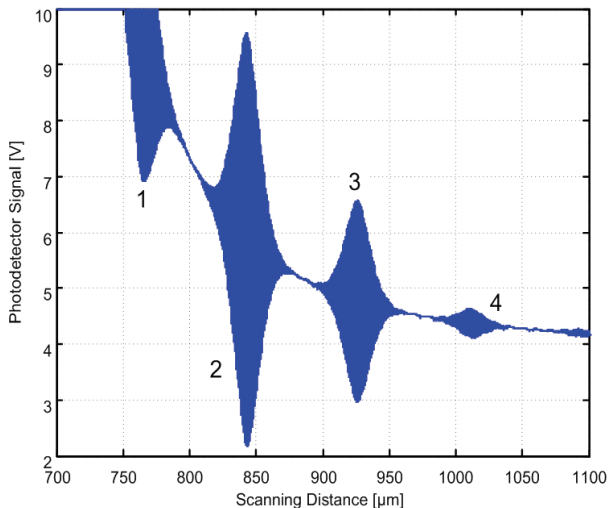
Fig. 2. Raw photodiode signal obtained by measurement of Si membrane thickness

From the raw signal it is difficult to extract the optical path difference between the interferograms. For that reason we need to make the basic signal processing consisted of high-pass filtering and envelope detection. In Fig. 3a is given the diagram of filtered signal without DC component together with the detected envelope in Fig. 3b. In this diagram is presented a part of the processed signal around the second interferometric pattern.

Based on the fitting of the detected envelope with the Gaussian curve it is possible to obtain the position of the zero optical path length difference. In Fig. 4a is given the detected envelope together with the fitted sum of the Gaussian functions (Fig. 4b). In Fig. 4b is presented part of the processed signal around the second interferometric pattern.

3. Experiment

Fig. 5 shows a close look to the sensing fiber directed against the etched side of the 3 inch {100} Si wafer in KOH solution according to the sensing configuration depicted in Fig. 1. We measured the thickness of the flat membranes and membranes with central boss. The last one is schematically presented in Fig. 6. The overall dimensions are 2x2 mm2.

We measured the membrane thickness and uniformity by scanning of one single membrane in x-y direction, subjecting several membranes in central part of the wafer and several membranes all around the periphery of the wafer.

4. Results and discussion

We calculated the optical path that light beam has passed through the membrane by determining of the distance difference between the central position of the second and third interferometric pattern. According to the algorithm we determined the central position of the interferometric patterns of interest by fitting of Gaussian functions. We determined this
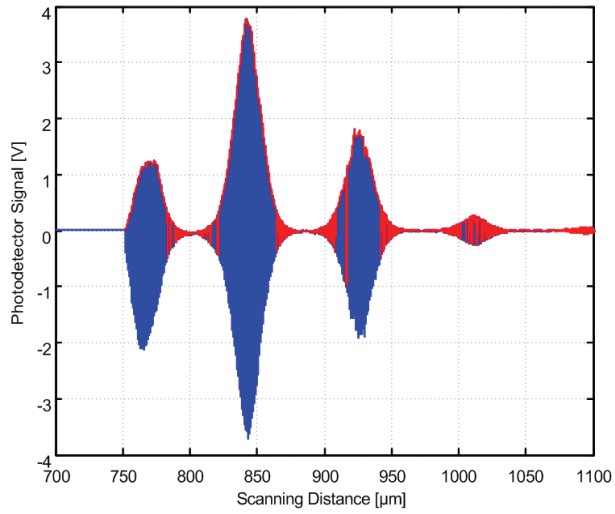
a)
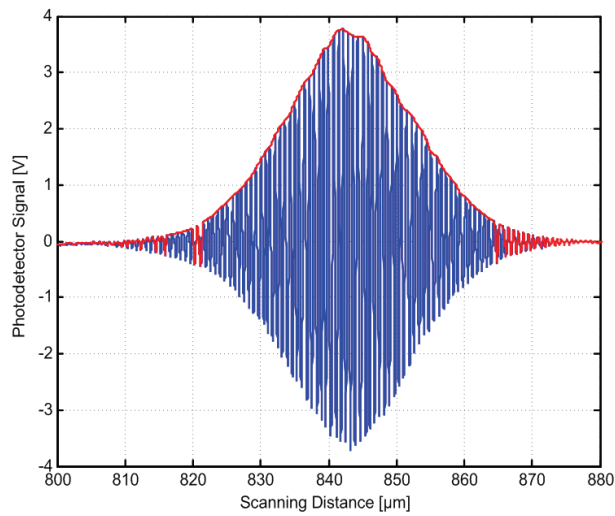
b)

Fig. 3. Filtered signal (a) and detected envelope (b). In the (b) diagram is presented part of the processed signal around the second interferometric pattern
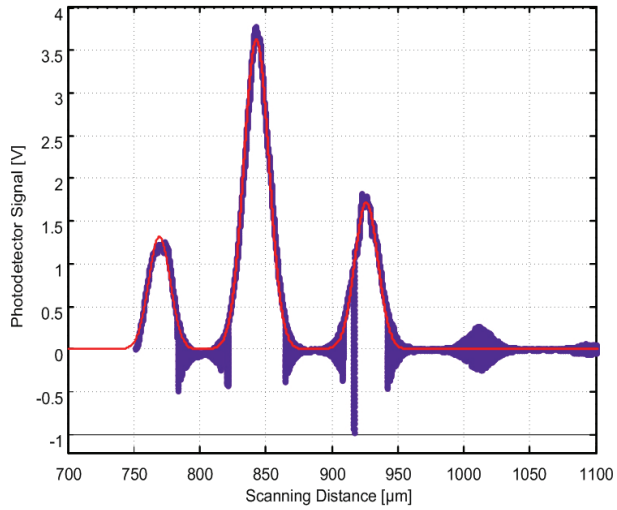
a)
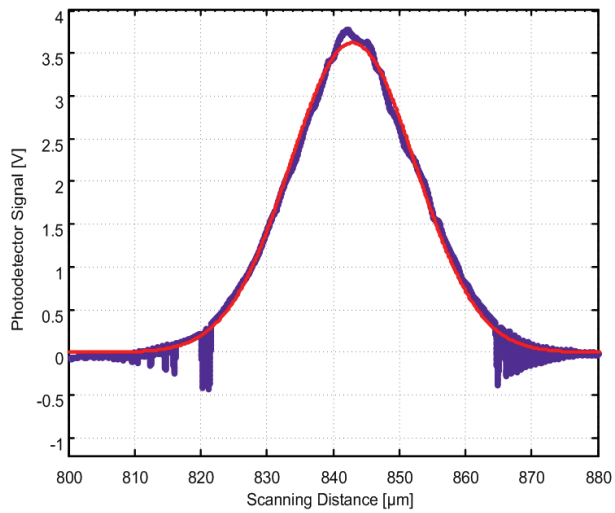
b)

Fig. 4. Detected envelopes (a) and fitted sum of the Gaussian functions (b). In the (b) diagram is presented part of the processed signal around the second interferometric pattern
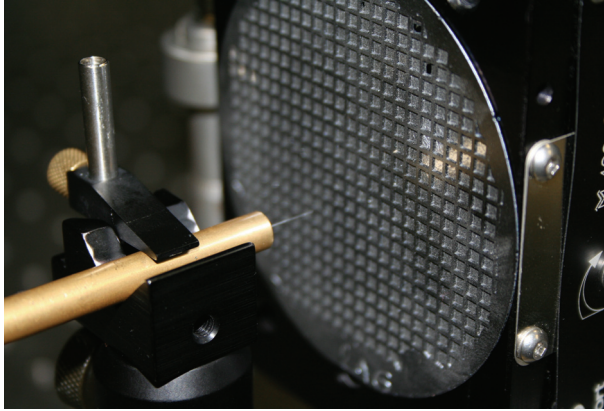
Fig. 5. Close look to the sensing fiber and 3 inch {100}Si wafer with chemically etched membranes
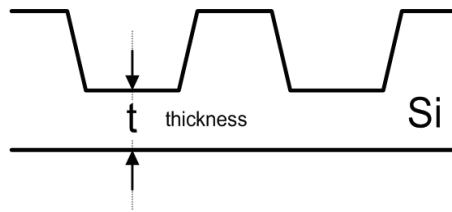
Fig. 6. Cross-section of one silicon membrane with central boss. t is the membrane thickness

position with accuracy of about 100 nm. The value of the difference between the central peaks is an optical thickness, t_{opt} of the membrane. Using the simple relation ($t_{opt}=t_{phys}\,n$) between the optical path and refraction index, n of Si (n=3.5085 @ 1300 nm, 23°C) we calculated the physical thickness, t_{phys} of the membrane.

Fig. 7 presents the results of calculation of physical thicknesses within the one subjected Si membrane with central boss along the 300 µm long line in peripheral zones of the membrane (left, right, up and down part of the membrane). Measuring points have been obtained by probing after every 50 µm. We obtained the average thickness of the membrane of 27.3 µm with uniformity in the range of ±0.4 µm. We also measured the roughness uniformity along all four peripheral zones of the Si wafer. We obtained average roughness of ±1.7 µm.

We also performed measurement of the membrane thickness of four membranes around the central one. In this measurement we obtained the average thickness of five central membranes of 27.8 µm and uniformity of about ±0.7 µm. To obtain the overall thickness and uniformity of the membranes at periphery of the wafer we measured the thickness of four membranes at the edges of the wafer. In this case we obtained average thickness of 25.6 µm and uniformity of about ±1.7 µm.

Using the same setup we measured the thickness and roughness of the membrane around the outer rim in the near proximity to the edge. Measuring points have been obtained by probing after every 10 µm in the range of 50 µm. The results are given in Fig. 8. The average thickness is 23.9 µm and the uniformity of ±0.11 µm. The roughness of the surface is ±0.23 µm.
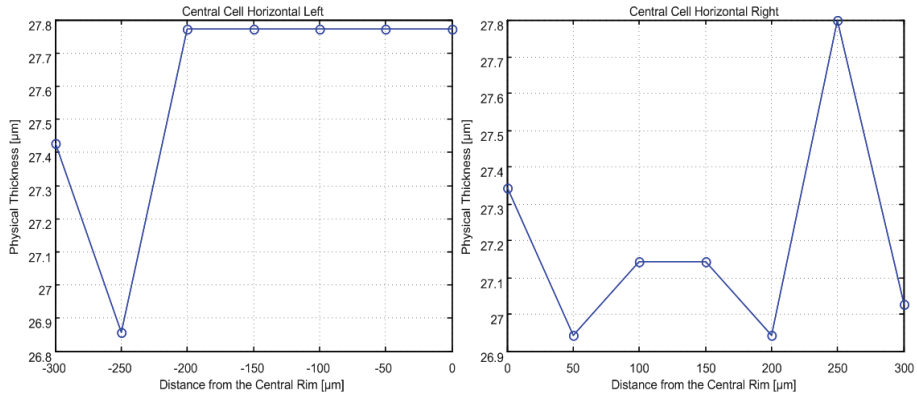
(a)
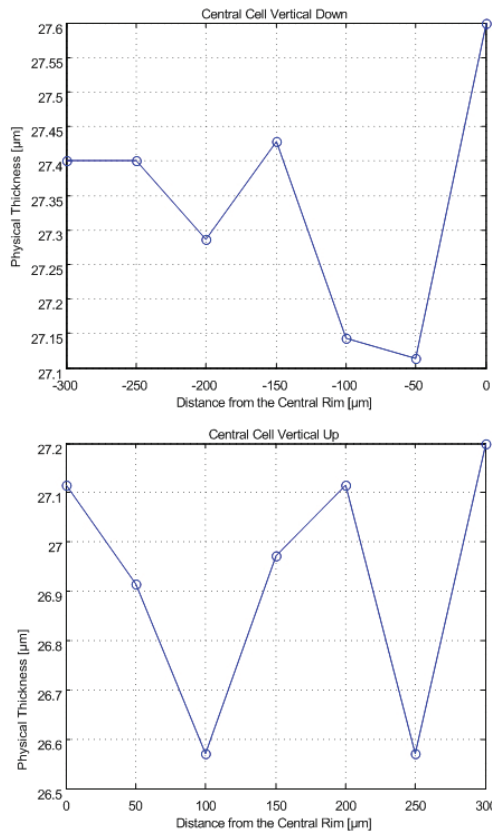
(b)

Fig. 7. Thicknesses within the one subjected Si membrane with central boss along the 300 μm long line in peripheral zones of the membrane (a) left and right; (b) up and down part of the membrane
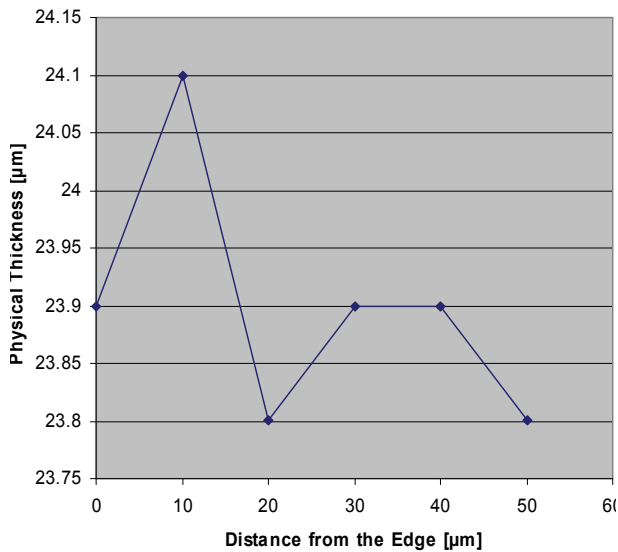
(a)
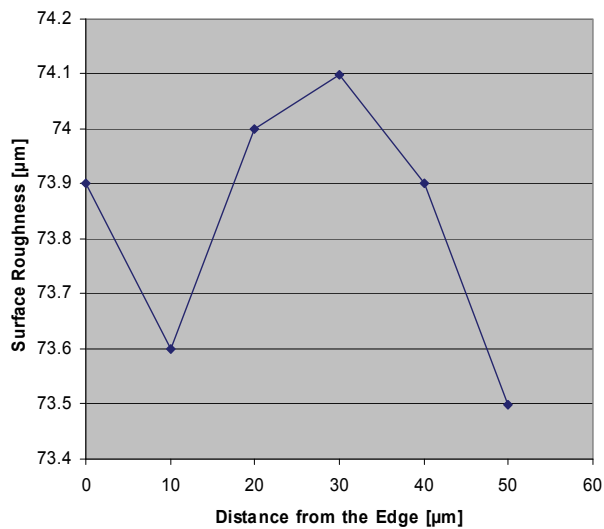
(b)

Fig. 8. Thickness (a) and roughness (b) of the membrane in the close proximity to the rim and edge

Fig. 9 presents the results of calculation of physical thicknesses within the one subjected flat Si membrane along the 200 µm long line in central zone of the membrane. Measuring points have been obtained by probing after every 10 µm. We obtained the average thickness of the membrane of 30µm with uniformity in the range of ±0.5µm. We also measured the thickness

uniformity along the central zone of the Si wafer, testing central part of 20 membranes. We obtained average thickness of 28,6 µm with scattering of ±1.6µm.
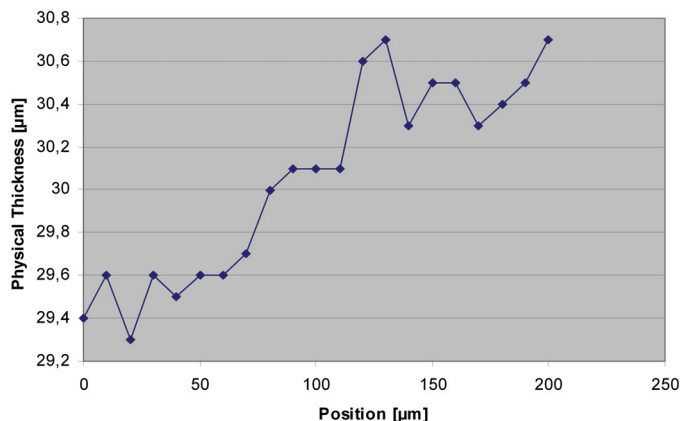
Fig. 9. Thickness uniformity within the one Si membrane of 2x2 mm in overall dimensions

The above results (Fig. 7 and 8) show that anisotropic chemical etching of the central Si membrane generates relatively rough surface of ±1.7 µm while the thickness uniformity is much better of ±0.4 µm. The first several neighbours have average membrane uniformity of ±0.7 µm. However, the average membrane uniformity of peripheral membranes at the same wafer of ±1.7 µm shows that chemical etching along the circumference of the wafer is affected by concentration variation of KOH solution. Similar results are obtained by flat membranes presented in Fig. 9. Such results are in accordance with findings given in (Dibi et al. 2000).

5. Conclusion

We presented here one contact-less optical technique based on low-coherence interferometry for measurement of thickness and uniformity of Si membranes. We performed a single-mode fiber-optic sensing configuration that is also applicable for the in situ measurement of membrane thickness. Space resolution was defined by diameter of spot of the impinging light of about 20 µm. The accuracy of the technique is about 100 nm.

6. Acknowledgement

7. References

Kovacs, G.T.A.; Maluf, N.T. & Petersen, K.E. (1998). Bulk micromachining of silicon, *Proceeding of the IEEE, Vol. 8*6, pp.1536-1551

Sugiyama, S.; Takigawa, M. & Igarski, I. (1983). Integrated piezoresistive pressure sensor with both voltage and frequency output, *Sensor and Actuators A*, 4 (1983) 113-120

Dibi, Z.; Boukabache, A. & Pons, P. (2000). Effect of the silicon membrane flatness defect on the piezoresistive pressure sensor response, *Proceedings of the 7th IEEE International Conference on Electronics, Circuits and Systems, ICECS 2000*, Vol. 2, pp. 853- 856

Bernstein, J.; Denison, M. & Greiff, P. (1988). Optical measurement of silicone membrane and beam thickness using a reflectance spectrometer, *IEEE Transactions on Electronic Devices*, 35 (1988) 801-803

Tosaka, H. ; Minami, K. & Esashi, M. (1995). Optical in situ monitoring of silicon diaphragm thickness during wet etching, *J. Micromech. Microeng.* , 5 (1995) 41-46

Mescheder, U. M. & Ch. Koetter, Ch. (1999). Optical monitoring and control of Si wet etching", *Sensors and Actuators,* 76 (1999) 425-430

Djinovic, Z.; M. Tomic, M. & Vujanic, A. (2005). Nanometer scale measurement of wear rate and vibrations by fiber-optic white light interferometry, *Sensors and Actuators*, A123-124 (2005) 92-98

Tomić, M; J. Elazar, J. & Djinović, Z. (2002). Low-coherence interferometric method for measurement of displacement based on a 3x3 fibre-optic directional coupler, *J. Opt. A: Pure Appl. Opt.* 4, (2002) 381-386

Nanomembrane:
A New MEMS/NEMS Building Block

Jovan Matovic[1] and Zoran Jakšić[2]
[1]Vienna University of Technology
[2]Institute of Chemistry, Technology and Metallurgy Belgrade
[1]Austria
[2]Serbia

1. Introduction

Although MEMS devices are exceptionally diversified, all of them are basically built from a very limited number of constituents manufactured in micrometer-size dimensions: plates, cantilevers, bridges, various channels. It is a combination of these basic building blocks that makes the almost limitless variety of MEMS devices. One may philosophize that at the very foundation of all technologies, MEMS and nanotechnologies included, one encounters atoms and molecules, and the combinations of such basic structural elements can be almost arbitrarily combined to obtain any structure and functionality. In the reality, however, there is a relatively small set of structures behind the existing nanostructures and nanosystems which includes

- Quasi-zero-dimensional structures – nanoparticles (including quantum dots, fullerenes/buckyballs, plasmonic nanoparticles, etc.). This class may be extended to include nanoholes/nanoapertures, nanocavities, etc. which can be seen as the complementary structures to nanoparticles (having a cavity instead of nanoparticle material and a solid instead of the surrounding vacuum or fluid).
- Quasi-1D: various nanorods, nanowires, nanotubes (both carbon and inorganic ones) – again including complementary structures.
- Quasi-2D: ultrathin films and membranes.

One may argue that this classification of nanostructural building blocks is actually similar to those already used e.g. in the fields of semiconductor quantum structures where one deals with quantum dots (zero-D), quantum wires (1D) or quantum wells (2D), or in the fields of photonic crystals and metamaterials (1D, 2D and 3D PBG structures), etc.

A plethora of other more complex nanostructures structures may be derived from those already listed and are actually built as their combinations. This includes e.g. quantum rings, various V-shaped nanostructures, nanostrips and ribbons, nanobuds, nanocubes, tripods, spheres, etc., and actually a whole menagerie of the nanotechnological 'zoo'.

It can be seen that plates/membranes appear in both MEMS and nanotechnology building blocks. In the latter, ultrathin membranes represent the basic quasi-2D nanostructure. In MEMS, membranes as structural elements belong to a wider class of microplates.

One possible classification of plates is based solely on their mechanical and geometrical properties. According to the theory of elasticity, plates are divided into three groups,

depending on their aspect ratio: the thick plates, the diaphragms and the membranes (Timoshenko, 1959). The thick plates balance the external distributed forces (pressure) by a combination of shearing and bending stress. The thick plates find application in MEMS as various supporting structures, rims and other structural elements within the MEMS devices. The diaphragms (thin plates) are somewhat arbitrary defined as plates with an aspect ratio 10 to 100. It is assumed that diaphragms resist the applied pressure solely by bending stress, which is a linear function of the external pressure. Such consideration is of course an approximation, as all three components of stress are simultaneously present in a loaded diaphragm (shear stress, bending stress and membrane stress). The diaphragms are widely used as the sensing element in MEMS pressure sensors.

The last class of plates are membranes, the relatively thinnest plates, with an aspect ratio exceeding 100. This definition is valid both in the macro world as in the MEMS systems. From the point of the theory of elasticity, membranes balance the external pressure exclusively by in-plane membrane stresses, i.e. without a bending rigidity component. A soap bubble or a rubber balloon are typical examples of membrane structures in the macro world. The membrane stress and deflection are highly nonlinear functions of the external load. The application of membranes in mechanical MEMS sensors is limited mostly to the capacitive pressure sensors and microphones where the deflection is small. Membranes find broader application in various bio/chemical sensors, microsleeves and MOEMS devices as e.g. micromirrors.

Regardless of the absolute dimensions or the aspect ratio of the plates used in typical MEMS devices, the basic physical properties of the plate materials remain unaltered when crossing from the macro to the micro world (Bagliom et al, 2007). This is valid for example for Young's modulus, thermal and electrical conductivity and optical properties of MEMS membranes. The values of these material parameters are identical in MEMS and in macro structures. The scaling issues which must be considered in micro-plates with lateral dimensions below 1 mm are limited to the influence of gravity forces, since the plate mass decreases with the third power of their dimensions.

A new situation, however, arises with a reduction of thickness of the common MEMS membranes under approximately 100 nm. Then the simple MEMS scaling laws are no more fully applicable. In membranes with a thickness of several nanometers and lateral dimensions of the order of a few millimeters, a number of new phenomena appear. For instance, quantum effects become marked in the heat and charge transport, as well as in electromagnetic interactions. Also, the fluid flow models valid for microstructures will not be applicable to the nanoflow featuring the structures with nanometric thickness. A plethora of other novel physical, chemical and mechanical effects appears in these structures, often leading to surprising or even outright counter-intuitive behavior. Such peculiarities classify ultrathin membranes to a separate group of MEMS/NEMS building blocks – the nanomembranes.

Nanomembranes may be defined as freestanding or free-floating artificial or natural structures with a thickness below 100 nm and a large aspect ratio which may exceed 1.000,000 (areas of several square centimeters or larger). Their low thickness is very near to the fundamental limit for solids, since typical nanomembranes may be below 5 nm thick. This approximates 15 atomic layers and makes the nanomembrane structure quasi 2D. In spite of such minute thickness and the extremely large aspect ratio such structures are self-supported – able to stand free in air or in vacuum, while in the case of free-floating structures they may be even monolayers with a thickness of about 0.3 nm.

Nanomembranes simultaneously belong to nano-objects (because of their thickness and their low-dimensional physics and chemistry) and to macroscopic objects (because of their large lateral dimensions). A nanomembrane is probably the only nanotechnological object which may be manipulated without special equipment and observed with a naked eye.

Currently the field of nanomembranes is still in its embryonic state regarding the related manufacturing processes, applications, and even the knowledge of their exact composition, structure and behavior. However, this novel paradigm attracts a growing attention of the scientific community and the number of related publication is growing exponentially.

In this chapter we deal with manufactured (artificial, man-made) freestanding nanomembranes and their peculiarities. We review the two major classes of nanomembranes, the inorganic and organic (typically polymeric) ones. We overview the approaches for their fabrication and functionalization. We stress the methods, procedures and structures developed by us. Finally we shortly present some applications of inorganic and organic nanomembranes. Their use as plasmonic structures for chemical, biochemical and biological sensing are presented in a separate chapter of this book.

2. Terminology and classification

The fields of science and technology are today more diverse than ever and are further diversifying at an accelerated pace. Many various fields utilize their own terms and idioms and sometimes these are in a discrepancy. This requires that linguistics and semantics follow the explosive growth of science.

The term 'membrane' is a good example, since it depends on the field where it is used. Its general meaning 'any thin, pliable material', and the other, almost equally often met 'a thin, pliable sheet or layer of animal or vegetable tissue serving to line an organ, connect parts, etc.' (Random House, 1992) will not have much meaning in technical sciences.

Membrane has a different meaning in microtechnologies (a relatively thin sheet with a large area, i.e. a large aspect ratio, often made of silicon and possibly having active elements in it), another in biochemistry, third in structural engineering. Previously this usually did not represent a problem, but science is becoming increasingly more multidisciplinary nowadays. Micro- and nanotechnologies include more and more diverse fields and misunderstandings are becoming unavoidable. For instance, if one speaks about a MEMS-based biochemical sensor incorporating a membrane, does this denote a thin, free-standing structure utilized as a (nano)sieve or an active layer containing a MOS transistor?

Nanomembranes may be seen as a special case of general membranes and actually the terminological confusion continues and even deepens here. In a large body of literature a nanomembrane represents a porous membrane with thickness that may be of the order of micrometers, even hundreds of micrometers, but is mesoporous or microporous (van Rijn, 2004). Thus the existence and dimensions of pores are used as the identifier of the whole membrane instead of its thickness and one arrives to a somewhat bizarre situation that the term 'nanomembrane' may denote a structure almost a millimeter thick and with a surface which may exceed square decimeters, only because it has nanometer-sized pores. This goes even further, and according to some sources the 'nanomembrane' may be any membrane if consisting of a nanostructured material (Fissell et al, 2006).

In this Chapter we use the term 'nanomembrane' to denote exclusively a freestanding film with a thickness below 100 nm. The bottom physical limit to the nanomembrane thickness is obviously a molecular or atomic monolayer, i.e. about 0.3 nm. As mentioned previously, the

lateral dimensions may be large at the same time, of the order of millimeters and even centimeters. It is clear that our definition of nanomembranes does not exclude the existence of nanopores or actually any other kind of functionalization.

Another term relatively often occurring in literature and related with nanomembranes is "ultrathin structure". Such a relative expression is somewhat misleading, and the attribute "ultrathin" is sometimes used for membranes with a thickness of the order of micrometers (Liu et al, 2004).

Some synonyms for the term "nanomembrane" encountered in literature include unbacked films, free standing films, freestanding membranes, ultra-thin (free standing) films, self-supported films, suspended nanofilms, free-floating films, atomic membranes, monolayer membranes, etc.

When membranes with nanometric thickness are mentioned in literature, very often one encounters structures deposited on some kind of a solid support. A common situation is met in nanofiltration and nanosieves, where the active nanofilm to be used as selective or barrier is deposited over a porous substrate with macroscopic, even mm-order thickness (Pientka et al, 2003). The porous substrate does not hinder the filtrate from passing through, while at the same time it ensures mechanical robustness. Some examples include polymer supports (Sackmann et al, 2000), ceramics (Jayaraman et al, 1995), zeolites (Tavolaro et al, 2007), etc.

This Chapter is dedicated exclusively to nanomembranes as freestanding or free-floating structures which are robust enough to ensure self-support. They may be in contact with gas from one side and with liquid from the other one, or may be completely surrounded by gas or by liquid. For some particular applications, the nanomembrane may be located in vacuum.

As far as their classification is concerned, the basic one would obviously be to artificial (man-made) and biological structures. In this Chapter we deal almost exclusively with the latter.

The artificial nanomembranes can be further divided into two large groups, the inorganic ("hard") ones and the organic ("soft") ones. Basically, one may say that inorganic structures are potentially more robust and are able to withstand more harsh conditions than the organic ones, including high temperatures and aggressive media. However, their simple structure does not leave much space for further functionalization and enhancement.

On the other hand, the organic nanomembranes, which also include the biological ones, are much more sensitive to environmental conditions and are destroyed even at moderately elevated temperatures, are chemically more sensitive and their mechanical robustness is typically poor. At the same time, their "toolbox" contains a virtually infinite number of different materials and the possibilities for functionalization are vast – the diversity of organic life forms and their functions being the prime example.

Obviously, one could also use combinations of both classes to make an infinite number of new composite membranes. It is also possible to use the processes found in biological structures to obtain biomimetic structures enhanced by novel functionalities offered by the nanotechnology.

3. Artificial nanomembranes and biomimetics

Biological structures are hierarchically organized. Out of a "toolbox" containing simpler building blocks more complex units are organized, and these units make new sets of building blocks for hierarchically higher toolboxes.

Nanostructures are near the bottom of this architectural pyramid of life. Biology means nanostructures. Every cell and every part of the cells, from the higher organisms down to bacteria, every virus, every prion are either nanosized structures or their conglomerates. Most of the building blocks of living organisms may be regarded as sophisticated 3D nanosystems.

The nanomembrane as a nanostructural unit is practically unavoidable along the path toward the top of the hierarchical pyramid of life. Typically a living cell includes lipid bilayer membranes incorporating various protein and lipid-based building blocks that enable the functionality of the cell. It divides the cytoplasm of the living cell from its environment and at the same time enables its active interaction with it. The important metabolic processes in each cell proceed through the nanomembrane, but also with its active participation. Throughout the wide variety of life forms, their very existence is enabled through the nanomembranes, which are definitely the most ubiquitous building element of life.

We may look at the living world through the eye of MEMS technologies. There are many reasons that their development nowadays is proceeding in two divergent directions. The first and the oldest approach is to use brute force when manipulating atoms and molecules during the fabrication of a device – the top-down method. A good example is the trend in the development of microprocessors and memory chips. To achieve a larger packaging density one needs cleaner rooms, more precise alignment tools and increasingly accurate process control. In spite of that, today's microprocessors with tens of millions of transistors are unable to emulate the behavior of a simple insect.

It is beyond any doubt that contemporary microtechnologies intensely contributed to an improvement of the overall quality of life – and continue to do so. Never before in any moment of the recorded history it was possible to keep information so safe and to interchange them so quickly. Practically one could not imagine whole fields of human activities without their results. Yet even such a vastly successful approach leaves much to be desired.

Let us consider an example: a swift with a weight of 50 grams and with a tiny brain possesses inertial, topographical, magnetic, solar, lunar and stellar navigation systems. It also has microactuators with sufficiently high power–weight ratio to fly non-stop for thousands of miles; it has built-in temperature controls; growth, reproduction and self-repair mechanisms; the ability to redesign itself both in the short term (as an adaptation to slightly changing conditions), and in the long term (resulting in another type of machine capable of other skills) (Vincent, 2000). Even our most sophisticated artificial machines are far from such abilities. Yet such an efficient and successful system was realized completely without fascinating facilities and equipment. The key is in the self-organization of the matter according to the nanoscopic blueprint contained in DNA.

Probably the best way to go would be to unite the both approaches, top-down and bottom up, into a single system utilizing the benefits of both. Biomimetics is surely important for this: why not mimic proven solutions from Nature? However, simultaneously one can introduce new ones, which maybe never existed before, into the same structures. Nanotechnologies already offered artificial structures without a known parallel in the real world and at the same time devices and systems closely mimicking the biological ones.

Nanomembranes are an excellent example since they integrate both. If their functionalities were sufficiently good that Nature promoted them into probably the most omnipresent building block of life, why not try and utilize the same already opened path in MEMS/NEMS?

We have an advantage along the way. We can mimic the nanomembranes offered by the Nature, but at the same time we are free to use the solutions and the functionalities the humanity arrived upon which do not have a parallel in the natural world, especially those introduced by nanoscience and nanotechnologies.

By continuing the development of these fundamental unit blocks of the living world, extending them both into the inorganic and further into organic and imparting them novel functionalities one could hardly expect that many exciting novel findings could not occur. This may well be the missing step to unify both of the existing approaches to the micro- and nanofabrication, thus offering unprecedented functionalities for the welfare of the whole humanity.

4. General notes on fabrication of artificial nanomembranes

Various methods for the fabrication of artificial nanomembranes are met within the technological cookbooks of both MEMS technologies and nanotechnologies. Different method tried until now will be presented in more detail further in this Chapter, but in this place we delineate a general procedure used with variations in many different situations. Figure 1 shows its most important technological steps.
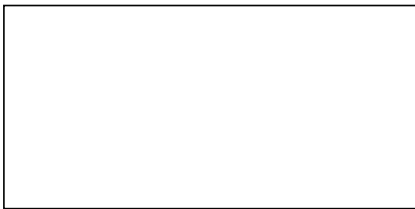
1. **Initial structure**: sacrificial solid substrate (the alternative is to use liquid as a substrate, in which case step 3 is skipped

2. **Deposition** of nanomembrane superstrate (precursor) which may or may not react or mix with the substrate
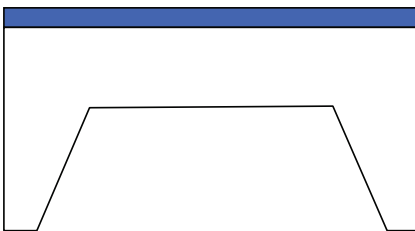
3. **Etching/Removal** of the sacrificial substrate
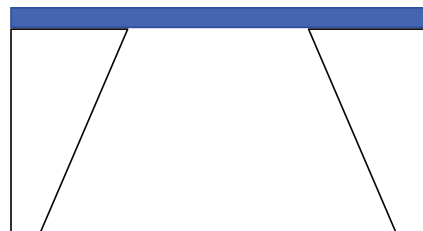
4. **Freeing the membrane** after the sacrificial structure is fully etched; the rims may remain as support

Fig. 1. Generic technological procedure for the fabrication of nanomembranes utilizing sacrificial support

Most of the approaches start from an initial structure to serve as a support and later to be removed to reveal the nanomembrane (Fig. 1, step 1). This sacrificial structure may be any

kind of solid, including the traditional materials of MEMS technologies, silicon and silicon dioxide. The choice does not end there, however, since practically any inorganic or organic solids may serve the purpose. Liquid media are also very convenient as supports for the growth of nanomembranes, actually in many situation they are even better choice – not to mention that most biological membranes are grown at a liquid-gas or liquid-liquid interface. A variation of this step is to use a solid or liquid support and apply over it an additional sacrificial layer.

The next step (Fig. 1, step 2) is where the creativity steps in and it differs widely among different methods. The nanometric structure to serve later as a nanomembrane itself is deposited on the support. Practically any material and any deposition method is available to this purpose. One may apply top-down deposition methods or use the bottom up approach. The real knowledge and art are contained in this step. Practically any research team in the field has its own approach and its 'trade secrets' connected with it. Some of the top-down methods used include RF or DC sputtering, electron beam technique, evaporation, electroplating, epitaxy, drop-evaporation method, various versions of chemical or physical vapor deposition, etc. Bottom-up methods include various self-assembly techniques.

The step 2 may also include further processing of the deposited material, to induce further changes in it or to cause its reaction or partial mixing with the substrate. For instance, one may utilize high-temperature annealing to change the crystalline structure/cause reordering of the material, change the structure or remove built-in stresses within the deposited nanolayer, fabricate nanopores, etc.

If a solid substrate is used or an additional sacrificial layer is applied, in step 3 one applies some of the etchants convenient for the particular kind of support material and selective towards the membrane material. For instance, if silicon is used, one may utilize anisotropic etchants like KOH, EDP or TMAH, or isotropic etchants like HF/Nitric/Acetic Acid (HNA).

The final step is also very important and may pose large challenges (Fig. 1, step 4). After the etching step is complete and the both sides of the membrane are exposed to the fluid environment, one needs to free the structure from the etchant.

The next two Sections deal with the specific methods to deposit inorganic and organic nanomembranes (step 2 above) and with the properties of thus obtained structures. In both sections the nanomembranes are classified according to their types and composition. For each type the specific fabrication methods and the properties of thus obtained structures are outlined. Such organization of the text was adopted because various types of nanomembranes vastly differ in both fabrication procedures and specific characteristics.

5. Inorganic nanomembranes

Although typically with a vastly simpler structure than the organic membranes (and especially the biological ones), the history of the inorganic membranes is, paradoxically, more recent than that of organic membranes. Of course, the first biological nanomembranes appeared simultaneously with the life itself. However, even the artificial organic nanomembranes are at least a century older than the inorganic ones. The existence of inorganic nanomembranes became possible only after the relevant fabrication procedures sufficiently matured to ensure repeatable and reliable production of freestanding nanostructures with extremely high aspect ratios.

In this text we first consider the inorganic structures, because their structure is the simplest one and their properties also shed a light to the function of their more complex organic counterparts.

There are several main classes of inorganic freestanding structures in dependence on their composition. The simplest one includes pure element nanomembrane and may be divided into three separate subclasses. The first one are pure metal structures; the second one that merits a separate consideration includes predominantly carbon-based nanomembranes, where probably the most important are diamond and diamondoid (diamond-like) ones. The third subclass belongs to elemental semiconductor membranes, where the key position belongs to the silicon structures, a consequence of the dominant position of this material in MEMS and NEMS (more than 90% of all microsystems are silicon-based).

Another large class includes simple inorganic compounds like oxides, nitrides, carbides and similar mechanically robust materials. Among the oxides, silicon dioxide takes a central position, again because of its dominance in MEMS.

The third important class includes glass and ceramic nanomembranes.

5.1 Pure metal nanomembranes

Probably the simplest nanomembranes are those consisting of pure metals. Some of the used materials include chromium, nickel, aluminum, platinum, palladium, silver, gold and similar. One also encounters titanium, tungsten, lead, tin, and copper. A common trait of these materials are that they are structural metals with better or worse mechanical characteristics. Most of them are used in microelectronics and in MEMS, but some of them are known for their use in catalysis. Their chemical inertness is usually rather good.

The history of pure metal nanomembranes can be tracked back to 1931 when Winch manufactured ultrathin freestanding gold films (Winch, 1931). The thickness was 80 nm, a remarkable feat at the time. Manufacturing process was sputtering of gold on the surface of a polished halite mineral (sodium chloride). After gold film deposition, the halite substrate was dissolved in water, leaving the nanomembrane floating free on the water surface.

Not much work on metallic nanomembranes was done in the following decades of the 20th century and most of it was concentrated on very specific applications. We mention here metallic film nanomembranes for the X-ray and the extreme UV spectroscopy, both for atmospheric and space application. The prevalent manufacturing technology was evaporation of metallic layer on a parting substance (Carpenter & Curico, 1950), (Hunter, 1982). A more advanced technology, evaporation of metals as Pt and Cr in the UHV chamber, was used in studies of electron ballistic transport in the ultrathin metallic films (Aristov et al, 1998), (Stepanov et al, 1998), (Glotzer 2004).

The areas of the first metal nanomembranes were typically smaller than 0.1 mm^2 and their aspect ratio was below 500. Most of them tended to be brittle and fragile. Metallic nanomembranes are inherently electrically conductive.

Typically the ultrathin pure metal membranes are fabricated utilizing the conventional microsystem technologies, i.e. they used the well known technique of sacrificial supporting structure (Striemer & Fauchet, 2006), as outlined in Section 4.

According to (Striemer & Fauchet, 2006), pure metal nanomembranes with surface areas up to about 10 mm^2 have been obtained and with aspect ratios up to 430,000:1

5.2 Metal-composite nanomembranes

As far as the terminology is concerned, the term "composite" within the context of nanomembranes means that their structure consists of a mixture of two or more materials. These may be alloys, single crystals, polycrystals, homogeneous mixture of nanoagglo-

merates, etc. In this Subsection we stress nanocomposite structures consisting of one or more metals with various additional ingredients, which may include oxides, silicon, etc. Such structures may be uniform or gradient, but they are always homogenized at a level close to the molecular/atomic – or at least at a dimensional level much smaller than the nanomembrane thickness.

In our work we introduced a simple method for the fabrication of nanomembranes with giant aspect ratios exceeding 1.000.000 (Matovic & Jakšić, 2009). We used only standard procedures of the MEMS technology, although with parameters deviating from the values common in the art. Examples of our metal-composite nanomembranes are shown in Figs. 2 and 3.
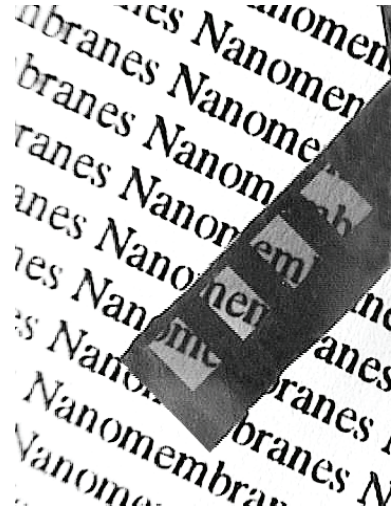
Fig. 2. Photo of an array of metal-composite nanomembranes. Area 1.5 x 2.5 mm, thickness 7 nm
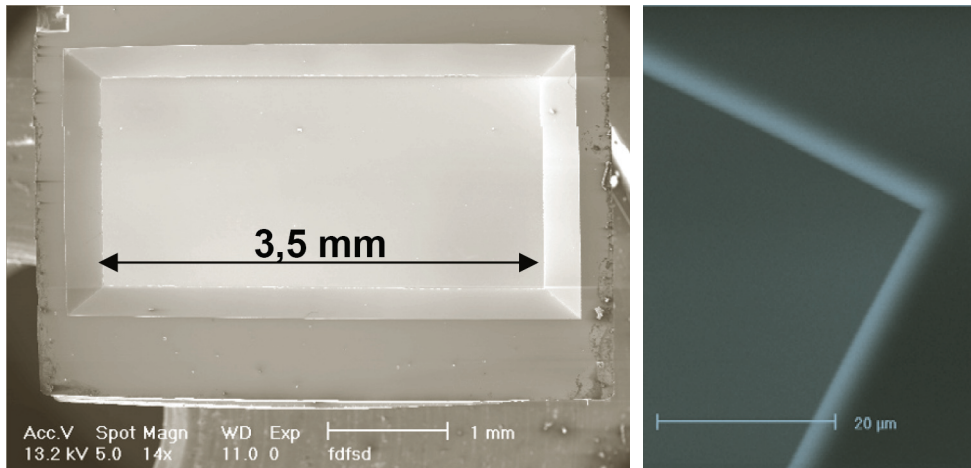
Fig. 3. Giant metal-composite nanomembranes fabricated at TU Vienna. Left: photograph of a 3.5 mm long nanomembrane, thickness 8 nm. Right: SEM of the edge of the same membrane (Matovic & Jakšić, 2009)

Our nanomembranes were made using 50-20 nm thin sputtered films of chromium deposited onto a silicon surface. The sputtering procedure was modified in order to enable formation of a metal-composite complex. A minute and well controlled amount of reactive gas was introduced into the system, which resulted in partial oxidation of the fabricated film. The sputtered chromium atoms were kept at 5 eV, allowing a penetration into the sacrificial silicon to a depth of about 0.7 nm. As the subsequent release process we used a back side etch procedure to completely remove sacrificial silicon and free the nanomembrane. The composition of a typical nanomembrane is shown in Fig. 4.
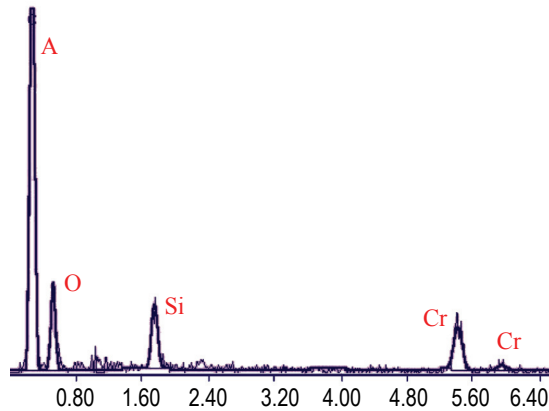
Fig. 4. Energy dispersive X-ray spectrometry analysis of the nanomembrane composition: metal ~52%, oxygen ~28%, silicon ~ 20%.

The area of metal composite nanomembranes can be increased to several millimeters square, even centimeter square areas. The thickness of our membranes was in the range 5-50 nm, as measured by profilometry after placing the membrane on a polished silicon wafer surface.

Until now only the use of oxides with a good adhesion towards the basic metal was confirmed. The application of some complex aggregates, line nanoceramics or clays was not investigated in metallic nanomembranes.

(Vendamme et al, 2006) reported the fabrication of 35 nm thick nanomembranes based on hybrid interpenetrating networks of organic and inorganic materials. They used simple spin coating to deposit their membranes on the support.

Metal composite nanomembranes are robust on macroscopic level (Fig. 5) and allow aspect ratios in excess of 1,000,000. They exhibit high flexibility and mechanical strength. They are optically transparent. (Vendamme et al, 2006) reported that their nanomembranes are sufficiently robust to be able to hold amounts of liquid 70,000 times heavier than their own weight, and at the same time sufficiently flexible pass without damages through holes 30,000 times narrower than their size.

For some composite films it was reported that they have a self-healing mechanism able to seal relatively large openings and to restore the structures driven beyond the plastic deformation limit (Jiang et al, 2004). In our own experiments with metal-composite nanomembranes we also observed this effect.

There are important differences in properties between polymer composite and metal composite nanomembranes:
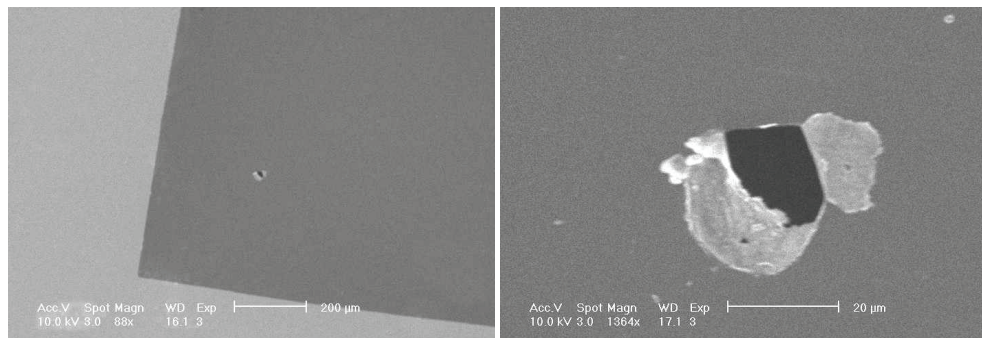
Fig. 5. The example of robustness of a metallic composite NM. SEM picture of a membrane pierced near its edge and a magnified photo of the hole. Even at a large magnification the cross-section of the hole walls is too thin to be observed. The lateral dimensions are 2.5 x 2.5 mm, thickness 10 nm.

- Metallic nanomembranes are mechanically more stable than their polymer counterparts. According to preliminary investigations, metallic nanomembranes do not creep under a prolonged stress. Overloaded, they break similar to brittle Si membranes, while under the same conditions polymer nanomembranes creep, see Figure.
- Young's modulus is significantly higher than in polymer composite nanomembranes, although the available data are still not sufficiently reliable
- Metallic nanomembranes are inherently electrically conductive
- Optical transmission is generally lower than in the case of polymer nanomembranes and strongly dependent on thickness. The transmission of nanomembranes 6-7 nm thick is about 75%, while nanomembranes thicker than 20 nm are practically opaque and show a mirror-like behavior. However, these properties may be readily engineered to obtain any desired transparency from 0% to 100%, for instance by utilizing multilayer structures and/or making use of plasmon resonance (Feng et al, 2005 1), (Feng et al, 2005 2).
- The surface roughness of metallic composite nanomembranes is significantly lower than in the case of polymer composite nanomembranes. Already now the mean roughness is below 2 nm, which points out to possible uses for MOEMS devices. Naturally, for some applications it may be desirable to have a higher roughness.
- The robustness of metal-composite nanomembranes is very high. They permit handling with only a modest level of precautions, including the manipulations with wafers supporting nanomembranes using bare fingers (Matovic & Jakšić, 2009).

5.3 Carbon, diamond, and diamond-like nanomembranes

Amorphous, diamond-like carbon films have exhibited excellent material properties such as chemical stability, wear resistance and optical transparency (Petrmichl & Veerasamy, 2001).. This resulted in their wide use as protective coatings in numerous applications.

Carbon-based diamondoid (diamond-like) nanomembranes were fabricated in Sandia by Friedmann & Sullivan, 1997). The hydrogenated forms of these films, a-C:H, and specifically the near-frictionless carbon (NFC) films developed at Argonne National Laboratory have exhibited the lowest ever recorded friction coefficient, 0.001, and ultra-low wear rates of

10^{-11}–10^{-10} mm^3 N^{-1} m^{-1} (Robertson, 2002). All of these functionalities are available in freestanding structures (Zhou et al, 2006).

It is even more important that freestanding diamond membranes could be readily combined with any of polymer, metal and other composite nanomembrane materials. Surprisingly, no results of investigations of such structures were published until now, although such synthesis of different technologies is actually very promising and could vastly expand the spectrum of possible applications.

It should be also mentioned that although diamond coatings were investigated as a wide bandgap semiconductor material and generally as microelectronics material, no such investigations were extended to freestanding ultrathin structures, although novel functionalities achieved through the use of quantum effects due to the low dimensionality of these structures could be expected.

5.4 Elemental semiconductor nanomembranes

Another class of membranes are those fabricated of single crystalline semiconductors. These may consist of elemental semiconductors, like silicon, compound semiconductors, for instance gallium nitride, silicon carbide, etc., or semiconductor combinations (e.g. lattice-mismatched strained Ge on Si, based on the concepts of strain sharing and critical thickness). Such nanomembranes are free-standing, relieved from elastic strain, flexible and dislocation free.

Obviously the central position among the nanomembranes based on elemental semiconductors belongs to silicon. Its technological basis is extremely well developed and mature and the material itself is called "the king of semiconductors". Its mechanical properties make it the material of choice for MEMS systems. Micrometer thick silicon membranes are the basis of many microsystem sensors, but also of a large number of other structures. Nanomembranes represent a natural continuation of the development of such MEMS building blocks to the NEMS area.

Another reason for the fabrication of Si nanomembranes is stretchable electronics (Kim et al, 2008). To this purpose large nanomembranes were made in silicon and applied to different substrates. Buckled, wavy Si nanoribbons and full Si nanomembranes (Choi et al, 2007) with a thickness anywhere between 10 nm and 100 nm have been reported. These were fabricated on elastomeric supports of poly(dimethylsiloxane) (PDMS). Full 2D stretchability of such membranes was obtained. Doping, dielectric layers, metallization and other structures were applied to such stretchable/compressible nanomembranes to fabricate MOSFET transistors and p-n junction diodes. Both n-doped and p-doped silicon were used for the fabrication of nanomembranes.

5.5 Oxide, nitride, carbide nanomembranes

Various materials belong to this group. Silicon dioxide is most often met. (Striemer, 2006) mentions nanomembranes made from TiO_2, MgF_2, Ta_2O_3, Y_2O_3, La_2O_3, HfO_2, ZrO_2. (Branton et al, 2005) describe the fabrication of silica and silicon nitride membranes. Silicon nitride membranes fabricated utilizing the stencil mask technique were reported in (Deshmukh et al, 1999), (Toh et al, 2004). Gallium arsenide nanomembranes were mentioned in (Choi et al, 2007).

It is noticeable that most of the mentioned materials are mechanically robust. They are either dielectrics or wide bandgap semiconductors. Among them, silicon dioxide is relatively least

robust and it is brittle, however it is extremely well known as a material and its technology is mature, with its roots in microelectronic technologies.

Many (but far from all) of the oxide and generally simple inorganic compound nanomembranes are brittle and fragile, not able to withstand large mechanical forces and with relatively modest aspect ratios.

5.6 Glass, ceramic nanomembranes

Glasses and ceramics were also mentioned as materials for experimental nanomembranes. They were used as porous materials, thus enabling the application as nanosieves.

Sodium borosilicate glass was used to fabricate porous membranes with a thickness of about 100 nm (Enke et al, 2002). To this purpose the initial ultrathin plate of glass was treated by phase-separation and combined acid and alkaline leaching. The pore dimensions in such a membrane were adjusted in the range between 1 and 120 nm.

6. Organic nanomembranes: macromolecules/ macromolecular composites

Organic chemistry and biochemistry offer an almost infinite number of macromolecular formulations for nanomembranes, followed with an extraordinarily wide spectrum of chemical, biochemical and physicochemical properties. All macromolecular nanomembranes share some common physical properties. Their mechanical properties are highly sensitive to humidity, temperature and solvents. Further, they have a low Young's modulus, low temperature range and creep under permanent stress. Since nanomembranes thickness is a few dozen of atoms, the instability of their physical properties under the influence of the mentioned agents is more pronounced than in solid bodies with the same composition.

Collodion (pyroxylin) was the first macromolecular nanomembrane (Yamauchi, 2000). It was discovered in 19th Century. Collodions were fabricated by solving the nitrocellulose in acetone. The solution was poured to the water surface and after acetone evaporated, it left a thin nitrocellulose film, floating on the water surface. The film was released using a wire loop. Thus a free-standing nanomembrane was obtained, and it was relatively stable and durable. Collodion is practically forgotten nowadays, but the described method of nanomembrane fabrication is today among the mainstream ones: one of the many examples that today's technologies often rest on a knowledge accumulated for a long period of time.

Some organic nanomembranes are based on liquid crystalline films (Sonin, 1999), but they are very unstable. Further, every year hundreds of papers are published on Langmuir-Blodgett (LB) method (Peterson, 1990), which allows one to construct amphiphile multilayers with a thickness ranging from 5 to 500 nm. A Langmuir–Blodgett film is formed by one or more organic monolayers on a liquid surface and is often transferred to a solid surface through its immersion into or emersion from the liquid. Other applicable methods are based on thiol or silane compounds, spin coating and thermal deposition of macromolecules onto a substrate. Unfortunately, these methods do not allow one to control the molecular order in the films.

Finally, there is a more recent method for film self-assembly that makes use of the alternate adsorption of oppositely charged macromolecules (polymers, nanoparticles, proteins) – the Layer-by-Layer technique (LbL) (Lvov et al, 1995). The assembly of alternating layers of oppositely charged long chained molecules is a simple process, closely mimicking the natural self-organized structures. The process provides the means to form 5–500 nm thick

films with monolayers of various substances growing in a pre-set sequence on a substrates at a growth step of about 1 nm per cycle. These nanomembranes have a lower molecular order than the LB films, but the advantages are high strength and easy preparation. This technique has been called in some places "molecular beaker epitaxy," meaning by this that with simple instruments (exploiting the materials self-assembly tendency) one can produce molecularly organized films similar to the ones obtained with highly sophisticated and expensive molecular beam epitaxy technology used in top-down approach.

Briefly, the LbL process proceeds as follows. A substrate is immersed into a dilute solution of a cationic polyelectrolyte, optimized for the adsorption of a single monolayer (ca 1 nm thick), rinsed and dried. The next step is the immersion of the polycation-covered substrate into a dilute dispersion of polyanions or negatively charged nanoparticles, also optimized for the adsorption of a monolayer, and then rinsed and dried. These operations complete one cycle of the self-assembly of a polyelectrolyte monolayer onto the substrate. (Fig. 6).

The subsequent sandwich units are analogously self-assembled. Linear polycation/polyanion multilayers can be assembled by similar means. Different molecules, proteins, enzymes, antigens and so on may be assembled in a pre-planned order in a single film.
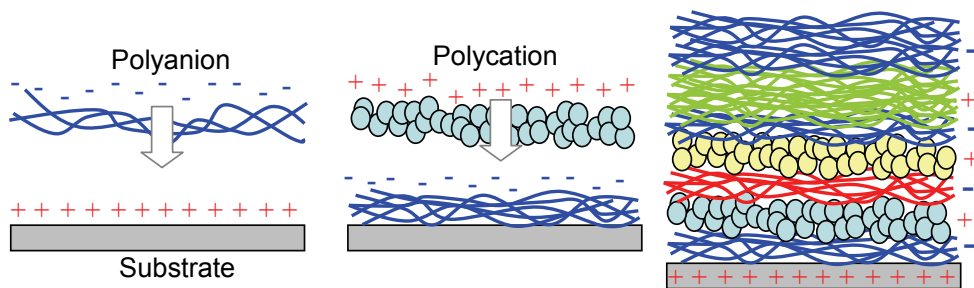
Fig. 6. Illustration of the layer-by-layer nanomembrane assembly technique. Left: assembly of the first layer; middle: assembly of the second layer; right: final structure.

7. Basic methods of nanomembrane functionalization

Nanomembrane functionalization is the essential step towards fundamental extension of their application. This includes an imparting additional desirable mechanical, electronic, chemical, biological, optical, magnetic, etc. nanomembrane properties. Thus it represents an essential factor in extending their practical applicability.

7.1 Nanomembrane functionalization by active fillers

An obvious method to enhance the performance of nanomembrane is to include fillers into their matrix, for instance nanoparticles with desired properties, and thus to create an enhanced nanocomposite (Zhang & Ali, 2007). In the current implementations the presence of fine fillers or clusters in the nanomembrane matrix mostly enriches the durability and overall mechanical properties of the nanomembrane. Such nanofillers are the passive (inactive) ones.

Inactive fillers can be replaced by nanoparticles already possessing an additional functionality, for instance catalytic (Pt, Ag, Rh), photocatalytic (TiO_2), light-emitting (SiC,

ZnSe), soft and hard magnetic nanoparticles, electrets or piezoelectric ceramics, chemically active substances, etc. The fillers may consist of nanoparticles or structures with inherent multifunctionality, like fullerenes of nanotubes. Fig. 7 shows some possible nanofillers.

The introduction of active nanoparticle fillers should be performed without compromising the existing mechanical properties of the nanomembranes.

Nanoparticle incorporation may be done during the production of the membranes themselves (using any of the various methods like incorporation by melt compounding, incorporation during polymerization step, blending and hot (dry/wet) isostatic pressing, plasma spraying techniques and co-evaporation / codeposition, etc).
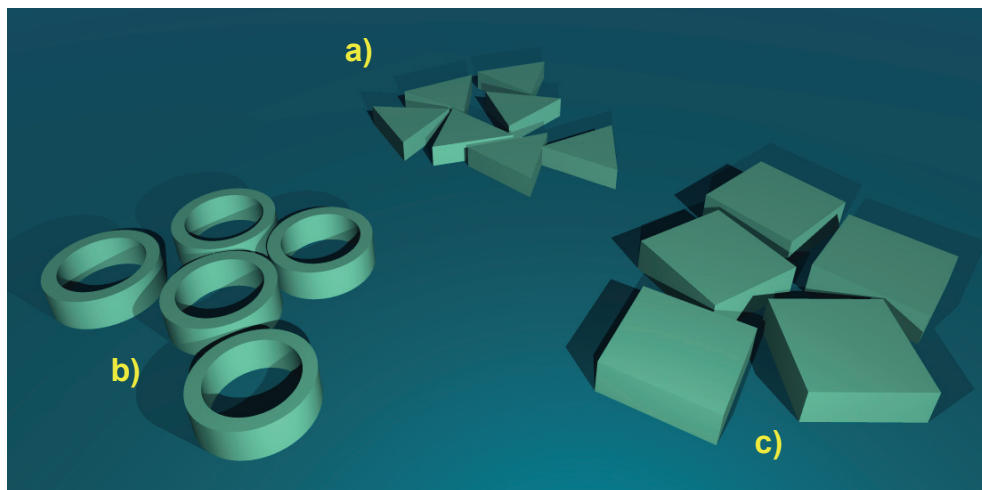
Fig. 7. Various forms of nano-fillers. a) triangular nanoprisms (may be fabricated in gold); b) nanorings (zinc oxide); c) nano-cubes (silver). The nano-building blocks are not to scale.

7.2 Lamination

The simplest way to introduce multiple functionalities to nanomembranes is to fabricate sandwich structures, i.e. to perform lamination. The fabricated superstructures may consist of two (the simplest case) or more strata. Each stratum introduces its own properties to the nanocomposite. For instance, one may be conductive; another may ensure mechanical strength and robustness. One may introduce a layer to serve as a plasmonic waveguide and another may be a ligand layer to attract a specific chemical or biochemical only.

Fig. 8 shows several possible configurations for nanomembrane functionalization through lamination. Obviously, the number of possible combinations of properties and functions is virtually limitless, and the basic limitation stems from the adhesion and reactions between different strata.

An important kind of functionalization through lamination is passivation treatment, necessary for mechanically sensitive or unstable surfaces, for those reactive, oxidizable, etc. Such treatment includes diamond-like materials, polymers, metals, semiconductors and their combinations. Chemical modification may be done by surfactants.
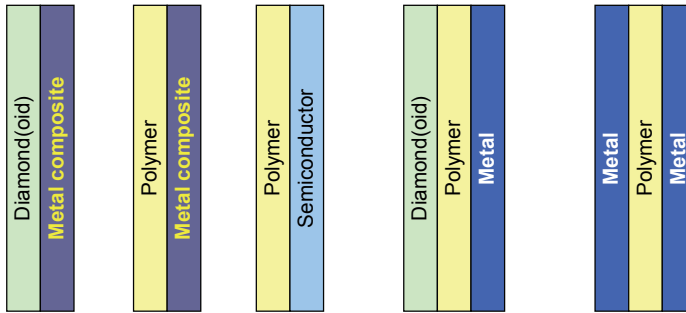
Fig. 8. Several possible basic kinds of nanomembrane functionalization through lamination.

7.3 Structural modification

Another large group of functionalization methods is based on patterning of nanomembrane structures. For instance, one may form 2D arrays of nanoholes/nanoapertures with controlled geometry in the nanomembranes. Fabrication of 2D nanohole (nanopore) arrays also belongs to nanopatterning and brings with itself another group of functionalities (Fig. 9). Nanopatterning of membranes may be done for instance by microcontact printing, using dry lift-off approaches and many other approaches (Hannink, 2006). Due to the extremely small thickness of nanomembranes the ratio between the pore width and the nanomembrane thickness remains small. Due to this, it is much easer to maintain the control of the pore dimensions during micro/nanofabrication, since the aspect ratio (diameter to width) of the pores themselves is low.
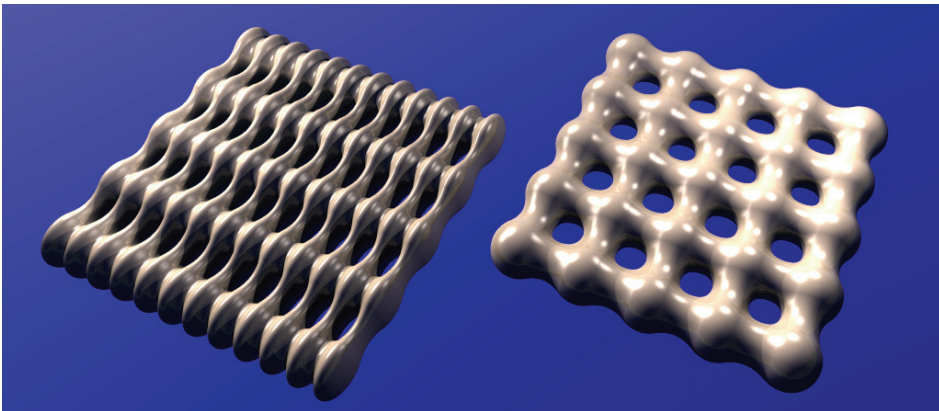
Fig. 9. Structural modification of nanomembranes through formation of 2D nanoaperture patterns; left: elliptical openings; right: circular openings.

A method of forming pores in freestanding silicon nanomembranes was described by (Striemer et al, 2007), Fig. 10. After depositing a silica-silicon-silica sandwich onto a Si sacrificial substrate one anneals the Si nanomembrane precursor and as a result pores are formed. The higher the annealing temperature, the larger is the pore size. It can be seen that the procedure is a variation of the generic sacrificial process for nanomembrane fabrication, as outlined in Section 4.
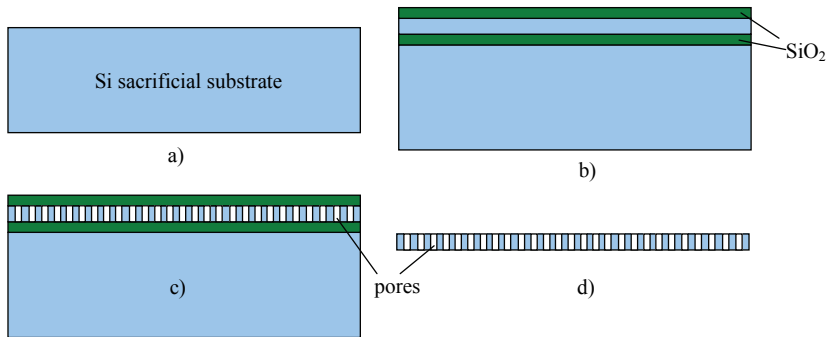
Fig. 10. Fabrication of pores in Si nanomembrane. a) Sacrificial silicon substrate; b) layers of silicon dioxide, silicon and again silicon dioxide are deposited; c) after annealing, pores are formed in silicon; d) sacrificial Si is etched away.
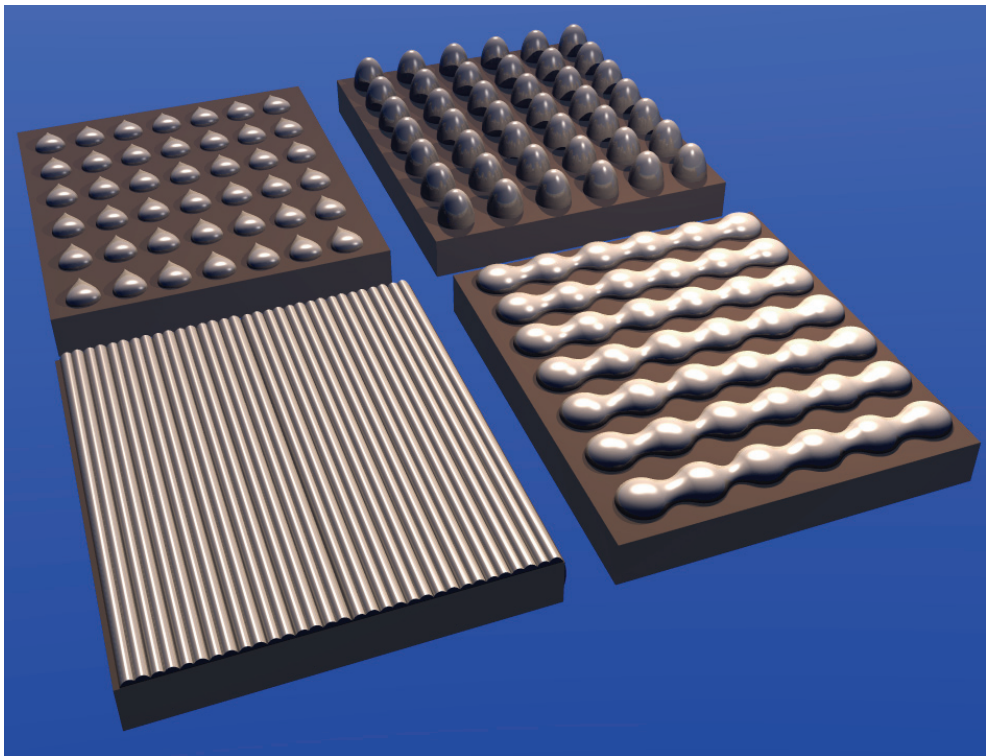
Fig. 11. Nanomembrane modification through various forms of surface modification.

7.4 Surface modification

An approach which may be regarded as a subgroup of the structural modification is the modification of the nanomembrane surface. However, this is a very rich and diverse group in itself and includes a number of subclasses. Some of those include

- nanopatterning/deposition of various planar and 3D structures directly to the nanomembrane surface. These patterns may modify the electromagnetic or optical behavior of the surface, enhance or suppress local wettability, improve attraction or repellence toward biological/living agents, etc.
- chemical, biochemical or biological activation through deposition of e.g. ligands, catalysts, etc.
- Passivation surface treatment, necessary for mechanically or chemically unstable nanomembranes (e.g. oxidizable, reactive, etc.)

Some surface modifications obtainable through the application of nanolithography are shown in Fig. 11. The alterations may be done in the nanomembrane material itself or through the deposition of an additional pattern which may assume the form of various pillars, stripes, etc.

8. Some basic applications of nanomembranes

Since nanomembranes are a novel concept which extends the range of MEMS & NEMS building blocks and practically introduces a new one, this means that whole branches of science and technology can be re-read and re-created through it, which obviously has a potential to create an enormous number of novel applications. In this Chapter we only give a brief summary of some of the potential applications. Practically it may be said that the nanomembranes are applicable everywhere where microfabrication, MEMS/NEMS and nanotechnologies are applicable. One of the reasons for this is their unique property to combine nanometric dimension in one direction with mm-, even cm in other dimensions, thus offering nanotechnological functionality with an ease of manipulation, use and characterization.

The potential applications of nanomembranes can be roughly divided into two groups:

1. *Quantitative and qualitative improvements of the known applications of conventional membranes.*
   New physical properties brought by the low-dimensionality of nanometric order are sure to result in improvements of the existing or already proposed functionalities
2. *Qualitatively new functionalities brought by low-dimensional physics of nanomembranes.*
   Nanomembranes possess a plethora of novel properties which are yet to be investigated and their applications found. Already those investigated and proposed until now seem extremely promising. For instance, nanomembranes represent a generalization not only of conventional membranes/diaphragms in MEMS/NEMS, but also of microbridges and nanocantilevers (obtainable by forming the nanomembrane). Novel applications may be expected from their use in e.g. subwavelength optics, electromagnetics; quantum behavior resulting from their extreme low-dimensionality also offers new functionalities, etc. One may claim that whole fields of technology and application should be reviewed through the possibilities offered by the existence of membranes an order of magnitude thinner than those conventionally used.

Among the obvious properties of nanomembranes is a vastly improved ratio of the effective area to the total volume of the structure compared to thicker structures. In Fig. 12 one may see two membranes contained in an identical volume, but with different thickness. Although the ratio of their thickness is about an order of magnitude only, one can see that the effective area is much larger for the thinner membrane. In reality the ratios are much larger, and may well be 3-6 orders of magnitude. Due to the simple scaling issues (surface

decreasing with the second power, volume with the third power of the geometrical dimensions) it is obvious that the surface to volume ratio may be very large. This is of immense importance for various catalytic/enzymatic processes. One may expect higher conversion efficiencies and simultaneously a significant reduction of the necessary amount of expensive metals (e.g. platinum, palladium) in catalytic systems like microreactors, microconvertors, etc.
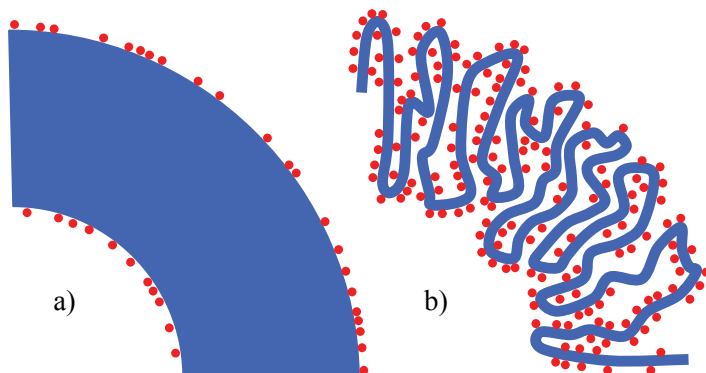
Fig. 12. Comparison of effective areas for adsorption of two membranes with different thickness, but contained in an identical volume. The number of adsorbed particles, being proportional to the area, is obviously larger in (b).

Among devices that could make use of the improved catalytic properties of nanomembranes are micro-batteries and battery arrays, various kinds of fuel generators and fuel converters, fuel purification systems (ranging from those for e.g. gasoline desulphurization to inexpensive and low-scale nuclear fuel purification). Especially, the nanomembranes may be used as a viable alternative to the conventional Nafion®-based membranes (Hamrock & Yandrasits, 2006) in proton-exchange fuel cells to improve their efficiency.

Another large field of nanomembrane application is for nanosieves and nanofilters. In nanomembranes the aspect ratio between the pore radius and the overall structure thickness is much smaller than in conventional porous membranes, which are typically 1000-10,000 times thicker than the particles to be separated (Striemer et al, 2007). A smaller aspect ratio means that the filtered particle has to pass a much shorter path through the filtering channel and consequently the transport rates/throughput will be larger. Also, in structures with a smaller aspect ratio it is also much easier to control and tune the cross section of the nanoaperture. This may result in an improved selectivity of the nanosieve and avoidance of the loss of filtrate within the membrane itself. In nanomembranes it is also easier to control the distribution of electrical charges within the channel, which will also improve filtering selectivity, this time by influencing the transport of polar molecules.

At the same time, the possibility to use novel strategies for nanomembrane fabrication enables to make them mechanically much more robust than the conventional structures. As an example, SiN_x nanomembranes incorporating nanohole arrays and intended for the use as filters and bioseparators) (Tong et al, 2004) about 10-30 nm thick are able to withstand single-sided overpressures of about 10 bar.

Nanosieves may be utilized for diverse applications, ranging from fluid-separation and gas-separation units and various size-exclusion-based separations and extractions to molecular

and microbial sieves for separation and purification of organic agents (for instance, absolute sterile filtration of bacteria & viruses). Among the important applications are DNA and microorganism separation, different proteomics devices and the fields of use include nanomedicine, genetic treatment and analysis, microbiology, biochemistry, bioseparation, but also environmental sensing and chemical/biochemical sensing generally (templates for nanosensors), homeland security, pathogen recognition, and also nanofabrication (e.g. masks for stencils) etc.

An important field of the application of nanomembranes is fuel purification and separation, including hydrogen recovery from gas mixtures and hydrocarbon fractionation. Another one is water desalination, detoxification and purification. The reverse osmosis by the use of nanomembranes was described in (Murad & Nitsche, 2004). Ion separation by macromolecular nanomembranes was elaborated in (Toutianoush & Tieke, 2001). An additional feature for filtering applications may be expected from a combination of nanomembranes with pores and the application of electrostatic potential, since the resulting buildup of the electric field enhances the selective rejection of ions (Schaldach et al, 2004).

Further, in biological/non-biological interfaces a nanomembrane may serve as an "ideal" biomimetic (enhanced or not) biological/nonbiological interface, convenient for chemical functionalization, patterning at the micro and nanometer scale, but also enabling engineering at macroscale (Watanabe & Ishihara, 2008). It may serve as a substrate for immobilization of proteins, antibodies, enzymes, cells through pattering a 2D arrays of anchoring units (DNA footprinting platform).

8.1 Main fields of nanomembrane application

In the further text, we present the main fields of nanomembrane application. For the convenience, we divided possible applications of nanomembranes into several main fields of interest and quoted only a few examples for each.

Energy production and conversion

- Fuel cells; other micro-power sources and microbattery arrays; nuclear fuel production/purification; selective removal, separation, extraction of hydrogen; gasoline desulphurization; petrochemistry; hydrocarbon fractionation; fuel production (environment-friendly, microreactor-based, biomimetic); fuel purification/separation (environment-friendly, microreactor-based).

Biomedicine

- Drug delivery systems; cancer treatment (selective drug delivery/selective enzyme removal); hemodialysis; blood treatment/filtration; medical sensors incl. implantable devices; diagnostic instruments; healthcare systems; immune system investigation.

Bioengineering

- DNA analysis/separation; gene analysis/research, genomics; cell biology and biotechnology; nanosieves; protein biomarker detection; virus and prion analysis; fabrication of diverse biomimetic nanostructures; biochemical sensing.

Chemical Engineering

- Food and beverage processing; seawater desalination; multicomponent gas mixtures separation; gas dehydration; other chemicals separation; reverse osmosis filtration;

chemicals analysis; microfiltration; ultrafiltration; nanofiltration; ultrapure chemicals production (environment-friendly, microreactor-based); μ-separators; atomization; emulsification.

Environmental Protection

- Recover volatile organic compounds from airstreams; wastewater treatment; air pollution control; recovery of valuable chemicals; water recycling; potable water production/purification; chemical, biochemical and biological sensors.

Toxicology, forensics and homeland defense

- Recognition of harmful inorganic, organic and biological agents.

MEMS/NEMS

Novel MEMS/NEMS devices (expanding the limits of all MEMS/NEMS building blocks to nanometer thickness and offering a whole new class of free-standing structures); various physical and chemical sensors for process industry, automotive industry, airspace, etc. including highly sensitive pressure and acoustic microsensors; nanosieves; self-recoverable micro and nanostructures; very high frequency microoscillators and microresonators; catalytic membrane microreactors; high temperature microreactors; "Lab on a chip"; molecular transport and sorting; nano printing and etching; stenciling using nanosieves.

Semiconductors

- e-projection lithography; subwavelength lithography based on "superlensing"; purification of microelectronics grade rinse water; interconnects and conductors; "superconductive" circuit components; flexible electronics.

Optics & Electromagnetics

- Free-standing integrated optics; thermal detectors; radiation and particle detectors; subwavelength optics; micro- and nano subwavelength waveguides; delay lines; plasmonic sensors; nanoplasmonic devices; nanophotonic structures; electromagnetic and optical metamaterials; cloaking devices; electric screening; supercapacitors; displays; projectors; light sources; nonlinear optical elements with second harmonic generation; extreme UV and X-ray applications.

9. Conclusion

Since nanomembranes are a novel concept which extends the range of MEMS & NEMS building blocks and practically introduces a new one, this means that whole branches of science and technology can be re-read and re-created through it, which may create an enormous number of novel applications. Nanomembranes need to be incorporated into coherent and ambitious programs of nanotechnology research, with aggressive funding and awareness-increasing campaigns. A care should be taken at that both about the fundamental and the applied aspects of research, since the recent developments clearly indicate that the field may have many promises and even surprises in stock.

A development of a novel technology or concept very rarely follows a smooth and gradual trend. Much more often one encounters an abrupt surge in development after the necessary conditions are met, not only scientific and technological, but also social and economic. In our opinion such is the situation with nanomembranes at the beginning of the 21st Century.

10. Acknowledgments

11. References

Aristov V.V., Kazmiruk V.V., Kudryashov V.A., Levashov V.I., Redkin S.I. Hagen C.W., and Kruit, P. (1998), Microfabrication of ultrathin free-standing platinum foils, *Surface Science* Vols. 402–404, pp. 337–340.

Bagliom, S; Castorina, S; and Savalli, S. (2007), *Scaling Issues and Design of MEMS*, John Wiley & Sons Ltd, ISBN 978-0-470-01699-2

Branton, D.; Gordon, R. G.; Chen, P.; Mitsui, T.; Farmer, D. B. and Golovchenko, J. (2005), Analysis of molecules by translocation through a coated aperture, *WIPO Patent* WO/2005/061373.

Carpenter F.E., Curico J.A, (1950), Preparation of Unbacked Metallic Films, *Review of Scientific Instruments*, Vol. 21, pp. 675- 676

Choi, W. M.; Song, J.; Khang, D.-Y.; Jiang, H.; Huang, Y. Y. and Rogers, J. A. (2007), Biaxially Stretchable "Wavy" Silicon Nanomembranes, *Nano Letters*, Vol. 7, No. 6, pp. 1655-1663.

Deshmukh, M. M.; Ralph, D. C.; Thomas, M. and Silcox, J. (1999), Nanofabrication Using a Stencil Mask, *Applied Physics Lett*ers Vol. 75, pp. 1631-1633.

Enke, D.; Friedel, F.; Janowski, F.; Hahn, T.; Gille, W.; Müller, R. and Kaden, H. (2002), Ultrathin porous glass membranes with controlled texture properties, *Studies in Surface Science and Catalysis*, Vol. 144, pp. 347-354.

Feng, S.; Elson, J. M.; Overfelt, P. L.; (2005) (1) Optical properties of multilayer metaldielectric nanofilms with all-evanescent modes," *Optics Express* Vol. 13, p. 4113-4124.

Feng, S.; Elson, J. M.; Overfelt, P. L.; (2005) (2) Transparent photonic band in metallodielectric nanostructures," *Physical Review B* Vol. 72, 085117

Fissell, W. H., IV; Humes, D. H.; Roy, S.; and Fleischman, A. (2006) Ultrafiltration membrane, device, bioartificial organ, and methods, *United States Patent* 20060213836

Friedmann, T.A. and Sullivan, J.P., (1997) Thick stress-free amorphous-tetrahedral carbon films with hardness near that of diamond, *Applied Physics Letters*, Vol. 71, No. 26, pp. 3820-3822.

Glotzer, S. C.; Solomon, M. J., and Kotov, N. A. (2004), Self-assembly: From nanoscale to microscale colloids, *AIChE Journal*, Vol. 50, No. 12, pp. 2978 - 2985.

Hamrock, S. J. and Yandrasits, M. A. (2006), Review: Proton Exchange Membranes for Fuel Cell Applications, *Journal of Macromolecular Science, Part C: Polymer Reviews*, Vol. 46, pp. 219-244

Hannink, R. H. J. (2006), *Nanostructure control of materials*, CRC, Woodhead Publishing Limited

Hunter W. R. (1982), Measurement of optical properties of materials in the vacuum ultraviolet spectral region, *Applied Optics* 2103, Vol. 21, No. 12, pp. 2103-2114.

Jayaraman, V.; Lin, Y. S.; Pakala, M. and Lin, R. Y. (1995), Fabrication of ultrathin metallic membranes on ceramic supports by sputter deposition", *Journal of Membrane Science* Vol. 99, No. 1, pp. 89-100.

Jiang, C.; Markutsya, S. & Tsukruk, V. V. (2004). Compliant, robust, and truly nanoscale free-standing multilayer films fabricated using spin-assisted layer-by-layer assembly. *Advanced Materials* 16, 157–161., ISSN: 0935-9648.

Kim, D.-H.; Ahn, J.-H.; Choi, W. M.; Kim, H.-S.; Kim, T.-H.; Song, J.; Huang, Y. Y.; Liu, Z.; Lu, C. and Rogers, J. A. (2008), Stretchable and Foldable Silicon Integrated Circuits, *Science* Vol. 320, pp. 507-511.

Liu, L.; Chakma, A.; and Feng, X. (2004), A novel method of preparing ultrathin poly(ether block amide) membranes, *Journal of Membrane Science*, Vol. 235, Nos. 1-2, pp. 43-52.

Lvov, Y.; Ichinose, I. and Kunitake, T. (1995), Assembly of Multicomponent Protein Films by Means of Electrostatic Layer-by-Layer Adsorption, *Journal of the American Chemical Society*, Vol. 117, pp. 6117-6122.

Murad, S. and Nitsche, L. C. (2004), The effect of thickness, pore size and structure of a nanomembrane on the flux and selectivity in reverse osmosis separations: a molecular dynamics study, *Chemical Physics Letters*, Vol. 397, Nos. 1-3, pp. 211-215.

Peterson, I.R. (1990), Langmuir Blodgett Films , *Journal of Physics D*, Vol. 23, No. 4, pp. 379-95.

Petrmichl, R. H. & Veerasamy, V. S. (2001), Hydrophobic coating with DLC & FAS on substrate, *United States Patent* 6312808.

Pientka, Z.; Brožová, L.; Bleha, M.; Puri, P. (2003), Preparation and characterization of ultrathin polymeric films, *Journal of Membrane Science*, Vol. 214, No. 1, pp. 157-161.

Random House Webster's Unabridged Dictionary, Random House Reference (2003).

Robertson J. (2002), Diamond-like amorphous carbon, *Material Science and Engineering* R Vol. 37, pp. 129-281.

Sackmann, E. and Tanaka, M. (2000), Supported membranes on soft polymer cushions: fabrication, characterization and applications, *Trends in Biotechnology*, Vol. 18, No. 2, pp. 58-64.

Schaldach, C. M.; Bourcier, W. L.; Paul, P. H. and Wilson, W. D. (2004), Electrostatic potentials and fields in the vicinity of engineered nanostructures, *Journal of Colloid and Interface Science*, 275 (2), pp. 601-611.

Sonin A. A. (1999), *Freely Suspended Liquid Crystalline Films*, John Wiley & Sons, ISBN-13: 978-0471971559

Stepanov, I.S.;, van Aken, R.H.; Zuiddam, M.R., and Hagen, C.W. ; Fabrication of Ultra-Thin Free-Standing Chromium Foils Supported by a Si_3N_4, Membrane-Structure with Search Pattern, *Microelectronic Engineering*, Vol. 46, No 1-4, pp. 435-438, ISSN 0167-9317.

Striemer, C. C., and Fauchet, P. M., (2006), Ultrathin nanoscale membranes, methods of making, and uses thereof, *United States Patent* 20060243655

Striemer, C. C.; Gaborski, T. R.; McGrath, J. L.; Fauchet, P. M. (2007), Charge- and size- based separation of macromolecules using ultrathin silicon membranes, *Nature* Vol. 445, pp. 749-753.

Tavolaro, A.; Tavolaro, P., and Drioli, E. (2007), Zeolite inorganic supports for BSA immobilization: Comparative study of several zeolite crystals and composite membranes, *Colloids and Surfaces B: Biointerfaces*, Vol. 55, No. 1, pp. 67-76.

Timoshenko, S. and Woinowsky-Krieger S. (1959), *Theory of Plates and Shells*, McGraw-Hill Co. 2 edition , ISBN-13: 978-0070647794

Toh, C. S.; Kayes, B. M.; Nemanick, E. J. and Lewis, N. S. (2004), Fabrication of Free-Standing Nanoscale Alumina Membranes with Controllable Pore Aspect Ratios, *Nano Letters* Vol. 4, pp. 767-770.

Tong, H. D.; Jansen, H. V.; Gadgil, V. J.; Bostan, C. G.; Berenschot, E.; van Rijn, C. J. M. and Elwenspoek, M. (2004), Silicon nitride nanosieve membrane. *Nano letters*, Vol. 4, No. 2, pp. 283-287.

Toutianoush, A. and Tieke, B. (2001), Ultrathin self-assembled polyvinylamine /polyvinylsulfate membranes for separation of ions", *Studies in Interface Science*, Vol. 11, pp. 415-425.

van Rijn, CJM (2004), *Nano And Micro Engineered Membrane Technology*, Elsevier, 2004, ISBN-13: 978-0-444-51489-9

Vendamme, R.; Onoue, S. Y.; Nakao, A.; Kunitake, T., (2006), Robust free-standing nanomembranes of organic/inorganic interpenetrating networks, *Nature Materials* Vol. 5, No. 6, pp. 494-501

Vincent, J. F. V. (2000), Smart by name, smart by nature, *Smart Materials and Structures*, Vol. 9, No. 3, pp. 255-259.

Watanabe, J. and Ishihara, K. (2008), Establishing ultimate biointerfaces covered with phosphorylcholine groups", *Colloids and Surfaces B: Biointerfaces*, Vol. 65, No. 2, pp. 155-165.

Winch, R. P., Photoelectric Properties of Thin Unbacked Gold Films, (1931). *Physical Review* Vol. 38, pp. 321- 324.

Yamauchi, A.; Shin, Y.; Shinozaki, M. and Kawabe, M. (2000), Membrane characteristics of composite collodion membrane: IV. Transport properties across blended collodion/Nafion membrane, *Journal of Membrane Science*, Vol. 170, No. 1, pp. 1-7.

Zhang, S. and Ali, N. (2007), *Nanocomposite Thin Films and Coatings: Processing, Properties and Performance*, Imperial College Press, London, ISBN-13: 978-1860947841.

Zhou, B; Wang, L ; Mehta, N ; Morshed, S.; Erdemir, A.; Eryilmaz, O. and Prorok, B. C. (2006), The mechanical properties of freestanding near-frictionless carbon films relevant to MEMS, *Journal of Micromechics and Microengineering*, Vol. 16, pp. 1374-1381

Nanomembrane-Enabled MEMS Sensors: Case of Plasmonic Devices for Chemical and Biological Sensing

Zoran Jakšić[1] and Jovan Matovic[2]
[1]Institute of Chemistry, Technology and Metallurgy Belgrade,
[2]Vienna University of Technology
[1]Serbia
[2]Austria

1. Introduction

The world is witnessing a rapidly increasing need for various types of sensing devices. The requirements are for smaller, smarter and more versatile sensors. The microelectromechanical (MEMS) devices represent a natural way to proceed along these lines and it is no wonder that the number of various device types, their complexity and the sheer number of various units are increasing at an accelerated pace (Gründler, 2007), (Martinac, 2008), (Merkoçi, 2009), (Toko, 2005).

Desirable features of a MEMS sensor include high sensitivity and selectivity, low noise, high robustness, long mean time between failures, small dimensions, characteristics adaptive to the widest possible range of operating conditions, possibilities of massively parallel multisensor operation and low cost. Possibilities of self-testing and even self-repair are also advantageous. Some of these requirements are contradictory and all of them set demanding challenges to the device designers.

Many of the desired properties are met in biological organisms which currently set the ultimate target in miniaturization, multiprocess operation and complexity. Thus one of the important today's paths of the MEMS development and further of the nanoelectromechanical systems (NEMS) goes toward biomimetics/bionics (Toko, 2005).

Another paradigm gaining momentum these days and also connected with biomimetics are artificial nanomembranes (Vendamme et al, 2006), (Watanabe et al, 2009), (Choi et al, 2007). These may be defined as engineered quasi-two-dimensional freestanding structures (the thickness being much smaller than their width and length and belonging to a range below 100 nm – whence the prefix "nano–"). Their natural counterparts, biological lipid bilayers are the most ancient and most omnipresent natural building blocks since they envelop all living cells which critically rely on them. Artificial nanomembranes are a product of MEMS technologies which are used to produce many of today's self-supported freestanding artificial nanomembranes (Watanabe & Kunitake, 2007), (Ni et al, 2005), (Li et al, 2007) with a thickness sometimes even reaching down to an atomic monolayer (Bunch et al, 2008). Very often the fabrication of nanomembranes includes the deposition of nanocomposite precursor over a sacrificial diaphragm (Mamedov et al, 2002). Self-supported nanomembranes

fabricated today often have giant aspect ratios, readily exceeding the value of 1,000,000 – e.g. (Matovic & Jakšić, 2009). Such dimensions make them a hybrid between micro and nanosystems, even between macroscopic systems and nanosystems, since their lateral dimensions may be of the order of centimeters, while the thickness remains nanometric. Nowadays they are seen as a building block for various MEMS systems (Vendamme et al, 2006) , (Jiang et al, 2004a).

Since a biological complexity is sooner or later expected to be reached by micro and nanosensors and at the same time nanomembranes are the basic natural building block, it is only obvious to merge these two concepts into a single one.

In this chapter we show how such a simple fusion of two paradigms may result (and actually is already resulting) in a large multitude and variety of results. This is happening in spite of the fact that the field of nanomembrane-enabled sensors itself is extremely young, the first papers starting to appear several years ago (Jiang et al, 2004a). Here we consider only the application of synthetic/engineered nanomembranes and exclude biological structures.

After a concise overview of some of promising uses of nanomembranes in microsensors generally, we concentrate to a single sensor type, that of chemical, biochemical or biological (CBB) sensors utilizing the effects of adsorption/desorption and the surface plasmon resonance (SPR) effect. We consider the possibility to use nanomembranes as a platform for long range surface plasmons. The role of self-supported ultrathin structures in improving coupling between propagating modes and surface-bound plasmons is also analyzed, as well as their application in SPR sensor selectivity boost.

2. MEMS sensors enhancement through nanomembranes

This Section shortly considers the use of nanomembranes in the enhancement of MEMS sensors generally. This includes various inertial, thermal and photonic devices (Gardner et al, 2001). Being ultrathin and ultra-lightweight and at the same time robust, nanomembranes appear essential for the miniaturization of sensors when scaling down from the microscale to the nanoscale.

In many MEMS sensors the basic method of signal readout is to use deflection of a freestanding elastic structural part. This part is typically a microcantilever, a microbridge or a miniscule diaphragm. For instance, in a piezoresistive pressure sensor the deflecting element is a micrometer-thick diaphragm with a built-in Wheatstone bridge. Applied pressure causes the diaphragm deviation from the equilibrium and thus changes the resistance of piezoresistors. Similar situation is encountered in various inertial MEMS sensors like accelerometers and inclinometers, where the membrane or bridge deflection is caused by the movement of an inertial mass. Another elastic part whose deflection is measured in applications is the microcantilever, well known as one of the basic building blocks in MEMS and NEMS. For instance, in scanning probe microscopy, which includes Atomic Force Microscopy and represents one of the basic techniques for characterization in nanotechnologies, it is the principal element, and the readout is often based on the optical lever principle.

In most of the mentioned situations either the elastic structural part is made relatively thick (the order of micrometers, which is the conventional approach) and with large lateral dimensions (millimeters) or it is made thinner and with smaller lateral dimensions. If one desires to fabricate a sensor array with a large number of elements which is at the same time as compact and as sensitive as possible, the latter appears obviously the better approach.

The ultimate in thickness of these building blocks is posed by the mechanical properties of the material itself, and the nanomembranes whose thickness can be of the order of several atomic or molecular monolayers certainly approach that limit.

Literature quotes the use of nanomembranes as the ultrathin freestanding structure to replace the conventional building blocks in deflection-based sensors (Jiang et al, 2004a). Among the obvious advantages of applying such a strategy are an increased sensitivity and a wider dynamic range. Various forms of micromachined freestanding ultrathin structures ensure much higher resonant frequencies than the conventional ones (extending into the GHz range).
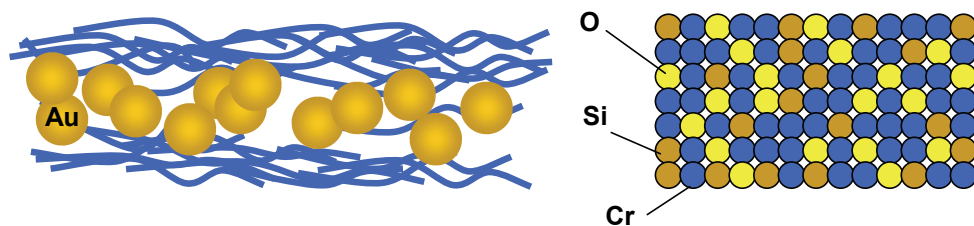
Fig. 1. Structures of composite nanomembranes convenient for mechanical and thermal sensors. Left: polymer matrix with gold nanoparticle filler. Right: metal-composite nanomembrane.

The nanomembranes for inertial sensors feature nanocomposites which may be e.g. polymer matrix filled with nanoparticles (Jiang et al, 2004b), metal-composite structures (Matovic & Jakšić, 2009), etc. Typically such structures are in a pre-stretched state. Their micromechanical properties can be readily adjusted by tuning the composition of the nanocomposite membranes. For instance, the amount of metal in a polymer-metal matrix will increase the elastic modulus of the nanocomposite. The measured elastic moduli for structures with gold nanoparticles were up to 10 GPa (Jiang et al, 2004a). The same structures can be obviously cut and formed in various ways and used in different shapes and with different anchorings as ultrathin microcantilevers and microbridges (Hua et al, 2004), (Zheng et al, 2002).

Some unique properties were observed in nanomembranes for MEMS sensors. Probably the most counter-intuitive one is their autorecovery feature, actually a mechanism of self-healing of overstretched structures (Jiang et al, 2004a). In our own experiments the metal-composite nanomembrane driven to the range of viscoelastic deformations did not remain distorted, but returned in a matter of tens of minutes to their original unstretched state. This property ensures a safety mechanism against accidental overstretching of nanometer-thin membranes and in final instance ensures a better stability of mechanical properties of inertial and pressure sensors based on nanomembranes, as well as a longer lifetime of such products.

The mechanical sensors based on nanomembranes also include acoustic imagers (Ballantine et al, 1997), (Kash, 1991). Acoustic sensitivities were reported at least an order of magnitude below the threshold of human hearing (Jiang et al, 2004a).

Another approach to sensing of mechanical movements is to utilize nanomembrane-based freestanding waveguides for evanescent field sensing. This was proposed for optical

measurement of deflection in micromirrors, gyroscopes, etc. and structures in the thickness range 30 nm to 100 nm were fabricated in Si_3N_4 (Altena, 2006).

Another large field of application of freestanding nanomembranes are thermal sensors (Kruse & Skatrud, 1997). The need for large area thermal arrays of miniature detectors in infrared technology and remote sensing is large (Rogalski, 2003), (Dereniak & Boreman 1996). Various thermal detectors include bolometers (Richards, 1994), pneumatic detectors/Golay cells (Golay, 1947), (Chévrier et al, 1995), microcantilever-based devices (Datskos et al, 2004) to which bimaterial detectors belong (Djuric et al, 2007), etc.

Thermal detectors are typically based on a large and thin absorbing area which reacts to thermal changes due to its irradiation by electromagnetic radiation and is sensitive in the whole electromagnetic spectrum. Nanomembranes obviously offer smaller thermal inertia and thus promise faster operation and higher specific detectivities. The assessments of polymer nanomembranes with gold nanoparticle fillers in thermal detectors show sensitivities several orders of magnitude higher than those for silicon membranes with the same diameter. For instance, temperature sensitivities below 1 mK were calculated for 55 nm thick, 200 μm diameter nanomembranes (Jiang et al, 2004a). Nanomembranes freely suspended over microfabricated cavities dedicated to infrared thermal detectors were reported in (Jiang et al, 2006).

A large field of application of nanomembranes in (nano)photonics is their use in enhancement of the operation of semiconductor infrared detectors (Rogalski, 2003). These detectors are actually quantum devices whose operation is based on generation of charge carriers in semiconducting material upon illumination in a given wavelength range. Their sensitivity spectrum is much narrower than that of thermal detectors and its cutoff frequency is determined by the bandgap of the given semiconductor. One of the fields of the application of nanomembranes in such detectors is the fabrication of resonant cavity structures, which may be implemented as multilayer dielectric mirrors or one-dimensional photonic crystals (Jakšić & Djurić, 2004), (Djurić et al, 1999), (Djurić et al, 2001). In addition to their application as the building blocks for the resonator reflectors, such freestanding structures may be applied in devices with tunable resonant frequency, where electrostatic field is used to deflect the membrane and adjust position to furnish the desired resonant peak (Ünlü & Strite, 1995).

Another field of application are both tunable and fixed filters for photodetectors in various wavelength ranges obtained by lamination of planar structures (Jakšić et al, 2005), (Maksimović & Jakšić, 2006). There is also a possibility to modify and tune the emissivity and absorptance by the application of such multilayers (Maksimović & Jakšić, 2005), up to the point of creating thermal antennas for visible and infrared radiation (Maksimović et al, 2008).

Finally, a large field of application of nanomembranes is in chemical, biochemical and biological sensors based on plasmon resonance. The rest of this Chapter is dedicated to this important topic.

3. Plasmonic sensors with ultrathin, freestanding films

3.1 CBB sensing systems

A chemical, biochemical or biological (CBB) sensor may be described as a device which generates an instrument- or observer-readable output proportional to the amount of a targeted chemical, biochemical or biological analyte in a given gaseous or liquid

environment. The output is most often electrical or optical. The most important issues regarding a CBB sensor are its sensitivity and selectivity towards a given analyte.

A general CBB sensing system (Fig. 2) consists of three main blocks, (1) the unit for separation/filtering and possibly reaction enhancement, (2) the detection unit – the main part of the sensor where the signal is generated and (3) the processing unit where signals are conditioned and communicated further.
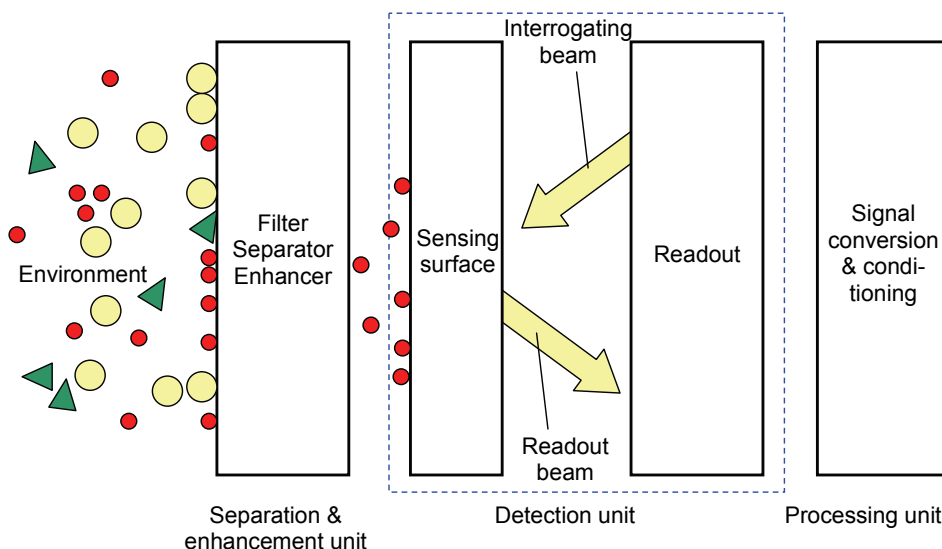
Fig. 2. Layout of a general CBB sensor consisting of (1) separation, filtering and enhancement unit; (2) detection unit and (3) signal conversion and conditioning unit.

We analyze the use of nanomembranes in the first two blocks. In the separation unit they are useful for filtering and generally molecular recognition if functionalized by nanopores, ion exchangers, absorbing fillers, etc., since their thickness enables a more accurate control of functionalization parameters than in larger structures. In the detection unit, especially of the kind used in nanoplasmonic devices, the nanomembranes are applicable as ultrathin, fully symmetric plasmon waveguides, strongly improving the device sensitivity.

3.2 Surface plasmon resonance sensors

Surface plasmons polaritons (SPP) are TM-polarized surface waves propagating along a metal–dielectric interface at optical frequencies (Fig. 3). Their wavelengths are extremely short and may even enter the X-ray range (Maier, 2007). The SPPs are evanescent perpendicularly to the active surface both toward the environment and toward the metal layer. In the situation when substrate and superstrate differ, the dispersion relation will allow two different modes of propagation, one on each interface (Maier, 2007).

Sensors based on the propagation of surface plasmon polaritons have become one of the most important tools in chemical, biochemical and biological sensing (Barnes et al, 2003), (Maier, 2007), (Homola, 2006), (Jung et al, 1998). They offer label-free, highly sensitive single-step measurements, real-time monitoring, require extremely small analyte samples (atomic/molecular monolayers suffice) and ensure a single generic framework for different

analytes. Multichannel devices are readily implemented in such configurations. No moving parts are required and the fabrication technology is simple – the conventional SPP resonance-based sensor is a planar metal surface with the plasma frequency in the wavelength range of interest. Good metals are used to this purpose, typically gold or silver. Being fully optical, these sensors are resistant to external electromagnetic disturbances. Finally, plasmon sensors are very convenient for miniaturization and the fabrication of ultracompact sensor arrays.
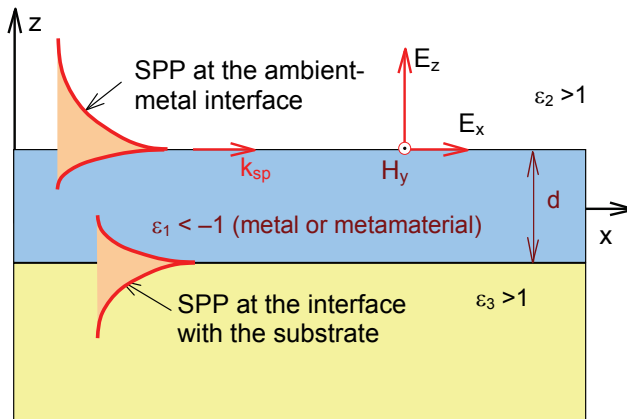
Fig. 3. Basic configuration of a guide for surface plasmon polariton propagation (metal-dielectric interface)

It is possible to use the same structure simultaneously for guiding SPP waves and for guiding the controlling electrical signals, since the active area is made of metal (Boltasseva et al, 2005). Also, the SPP components generally have high field localizations, thus promote the use of nonlinear photonic materials, ensuring the possibility for integration of active all-optical components (Zayats & Smolyaninov, 2003).

The operation of the SPP sensors is based on the modification of the propagation of surface plasmons polaritons at the sensor (metal)-environment (dielectric) interface. The analyte from the environment is bound either directly to the plasmonic surface, or (much more often) to a target-specific ligand layer. In both cases the surface refractive index is modified exactly in the position where the maximum of the SPP wave is located, since SPP waves are confined to the metal-dielectric interface and evanescent in perpendicular direction. In this way the maximum response is ensured. SPP resonance sensing is essentially thin film refractometry, where a change in the analyte concentration from c to $c + \Delta c$ causes a refractive index change at the metal-environment surface n to $n + \Delta n$ due to perturbed propagation conditions for the surface waves.

The obvious idea here is to use a nanomembrane as a waveguide for plasmons. Since a surface plasmon polariton is a quasi-planar electromagnetic wave decaying evanescently in both perpendicular directions, it is logical to utilize as a support for it a metal or metal-composite nanomembrane which is also quasi-planar.

Plasmons in nanomembranes with metal fillers were reported in (Jiang et al, 2004b), where gold nanoparticles were used in a polymer matrix and the packing density of the gold spheres varied from below 2% to about 25%. Experimental structures are typically light blue due to a plasmon resonance peak corresponding to the plasma frequency in visible.

3.3 Long range SPP sensors utilizing nanomembranes

One of the problems in structures with conventional SPP is large signal attenuation, a consequence of a high imaginary part of the propagation constant due to the ohmic losses/absorption in metals (Zayats & Smolyaninov, 2003). In such waveguides the propagation length are typically limited to a range from tens (visible range) to hundreds of micrometers (near infrared). Another problem is their coupling with propagating modes, since typically elaborate schemes using e.g. prism couplers or diffractive gratings must be used.

The way to overcome most of these shortcomings is to use long-range (LR) surface plasmon polaritons (Sarid, 1981), (Burke et al, 1986), (Charbonneau et al, 2000), (Berini, 2000). These are SPPs which propagate along metal strips with nanometric thickness (typically 10-40 nm) immersed into dielectric.
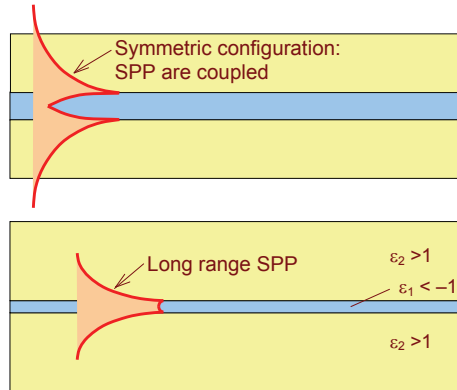
Fig. 4. Generation of long-range surface plasmons polaritons through coupling of top and bottom modes on ultrathin metal sheets; top: metal guide is surrounded from both sides with identical dielectric; bottom: metal sheet is smaller than the decay length and LR plasmon appears

In the case when the substrate and the superstrate are described by identical permittivity (the case of full immersion of the metal sheet in homogeneous dielectric), the structure is symmetrical in electromagnetic sense. The two propagating modes on the top and bottom surface then couple and propagate together, Fig. 4 top. If the metal sheet between the two identical media is sufficiently thin to make the interaction between the top and the bottom SPR non-negligible, these two modes couple and merge into a single one, Fig. 4 bottom. The degeneracy for that mode is then removed and its dispersion splits into two branches, one for low-frequency mode (odd), and the other for high-frequency mode (even). The even modes have a very short propagation path. The propagation constant of the odd modes decreases, being proportional to the square of the film thickness. This means that the attenuation of the odd mode will be very low and thus its propagation length large. Thinner films and more symmetrical structures will have longer propagation paths.

A typical trait of an LR SPP is that its fields are mostly contained outside of the metal part. Since the field concentration is much lower in the metal sheet, the propagation losses are consequently also much lower.

The imaginary part of their propagation constant being approximately zero, the LR SPP ensure much larger propagation paths, typical propagation losses being below 6 dB/cm (Boltasseva et al, 2005).

Long-range surface plasmon sensors are especially convenient for biological sensors, since the confinement of the plasmon waves is smaller than in other SPP devices and thus the larger biological samples are more easily encompassed (Berini et al, 2008)

Probably the most important cause of the signal attenuation in LR SPP structures is its deviation from symmetry (Park & Song, 2006). Fig. 5 shows a calculated curve of attenuation for a metal nanomembrane immersed in dielectric.
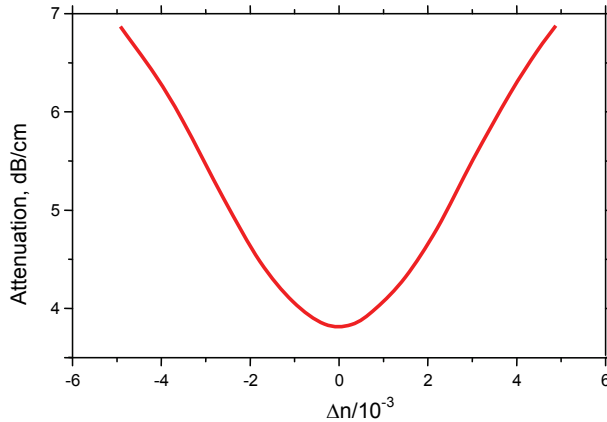
Fig. 5. Calculated LR-SPP propagation loss versus asymmetry of dielectric given as the refractive index difference. Membrane thickness 12.5 nm, material gold, refractive index of dielectric immersion 1.5, wavelength 1.55 μm.

It is visible that even very small deviations from symmetry introduce large losses into the waveguide.

The use of metal or metal-composite nanomembranes at the same time gives a platform for LR SPP and ensures its complete symmetry. Their thickness is typically from 4 nm up, thus very low losses are ensured. A layout of a nanomembrane-based LR SPP guide is shown in Fig. 6. The structure itself is extremely simple, being a freestanding planar nanomembrane sheet.
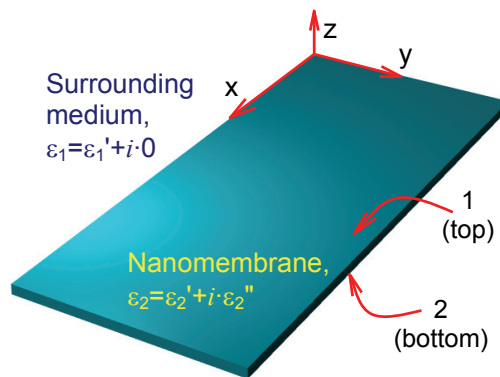
Fig. 6. Basic configuration of a freestanding nanomembrane guide for long-range surface plasmon polariton propagation (metal-dielectric interface)

The issue of coupling between the propagating modes and the plasmon waveguide is dealt with further in this text.

3.4 Detection limits and novel structures

One of the problems with probably all types of sensors are their ultimate limits of detection, which are connected with various extrinsic and intrinsic mechanisms of noise. Of these, the latter ones include mechanisms that are fundamental to the sensing process itself. In the case of plasmonic sensors, such fundamental mechanisms include adsorption-desorption noise which is connected with the operation of the SPR devices themselves, thermal (Johnson-Nyquist) noise, 1/f noise and zero-point noise (noise due to quantum fluctuations) (Jakšić et al, 2007), (Jakšić et al, 2009a). It is interesting to note that at least some of these noise sources are expected to affect the operation of nanomembrane-based SPR sensors less than that of other types of plasmonic sensors. For instance, it is expected that the adsorption-desorption noise will be smaller in nanomembrane-based long-range plasmon sensors than in other types, since this noise will decrease with increasing the active detection area (Jakšić, O. et al, 2009). Zero-point noise should also decrease in the case of LR SPR devices.

One of the ways to shift the ultimate detection limits and at the same time to ensure new degrees of freedom in sensor design is to utilize novel structures, optimized for higher sensitivities and lower noise. A possible pathway is to pattern or shape the nanomembrane surfaces, for instance by focused ion beam patterning (Gierak et al, 2007). A large opportunity window opened by the advent of nanoplasmonics (Maier, 2007), and especially with the introduction of electromagnetic metamaterials (Pendry et al, 1999). Such structures may be defined as artificial structures with electromagnetic response not readily found in nature. A typical and well-known type metamaterials are the structures with negative value of refractive index (Veselago, 1968), also known as left-handed structures (Ramakrishna & Grzegorczyk, 2009), thus named because the triplet electric field vector, magnetic field vector and wavevector form a left-handed set, contrasted to the "normal" materials where this set is always right-handed. Patterned and laminar metal-dielectric nanomembranes are a useful building block for quasi-2D metamaterial structures, the metasurfaces, intended for the operation in the optical range. Actually the metal nanomembranes themselves may be regarded as left-handed metamaterials in certain situations, since some electromagnetic modes propagating on them show the properties of negative effective refractive index (Smolyaninov, 2008). Generalized plasmonic sensors based on left-handed metamaterials were described in various references (Ishimaru et al, 2005), (Jakšić et al, 2007), (Bingham et al, 2008).

4. Freespace coupling with interrogating beam

An important issue in plasmonic sensors, regardless of the active surface type, is their coupling with light sources and the readout systems, i.e. the matching of propagating planar waves of optical radiation with evanescent SPP waves. The wavevector of SPP is always larger than the wavevector in free space and at optical frequencies the wavelengths of the SPP may become very small, even reaching nanometric lengths (Maier, 2007), (Raether, 1988), (Barnes et al, 2003). Thus it is necessary to impart the missing momentum to the interrogating beam (propagating planar wave) in order to enable coupling – i.e., to ensure phase matching between the two waves.

In coupling it is important to ensure that the maximum percentage of the incoming free-space mode is converted to SPP (and vice versa for the output). At the same time, it is important to ensure the smallest leakage and scattering losses. There are various schemes to ensure coupling between plasmonic devices and propagating modes. They may be roughly divided into four groups: prism couplers, endfire couplers, near-field probe couplers and those utilizing topological surface defects.

Historically the oldest methods are those utilizing prism couplers (Fig. 7). These include the Kretschmann configuration (Kretschmann & Raether, 1968) (Fig. 7a) which is still the prevailing readout method in plasmon sensors, as well as the Otto coupler (Otto, 1968) (Fig. 7b). Both of these methods utilize attenuated total reflection. Another method to excite the SPP is to use end-fire coupling, where the incident beam is in plane with the plasmonic surface (Fig. 7c) (Stegeman et al, 1983), (Berini et al, 2007).
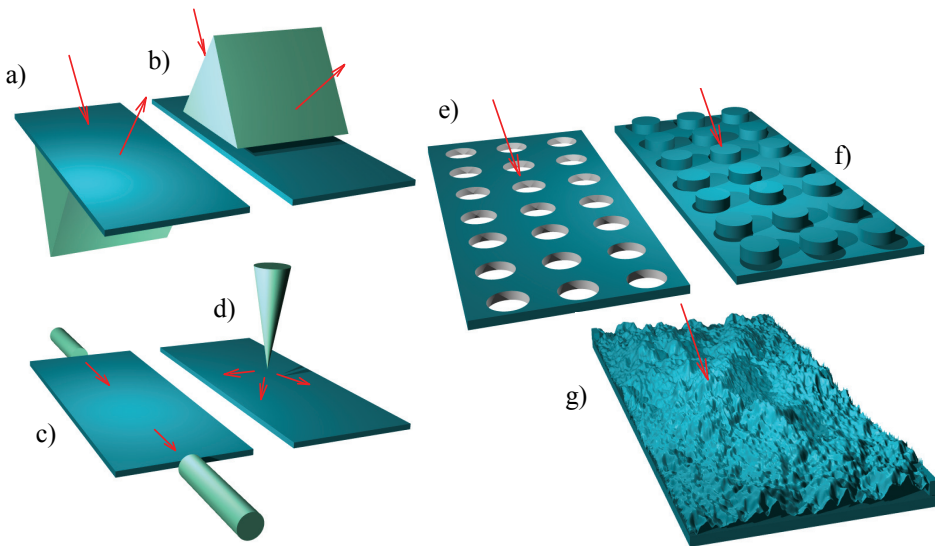
Fig. 7. Couplers plasmon-propagating a) Prism couplers in Kretschmann configuration;
b) Prism couplers in Otto configuration; c) end-fire coupling; d) near-field probe excitation;
Various methods of coupling through topological surface defects which may consist of
e) gratings consisting of nanohole arrays, f) surface protrusions or may be g) disordered
surface corrugations.

An important group of couplers utilize various near-field probes (the use of the "forbidden light' outside the light cone), (Fig. 7d) where local excitation in evanescent field is utilized and the beams tunnel from the impingement point to the metal-dielectric interface which supports SPP (Hecht et al, 1996), (Bouhelier & Novotny, 2007), (Maier et al, 2004).

Finally a large and very important group are couplers utilizing topological surface defects (Barnes et al, 2003), (Ritchie et al, 1968). These include grating couplers which may consist of periodic arrays of either subwavelength apertures (Fig. 7e) (Devaux et al, 2003) or surface protrusions (e.g. various pillars, bumps, etc.) (Worthing & Barnes, 2001), Fig. 7f. The arrays may be 2D like those shown in Fig. 7e, f) or 1D (grooves or stripes) and may have various shapes, e.g. rectangular, triangular, wavy, etc.

The couplers may be also disordered (this layout may be understood as a superposition of a large number of gratings with different periods) – Fig. 7g. (Ditlbacher et al, 2002)

In the case of freestanding nanomembranes and LR SPP sensors it is important to couple these structures with propagating modes with the least disturbance to the symmetry, thus preferably without a direct physical contact with the nanomembrane. One could use the shaping of a dielectric substrate (which, however, would perturb the electromagnetic symmetry of the structure), endfire coupling (which introduces alignment and coupling efficiency issues; it is known that the percentage of coupled light in this method is extremely low) or Otto prisms (bulky structure which makes the device significantly more complex).

We proposed an alternative approach which uses direct sculpting of the nanomembrane and is applicable without special alignment procedures (Jakšić et al, 2009). The idea of our approach is to incorporate the coupling structures into the freestanding nanomembrane itself, without any substrate to hold them. In this way the substrate and the superstrate remain fully index matched throughout the measurement. At the same time, the structure remains generally applicable, since the analyte does not have to be matched to the pre-fabricated device substrate.

The sculpted structures are small perturbations of the much larger nanomembrane, their dimensions being of the order of micrometers, while the membrane dimensions are measured in millimeters, even centimeters.

The surface is sculpted into a 2D array of protrusions (Fig. 8 a) which serve as a coupling diffractive grating (Kashyap, 1999). The basic approach to nanomembrane sculpting is illustrated in Fig. 8 b, c.
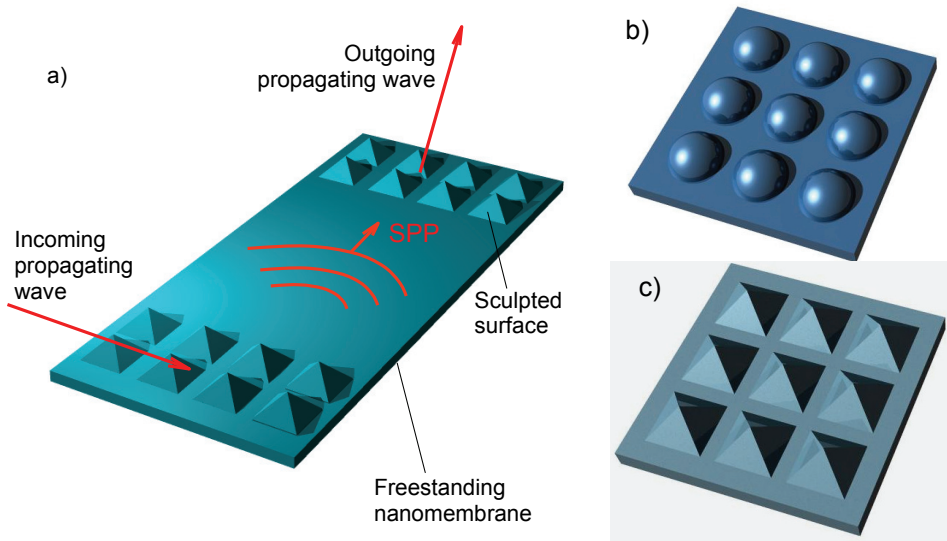
Fig. 8. a) Propagating wave to surface plasmon couplers using surface sculpting. b) Drawing of hemispherical surface relief for nanomembrane sculpting fabricated by isotropic etching through circular openings in photolithographic mask; c) Drawings of pyramidal surface relief for nanomembrane sculpting fabricated by anisotropic etching of silicon with (100) surface orientation through square windows aligned along [110].

The fabrication of surface-sculpted protrusion arrays is based on the deposition of a nanometric membrane precursor over a sacrificial layer (Mamedov et al, 2002), (Jiang et al, 2004b) and its subsequent release in etching solution, the same method used to produce metal-composite nanomembranes. The difference is that prior to depositing the membrane precursor, one first etches an array of micrometer pits in the sacrificial structure. The layout of the pits is determined by the applied photolithographic mask, thus defining the diffractive grating for the coupling. The shape of the pits themselves is defined by the chosen etching method. If isotropic etching is used, the pits are hemispherical or ellipsoidal. If anisotropic etching is utilized, one can produce for instance pyramids or truncated pyramids. The membrane precursor is subsequently deposited over the pits and upon release the fabricated nanomembrane retains the shape of the pits.

If the sacrificial layer pits are etched by isotropic etchant, the final protrusion are hemispherical, Fig. 8b. Variations to this generic form can be obtained by using different shapes of photolithographic masks (for instance, elliptical openings, but also various polygons) and by adjusting the etching duration to obtain either flatter or more voluminous structures.

The structure in Fig. 8c is obtained in an analogous manner, but using anisotropic etching of single crystalline silicon sacrificial layer with (100) surface orientation through square windows aligned along [110]. The variations to this generic form includes truncated pyramids and actually all standard forms obtainable by anisotropic etching.

Figure 9 shows the fabricated 15 µmx10 µm pyramid sculpted on the surface of a metal-composite nanomembrane with a thickness of 20 nm. It is known that nanomembranes become intrinsically stretched during their low-temperature annealing (Matovic & Jakšić, 2009) and thus the sculpted surface features retain their shape in spite of the minute thickness of their walls.
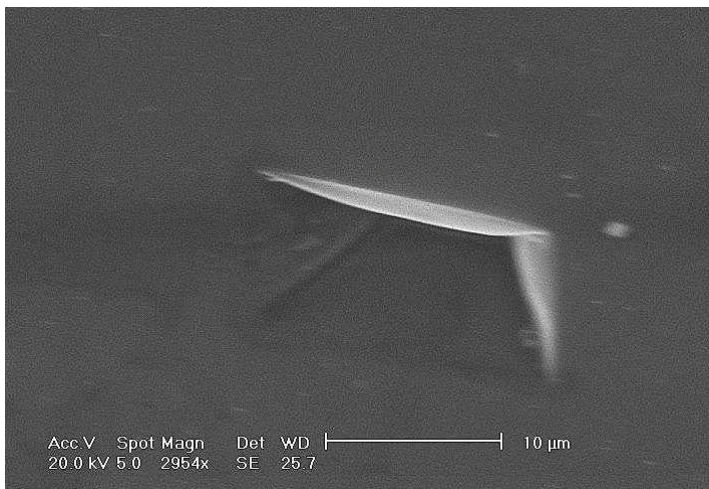
Fig. 9. Scanning electron microscope photo of a 15 µmx10 µm pyramid sculpted on the surface of a metal-composite nanomembrane, thickness 20 nm

We believe the tailorability makes the 3D surface-sculpted nanomembranes a valid alternative to other coupling methods for freestanding LR SPP sensor structures.

5. Selectivity enhancement through nanomembranes

The issue of selectivity is of the utmost importance for the CBB sensors, since among the myriad of possible analytes, especially the organic and biological ones, many are found with almost identical properties. The basic behavior of a plasmonic CBB sensor is determined by the adsorption/desorption properties of a given analyte toward the sensor active surface. The conventional wisdom in traditional CBB sensors is to deposit a ligand layer onto the sensor surface to adsorb the targeted analyte only (e.g. a particular DNA sequence, a desired enzyme, etc.) This is a well developed approach and is obviously also applicable for the nanomembrane-based (nano)plasmonic CBB sensors.

In this Section we turn attention to the fact that the use of nanomembranes opens additional possibilities to improve the selectivity of CBB sensors. At the same time these additional methods practically do not change the dimensions of the sensors. Probably the most obvious approach is to use nanomembranes as filtering/separation units.

It is well known that molecular separation is among the basic applications of conventional membranes (Yampolskii et al, 2006), (Böddeker, 2007), (Hoffman, 2004), (Sata, 2004), and they form the basis of the large separation science. It is the common knowledge that thinner membranes typically mean higher throughput and thus shorter reaction times. This property is of great importance for the CBB sensors, as very often they are required to operate in real time. Isotropic conventional membranes are usually at least tens of micrometers thick. Anisotropic membranes, including those with multistage hierarchical structure (Fendler, 1994) have much thinner active layers, but require thick porous substrates as supports. Only with synthetic nanomembranes the possibility appeared to have structures with a thickness in the nanometer range and at the same time robust enough not to require any additional supports. Thus for the first time one simultaneously obtains nanometric thickness, mechanical strength and a high throughput. At the same time they offer larger possibilities for nano-customization since one can tailor their properties with larger precision: it is always easier to keep for instance nm-dimensioned pore aspect ratio through a cross-section with nanometric thickness, than to keep it in a structure 1,000 or 10,000 times thicker.

In this Section we consider the basic ways to improve selectivity of nanomembrane-based plasmonic CBB sensors using structures similar to the ones which comprise the sensors themselves.

There are four main mechanisms of membrane-enabled separation (Baker, 2004) which can be implemented by nanomembranes utilizing some of the available methods for their functionalization (Fig. 10):

1. Pore-flow molecular sieving, where separation is basically done by passage of molecules through a system of randomly distributed pores which are at the same time interconnected. Thus such a structure operates in a manner similar to a conventional particle filter: molecules larger than the pore size will be rejected, while smaller ones will pass through the nanomembrane. Here thinner membranes will obviously offer a possibility for a better control of pore size and shape and at the same time enable a higher throughput. Such membranes may have an isotropic or anisotropic pore distribution across the cross-section. Fig. 10a shows a molecular sieve with anisotropic distribution of pores, the Loeb-Sourirajan structure.
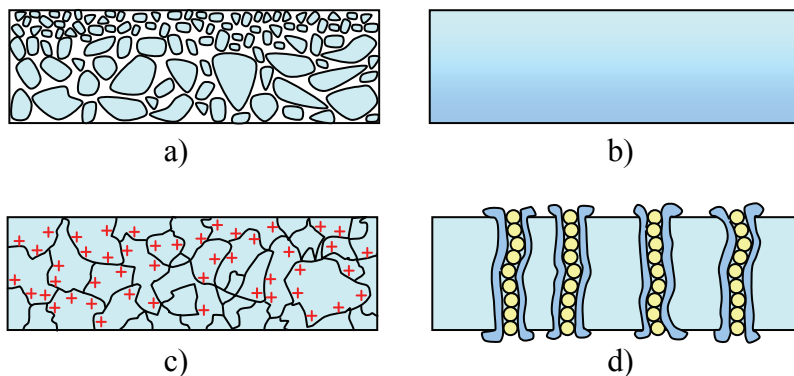
Fig. 10. Schematic presentation of selectivity enhancement through nanomembrane-enabled separation. a) Cross-section of nanomembrane with nanopores (Loeb-Sourirajan anisotropic structure); b) nonporous dense nanomembrane utilizing solution-diffusion mechanism; c) electrically charged membrane (ion exchanger); d) nanomembrane with gated ion channels.

2. Solution-diffusion through dense membranes where pores either do not exist or are at least smaller than the particle effective cross-section as defined by their thermal motion at a given temperature. The transport through such nanomembranes is a combination of particle solution and their diffusion across the structure, driven by concentration gradient, electrostatic field or pressure (Fig. 10b). These membranes can also be isotropic or anisotropic. They are used e.g. in pervaporation, reverse osmosis and often are regarded the method of choice in gas separation.

3. Ion exchange, where nanomembranes contain electrically charged particles (Fig. 10c). Most often such membranes are nanoporous, although they also may be dense. Positive or negative ions are fixed throughout the nanomembrane, usually at the pore walls. They bind the opposite ions and exclude the same charge. Structures with positive ions are denoted as anion exchangers, and those with negative ones are cation exchangers. This is the mechanism encountered in electrodialysis.

4. Gated ion channel flow, where the nanomembrane includes an ion-transmitting channel, typically consisting of proteins (Fermini & Priest, 2008) but which generally may be fabricated in various organic and inorganic materials. The channel is gated by external stimuli, typically by applied voltage or by ligands, which control the ion transport through the channel, Fig. 10d. This transport is highly selective. For instance, aquaporin proteins allow the flow of water molecules but are impermeable to protons, although these are much smaller than the water molecules. In other nanochannels transport of protons occurs, while the nanomembrane remains impermeable to other molecules (the Grotthuss mechanism, or proton hopping, where protons "hop" along a one-dimensional chain of water molecules – the "water wire") (Chung, 2007). Various ion channels exist for sodium, potassium ions, hydrogen, etc. This mechanism is fundamental for the ion transport through lipid bilayer nanomembranes in biological cells and represents the basis of the life on earth, while its biomimetic counterparts ensure another route to biosensor selectivity enhancement (Martin, 2007).

One can see that artificial pores are important for most of the described approaches. Nanopores are used to characterize DNA and RNA (Kasianowicz et al, 1996), up to the point of discriminating molecules differing by a single nucleotide (Vercoutere, 2003).

It is necessary to discern between different nomenclatures regarding the pore size, which may be the source of some confusion. According to the IUPAC, the term macropores is used to denote pores larger than 50 nm, mesopores have diameters between 2 nm and 50 nm and micropores are smaller than 2 nm (Rouquerol et al, 1993). In many literature sources all pores with diameters below 100 nm are termed nanopores (Aksimentiev et al, 2009).

There are several approaches to producing biomimetic pore complex capable of selectively transporting various biological analytes. Recently, an artificial scaffold for the nuclear pore complex-based gate was produced using natively disordered proteins termed FP nups. Such synthetic scaffold can be used as a generic platform for ultra-selective biomimetic nanopores (Jovanovic-Talisman et al, 2009). Another pathway is to avoid proteinaceous nanopores and to utilize other materials, including inorganic ones. Some materials used include silicon, silicon nitride, polyethylene terephthalate, silicon dioxide and many others (Aksimentiev et al, 2009). The design freedom offered by these approaches may lead to nanomembrane separators operating under less restrictive ranges of external parameters (temperature, electrolyte concentrations, pressures, pH values, bias, etc.). It is possible to tailor nanopores to practically any desired size, which means the freedom to optimize the pore geometry for various targeted analytes.

An important question is the method of combining the different parts of a CBB sensing device into a unified system. The unit for separation/enhancement may be integrated with the detection unit in various ways (Fig. 11). One may use two (or more) nanomembranes as separate elements so that the analyte-containing fluid is flowing sequentially through them, as shown in Fig. 11a. A modification of this approach is to utilize lamination of two (or more) separate nanomembranes into a single one, Fig. 11b. Finally, it is possible to aggregate detection and separation/enhancement functions into a single monolithic structure with one or more different active fillers which will perform affinity capture of the analyte and at the same time ensure tuning of the readout beam, Fig. 11c.
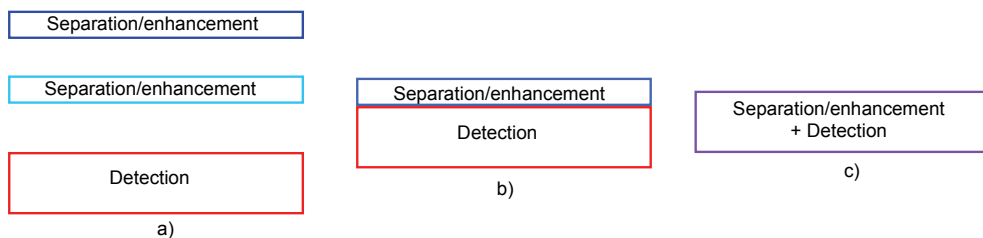
Fig. 11. Basic configurations of nanomembrane units for separation, enhancement and detection. a) The "separate separator" configuration where filtering structures are physically separated from the plasmonic waveguide; b) laminated structure, where separator/enhancer captures analyte; examples include e.g. solution/diffusion membranes, but also conventional ligand layers; c) aggregated/monolithic multipurpose unit.

6. Conclusion

The combination of the two paradigms, that of MEMS CBB sensors and that of nanomembranes, not only has potentials to vastly improve the performance of today's devices but also to introduce numerous completely new functionalities. A very convenient platform to this purpose are LR SPR waveguiding devices. They are useful for larger

biological samples because of their nanometric thickness and full electromagnetic symmetry. At the same time they offer the potential of lower adsorption-desorption noise due to larger effective surfaces. Their applications are many, from process industry and transportation, through medicine and life science, to environmental protection and antiterrorist defense.

A higher degree of design freedom is obtained because of the possibility to combine various organic and inorganic materials and at the same time to use biomimetic solutions learning from the membranes surrounding living cells, for instance highly selective gated ion channeling through nanopores. The possibility to functionalize nanomembranes through lamination, introduction of various active fillers and nanopatterning actually adds additional possibilities to those met in nature.

Further, the application of nanoplasmonics in this context opens a wide field for sensors which can be seen as a generalization of the already very potent and widely utilized SPR devices. Many novel types can be expected there, especially those which utilize the new class of nanoplasmonic structures, the electromagnetic metamaterials.

One may safely say that the introduction of the artificial analogue to the basic biological building block, the nanomembrane, and its combination with much wider range of materials and structures than the naturally occuring ones, opens an almost unlimited field for further research of micro and nanosensors.

Finally, one should mention that the described nanostructures obviously have a much wider applicability besides sensing, and include such fields as catalysis, biointerfaces, separation, microreactors, various environmental and biomedical applications and many more.

7. Acknowledgments

8. References

Aksimentiev, A.; Brunner, R. K.; Cruz-Chu, E. & Comer, J. & Schulten, K. (2009). Modeling transport through synthetic nanopores, *IEEE Nanotechnology*, Vol. 3, 20-28, 1536-125X

Altena, G. (2006). *Evanescent field sensing by Si_3N_4 nanomembranes*, PhD Thesis, University of Twente, Enschede.

Baker, R. W. (2004). *Membrane technology and applications*, 2nd ed, John Wiley & Sons Ltd, Chichester, ISBN: 978-0-470-85445-7

Ballantine, Jr., D. S.; White, R. M.; Martin, S. J.; Ricco, A. J.; Frye, G. C.; Zellers, E. T. & Wohltjen, H. (1997). *Acoustic Wave Sensors – Theory, Design, and Physico-Chemical Applications*, Academic Press, San Diego, ISBN: 0120774607.

Barnes, W. L.; Dereux, A. & Ebbesen, T. W. (2003). Surface plasmon subwavelength optics, *Nature*, Vol. 424, No. 6950, 824–830, ISSN 0028-0836.

Berini, P. (2000). Plasmon-polariton waves guided by thin lossy metal films of finite width: Bound modes of symmetric structures, *Physical Review B*, Vol. 61, 10484–10503, ISSN: 1098-0121.

Berini, P.; Charbonneau, R. & Lahoud, N. (2007). Long-Range Surface Plasmons on Ultrathin Membranes, *Nano Letters*, Vol. 7, No. 5, 1376 -1380, ISSN: 1530-6984.

Berini, P.; Charbonneau, R. & Lahoud, N. (2008). Long-Range Surface Plasmons Along Membrane-Supported Metal Stripes", *IEEE Journal of Selected Topics in Quantum Electronics* Vol. 14, No. 6, 1479-1495, ISSN: 1077-260X.

Bingham, C.M.; Tao, H.; Liu, X.; Averitt, R.D.; Zhang, X. & Padilla, W.J. (2008). Planar wallpaper group metamaterials for novel terahertz applications. *Optics Express*, Vol. 16, 18565-18575, ISSN: 1094-4087.

Boltasseva, A. (2004). *Integrated-Optics Components Utilizing Long-Range Surface Plasmon Polaritons*, PhD Thesis, Technical University of Denmark.

Boltasseva, A.; Nikolajsen, T.; Leosson, K.; Kjaer, K.; Larsen, M. S. & Bozhevolnyi, S. I. (2005). Integrated optical components utilizing long-range surface plasmon polaritons, *Journal of Lightwave Technology*, Vol. 23, No. 1, 413–422, ISSN: 0733-8724.

Böddeker, K. W. (2007). *Liquid Separations with Membranes: An Introduction to Barrier Interference*, Springer-Verlag Berlin Heidelberg, ISBN-13 978-3-540-47451-7

Bouhelier, A. & Novotny, L. (2007). Near-field optical excitation and detection of surface plasmons, in *Surface Plasmon Nanophotonics*, Brongersma M.L. and Kik, P.G. (eds.), pp. 139–153, Springer Verlag, Berlin, ISBN: 978-1-4020-4349-9.

Bunch, J. S.; Verbridge, S. S.; Alden, J. S.; van der Zande, A. M.; Parpia, J. M.; Craighead, H. G. & McEuen, P. L. (2008). Impermeable Atomic Membranes from Graphene Sheets, *Nano Letters*, Vol. 8, No. 8, 2458-2462, ISSN: 1530-6984.

Burke, J. J.; Stegeman, G. I. & Tamir, T. (1986). Surface-polariton-like waves guided by thin, lossy metal films," *Physical Review B*, Vol. 33, No. 8, 5186–5201, ISSN: 1098-0121.

Charbonneau, R.; Berini, P.; Berolo, E.; & Lisicka-Shrzek, E. (2000). Experimental observation of plasmon-polariton waves supported by a thin metal film of finite width, *Optics Letters*, Vol. 25, 844–846, ISSN: 0146-9592.

Chévrier, J.-B.; Baert, K.; Slater,T. & Verbist, A. (1995) Micromachined infrared pneumatic detector for gas sensor, *Microsystem Technologies*, Vol. 1, No. 2, 71-74, ISSN 0946-7076

Choi,W. M.; Song, J.; Khang, D.-Y.; Jiang, H.; Huang, Y. Y.; & Rogers, J. A. (2007). Biaxially Stretchable "Wavy" Silicon Nanomembranes, *Nano Letters*, Vol. 7, No. 6, 1655-1663, ISSN: 1530-6984.

Chung, S.H. (2007). *Biological Membrane Ion Channels: Dynamics, Structure, and Applications*, Springer Verlag, Berlin Heidelberg, ISBN:0387333231.

Datskos, P. G.; Lavrik, N. V. & Rajic, S. (2004). Performance of uncooled microcantilever thermal detectors, *Review of Scientific Instruments*, Vol. 75, 1134-1148, ISSN: 0034-6748

Dereniak E. L. & Boreman G. D. (1996). *Infrared Detectors and Systems*. Wiley-Interscience, ISBN: 978-0471122098

Devaux, E.; Ebbesen, T. W.; Weeber, J.-C.& Dereux, A. (2003). Launching and decoupling surface plasmons viamicro-gratings. *Applied Physics Letters*, Vol. 83, No. 24, 4936–4938, ISSN: 0003-6951.

Ditlbacher, H.; Krenn, J. R.; Félidj, N.; Lamprecht, B.; Schider, G.; Salerno, M.; Leitner, A. & Aussenegg, F. R. (2002). Fluorescence imaging of surface plasmon fields. *Applied Physics Letters*, Vol. 80, No. 3, 404–406, ISSN: 0003-6951.

Djurić, Z.; Jakšić, Z.; Randjelović, D.l Danković, T.; Ehrfeld, W. & Schmidt, A. (1999). Enhancement of Radiative Lifetime in Semiconductors Using Photonic Crystals, *Infrared Physics & Technology*, Vol. 40, No. 1, 25-32, ISSN: 1350-4495.

Djurić, Z.; Jakšić, Z.; Ehrfeld, W.; Schmidt, A.; Matić, M. & Popović, M. (2001). Photonic Crystal Enhancement of Auger-Suppressed Detectors: A Way to Background-Limited Room-Temperature Operation in 3-14 Micrometer Range, in *Nanoscale Linear and Nonlinear Optics*, Eds. M. Bertolotti, C. M. Bowden, C. Sibilia, *AIP Proceedings*, Melville, New York, Vol. 560, 418-424, ISBN 1-56396-993-9

Djurić, Z., Randjelović, D., Jokić, I., Matović, J. & Lamovec, J. (2007). A new approach to IR bimaterial detectors theory, *Infrared Physics & Technology* Vol. 50, No. 1, 51-57, ISSN: 1350-4495

Fendler, J.H. (1994). *Membrane-Mimetic Approach to Advanced Materials*, Springer-Verlag Berlin Heidelberg, ISBN: 0387572376

Fermini, B. & Priest, B. T. eds. (2008). *Ion Channels*, Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-79728-9

Gardner, J. W.; Varadan, V. K. & Awadelkarim, O. O. (2001). *Microsensors, MEMS, and Smart Devices*, John Wiley & Sons, ISBN-13: 978-047186.

Gierak, J.; Madouri, A.; Biance, A.L.; Bourhis, E.; Patriarche, G.; Lucot, C. U. D.; Lafosse, X.; Auvray, L.; Bruchhaus, L. & Jede, R. (2007). Sub-5 nm FIB direct patterning of nanodevices, *Microelectronic Engineering* Vol. 84, 779–783

Golay, M.J.E. (1947) Theoretical consideration in heat and infra-red detection, with particular reference to the pneumatic detector. *Review of Scientific Instruments* Vol. 18, 347–356, ISSN: 0034-6748

Gründler, P. (2007). *Chemical Sensors: An Introduction for Scientists and Engineers*, Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-45742-8

Hecht, B.; Bielefeld, H.; Novotny, L.; Inouye, Y. & Pohl, D. W.: (1996). Local excitation, scattering, and interference of surface plasmons. *Physical Review Letters*, Vol. 77, No. 9, 1889–1892, iSSN: 0031-9007.

Hoffman, E. J. (2004). *Membrane Separations Technology: Single-Stage, Multistage, and Differential Permeation*, Elsevier, Amsterdam, ISBN: 0750677104

Homola, J. Ed. (2006). *Surface Plasmon Resonance Based Sensors*, Springer, ISBN: 978-3-540-33918-2, Berlin-Heidelberg, ISBN: 978-3-540-33918-2.

Hua, T., Cui, T. & Lvov, Yu. M. (2004). Ultrathin cantilevers based on polymer-ceramic nanocomposite assembled through LbL adsorption, *Nano Letters* Vol. 4, 823–825, iSSN 1530-6984.

Ishimaru, A.; Jaruwatanadilok, S. & Kuga, Y. (2005). Generalized surface plasmon resonance sensors using metamaterials and negative index materials, *Progress in Electromagnetic Research – PIER* Vol. 51, 139–152, ISSN: 1559-8985

Jakšić, O.; Jakšić, Z. & Matović, J. (2009). Adsorption-desorption noise in plasmonic chemi-cal/biological sensors in multiple analyte environment", *Proceedings of SPIE* vol. 7362 *Microtechnologies for the New Millennium*, pp. 73621F-1-12, ISBN: 9780819476364.

Jakšić, Z. & Djurić, Z. (2004). Cavity Enhancement of Auger-Suppressed Detectors: A Way to Background-Limited Room-Temperature Operation in 3-14 Micrometer Range, *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 10, No. 4, 771- 776, ISSN: 1077-260X.

Jakšić, Z., Maksimović, M. & Sarajlić, M. (2005). Silver-silica transparent metal structures as bandpass filters for the ultraviolet range, *Journal of Optics A: Pure and Applied Optics*, Vol. 7, No. 1, 51-55, ISSN: 1464-4258.

Jakšić, Z.; Jakšić, O.; Djurić, Z. & Kment, C. (2007). A Consideration of the Use of Metama-terials for Sensing Applications: Field Fluctuations and Ultimate Performance", *Journal of Optics A: Pure and Applied Optics* Vol. 9, S377–S384, ISSN: 1464-4258.

Jakšić, Z.; Jakšić, O. & Matović, J (2009a) "Performance limits to the operation of nanopla-smonic chemical sensors – noise equivalent refractive index and detectivity", *Journal of Nanophotonics*, Vol. 3, pp. 031770-1-13, 6 April, ISSN: 1934-2608

Jakšić, Z. & Matović, J. (2009b). Coupling between propagating and evanescent modes in freestanding nanomembrane-based plasmon sensors using surface sculpting", *3rd Vienna International Conference Nano-Technology – VIENNANO´09*, March 18-20, Vienna, Austria, 187-194, ISBN 978-9-901657-33-7

Jiang, C.; Markutsya, S.; Pikus, Y. & Tsukruk, V. V. (2004b). "Freely suspended nanocomposite membranes as highly sensitive sensors", *Nature Materials* 3, pp. 721-728, ISSN: 1476-1122.

Jiang, C.; Markutsya, S. & Tsukruk,V. V. (2004b). Compliant, robust, and truly nanoscale free-standing multilayer films fabricated using spin-assisted layer-by-layer assembly. *Advanced Materials* 16, 157–161., ISSN: 0935-9648.

Jiang, C.; McConney, M. E.; Singamaneni, S.; Merrick, E.; Chen, Y.; Zhao, J.; Zhang, L. & Tsukruk, V. V. (2006). Thermo-Optical Arrays of Flexible Nanoscale Nanomembra-nes Freely Suspended over Microfabricated Cavities as IR Microimagers, *Chemistry of Materials*, Vol. 18, 2632-2634, ISSN: 0897-4756.

Jovanovic-Talisman, T.; Tetenbaum-Novatt, J.; McKenney, A. S.; Zilman, A.; Peters, R.; Rout, M. P.; Chait, B. T. (2009). Artificial nanopores that mimic the transport selectivity of the nuclear pore complex , *Nature* Vol. 457, No. 7232, 1023-1027. ISSN 0028-0836.

Jung, L. S.; Campbell, C. T.; Chinowsky, T. M.; Mar, M. N. & Yee, S. S. (1998). Quantitative Interpretation of the Response of Surface Plasmon Resonance Sensors to Adsorbed Films, *Langmuir*, 14 (19), pp. 5636 -5648, ISSN: 0743-7463.

Kash, A. (1991). Acoustic imager, *Journal of the Acoustical Society of America*, Vol. 89, No. 6, 3034-3034, ISSN: 0001-4966

Kashyap, R. (1999). *Fiber Bragg Gratings*, Academic Press, New York, ISBN-10: 0124005608.

Kasianowicz, J. J.; Brandin, E.; Branton, D. & Deamer, D. W. (1996). Characterization of individual polynucleotide molecules using a membrane channel, *Proceedings of the National Academy of Sciences USA*, Vol. 93, 13770–13773, ISSN: 0027-8424.

Kretschmann, E. & Raether, H. (1968). Radiative decay of nonradiative surface plasmons excited by light, *Zeitschrift für Naturforschung A* Vol. 23, 2135–2136, ISSN: 0932-0784.

Kruse, P. W. & Skatrud, D. D. eds (1997), *Uncooled infrared imaging arrays and systems*, Academic Press, San Diego Tokyo, ISBN-10: 0127521550

Li, Y.; Kunitake, T.; Onoue, S.; Muto, E. & Watanabe, H. (2007). Fabrication of Large, Free-standing Nanofilms of Platinum and Platinum–Palladium Alloy, *Chemistry Letters* Vol.36, No.2, 288-289, ISSN: 0366-7022.

Maier, S. A.; Barclay, P. E.; Johnson, T. J.; Friedman, M. D. & Painter, O. (2004) Low-loss fiber accessible plasmon waveguides for planar energy guiding and sensing. *Applied Physics Letters*, Vol. 84 (20), pp. 3990–3992, ISSN: 0003-6951.

Maier, S. A. (2007). *Plasmonics: Fundamentals and Applications*, Springer, Berlin, ISBN: 978-0-387-33150-8.

Maksimović, M. & Jakšić, Z. (2005). Modification of thermal radiation by periodical structures containing negative refractive index metamaterials, *Physics Letters A*, Vol. 342, No. 5-6, 497-503, ISSN: 0375-9601.

Maksimović, M. & Jakšić, Z. (2006). Emittance and absorptance tailoring by negative refractive index metamaterial-based Cantor multilayers, *Journal of Optics A: Pure and Applied Optics*, Vol. 8, 355-362, , ISSN: 1464-4258.

Maksimović, M.; Hammer, M. & Jakšić, Z. (2008). "Thermal radiation antennas made of multilayer structures containing negative index metamaterials", *Proceedings of SPIE 6896, Integrated Optics: Devices, Materials, and Technologies XII*, Christoph M. Greiner, Christoph A. Waechter, Editors, 689605, Feb. 12, 1-11, ISBN: 978081947071

Mamedov, A. A.; Kotov N. A.; Prato, M.; Guldi, D. M.; Wicksted, J. P. & Hirsch, A. (2002). Molecular design of strong single-wall carbon nanotube/polyelectrolyte multilayer composites. *Nature Materials* Vol. l, 190–194, ISSN: 1476-1122.

Martin, D. K., ed. (2007) *Nanobiotechnology of Biomimetic Membranes*, Springer Science, New York, ISBN-10: 0-387-37738-7.

Martinac, B. ed. (2008). *Sensing with Ion Channels*, Springer-Verlag Berlin Heidelberg, ISBN: 978-3-540-72683-8.

Matović, J. & Jakšić, Z. (2009). Simple and reliable technology for manufacturing metal-composite nanomembranes with giant aspect ratio", *Microelectronic Engineering*, Vol. 86, 906-909, ISSN: 0167-9317.

Merkoçi, A. ed. (2009). *Biosensing Using Nanomaterials*, John Wiley & Sons, Inc., Hoboken, New Jersey, ISBN 978-0-470-18309-0.

Ni, H.; Lee, H.-J. & Ramirez, A. G. (2005). A robust two-step etching process for large-scale microfabricated SiO_2 and Si_3N_4 MEMS membranes, *Sensors and Actuators A* Vol. 119, 553–558, ISSN: 0924-4247.

Otto, A. (1968). Exitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection", *Zeitschrift für Physik*. Vol. 216, 398.

Park, S. & Song, S. H. (2006). Polymeric variable optical attenuator based on long range surface plasmon polaritons, *Electronics Letters*, Vol. 42, No. 7, 402–404, ISSN: 0013-5194.

Pendry, J. B.; Holden, A. J.; Robbins, D. J. & Stewart, W. J. (1999). Magnetism from conductors and enhanced nonlinear phenomena. *IEEE Transactions on Microwave Theory and Technology* Vol. 47, 2075-2081, ISSN: 0018-9480

Raether, H. (1988). *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*. Springer-Verlag, Berlin Heidelberg, ISBN-13: 978-0387173634.

Ramakrishna, S. A. & Grzegorczyk, T. M. (2009). *Physics and Applications of Negative Refractive Index Materials*, SPIE Press Bellingham, Washington and CRC Press, Taylor & Francis Group, Boca Raton. ISBN: 9780819473998

Richards, P. L. (1994). Bolometers for infrared and millimeter waves, *Journal of Applied Physics* Vol. 76, No. 1, 1-24, ISSN: 0021-8979.

Ritchie, R. H.; Arakawa, E. T.; Cowan, J. J. & Hamm, R. N. (1968). Surface-plasmon resonance effect in grating diffraction, *Physical Review Letters*. Vol. 21, 1530–1533, ISSN: 0031-9007.

Rogalski, A. (2003). Infrared detectors: status and trends. *Progress in Quantum Electronics* Vol. 27, 59–210, ISSN: 0079-6727.

Rouquerol, J.; Avnir, D.; Fairbridge, C. W.; Everett, D. H.; Haynes, J. H.; Pernicone, N.; Ramsay, J. D. F.; Sing, K. S. W. & Unger K. K. (1994). Recommendations for the characterization of porous solids, *Pure and Applied Chemistry*, Vol. 66, No. 8, 1739-1758, ISSN: 0033-4545.

Sarid, D. (1981). Long-range surface-plasma waves on very thin metal films, *Physical Review Letters*, Vol. 47, No. 26, 1927–1930, ISSN: 0031-9007.

Sata, T. (2004). *Ion Exchange Membranes: Preparation, Characterization, Modification, Application*, The Royal Society of Chemistry, Cambridge, ISBN 0-85404-590-2.

Smolyaninov, I. I. (2008). Transformational optics of plasmonic metamaterials, *New Journal of Physics*, Vol. 10 115033-1-8, ISSN: 1367-2630

Stegeman, G. I.; Wallis, R. F. & Maradudin, A. A. (1983). Excitation of surface polaritons by end-fire coupling, *Optics Letters*, Vol. 8, No. 7, 386–388, ISSN: 0146-9592.

Toko, K. (2005). *Biomimetic Sensor Technology*, Cambridge University Press, Cambridge, ISBN 0-521-59342-5.

Ünlü, M. S. & Strite, S. (1995). Resonant cavity enhanced photonic devices, *Journal of Applied Physics*, Vol. 78, 607-639, ISSN: 0021-8979.

Vendamme, R.; Onoue, S.-Y.; Nakao, A. & Kunitake, T. (2006). Robust free-standing nanomembranes of organic/inorganic interpenetrating networks, *Nature Materials*, Vol. 5, 494-501, ISSN: 1476-1122.

Vercoutere, W. A. (2003). Discrimination among individual Watson-Crick base pairs at the termini of single DNA hairpin molecules, *Nucleic Acids Research*, Vol. 31, 1311–1318, ISSN: 0305-1048.

Veselago, V. (1968). The electrodynamics of substances with simultaneously negative values of ε and μ," *Soviet Physics Uspekhi*, Vol. 10, 509–514, ISSN 0038-5670

Watanabe, H. & Kunitake, T (2007). A Large, Freestanding, 20 nm Thick Nanomembrane Based on an Epoxy Resin, *Advanced Materials*, Vol. 19, 909–912, ISSN: 0935-9648.

Watanabe, H.; Muto, E.; Ohzono, T.; Nakao, A.; & Kunitake, T. (2009). Giant nanomembrane of covalently-hybridized epoxy resin and silica, *Journal of Material Chemistry*, Vol. 19, 2425–2431, ISSN: 0959-9428.

Worthing, P. T. & Barnes, W. L. (2001). Efficient coupling of surface plasmon polaritons to radiation using a bi-grating, *Applied Physics Letters* Vol. 79, 3035–3037, ISSN: 0003-6951.

Yampolskii, Y.; Pinnau, I. & Freeman, B. D. eds (2006) *Materials Science of Membranes for Gas and Vapor Separation*, John Wiley & Sons, Ltd. ISBN: 0-470-85345-X

Zayats, A. V. & Smolyaninov, I. I. (2003). Near-field photonics: surface plasmon polaritons and localized surface plasmons, *Journal of Optics A: Pure and Applied Optics* Vol. 5, S16–S50, ISSN: 1464-4258.

Zhang, S.; Fan, W.; Panoiu, N.; Malloy, K.; Osgood, R. M. & Brueck, S. R. J. (2005). Experimental demonstration of near-infrared negative-index metamaterials. *Physical Review Letters,* Vol. 95, 137404-1-4, ISSN: 0031-9007.

Zheng, H.; Lee, I.; Rubner, M. & Hammond, P. (2002). Controlled cluster size in patterned particle arrays via directed adsorption on confined surfaces. *Advanced Materials,* Vol. 14, 573–577, ISSN: 0935-9648.

Specific Serum-free Conditions can Differentiate Mouse Embryonic Stem Cells into Osteochondrogenic and Myogenic Progenitors.

Hidetoshi Sakurai, Yuta Inami, Naomi Nishio, Sachiko Ito,
Toru Yosikai, Haruhiko Suzuki and Ken-Ichi Isobe
*Department of Immunology, Nagoya University Graduate School of Medicine, 65
Tsurumai-cho, Showa-ku, Nagoya, 466-8560,
Japan*

1. Introduction

Embryonic stem (ES) cells and induced pluripotent stem (iPS) cells have great potentials for cell-based therapies based on their abundant potentials of self renewal and differentiation into all cell lineages (1,2). A serum-free ES cell differentiation system has an advantage for clinical applications because it can efficiently induce a specific cell lineage, and can avoid the risk of viral or prion infection by biomaterials. This study was initiated to examine how to efficiently induce paraxial mesodermal progenitor cells from ES cells in serum-free cultures. BMP4 acted as a key factor to promote the primitive streak-type mesoderm in both mouse development and ES cell differentiation in culture. Many lateral mesodermal derivatives such as hematopoietic cells, endothelial cells and cardiomyocytes, and intermediate mesoderm derivatives such as renal progenitors have been induced by BMP4 stimulation. However, differentiation of paraxial mesodermal cells from ES cells in serum-free culture has remained elusive. In this study, we developed a simple culture system with BMP4 and lithium chloride (LiCl) in serum-free conditions to promote two types of paraxial mesodermal progenitors, osteochondrogenic progenitors, and myogenic progenitors, which were identified using the paraxial mesodermal marker PDGFR-α.

2. Materials and methods

2.1 Cell culture and in vitro ES cell differentiation

CCE ES cells and ES cells expressing the *LacZ* gene (CCE/nLacZ) were kindly gifted by Dr. Nishikawa.

For serum-free culture of ES differentiation, type IV collagen (Nitta Gelatin), serum-free culture medium, SF-O3 (Sanko Junyaku), 0.2% Bovine serum albumin (BSA), 2-mercaptoethanol, and recombinant human BMP4 (R&D systems) were used. For myogenic mesodermal progenitor cell differentiation, initial induction by BMP4 was the same as osteochondrogenic progenitors. Three days after BMP4 treatment, the medium was changed entirely to SF-O3 (Sanko Junyaku) supplemented with 2.5mM LiCl, and cells were cultured in the medium for four days. For further myocyte induction, the cells which were sorted as

described below were re-cultured on a collagen type I-coated 24-well dish (Iwaki) in SF-O3 with 2ng/ml IGF-1 (R&D systems), 10ng/ml HGF (R&D systems) and 2ng/ml bFGF (R&D systems). Three days after re-culture, the medium was changed entirely to SF-O3 with 2ng/ml IGF-1. Four days after the medium was changed, the medium was changed again to SF-O3 with 2ng/ml IGF-1 and 10ng/ml HGF, and the cells were cultured for seven days.

2.2 Antibodies, cell staining, FACS analyses and cell sorting

Rat monoclonal antibodies (MoAbs), APA5 (anti-PDGFR-α), AVAS12(anti-VEGFR2) and ECCD2 (anti-E-cadherin) were kindly gifted by Dr. Nishikawa. Phycoerythrin-conjugated streptavidin (BD Pharmingen) was used to detect the biotinylated-APA5 antibody. ECCD2 and AVAS12 were directly conjugated by a standard method using allophycocyanin (APC). Cultured cells were harvested and collected in 0.05% trypsin-EDTA (GIBCO). Single-cell suspensions were stained as previously described (3) and analyzed or sorted by FACSCalibur or FACSVantage-HG (Becton Dickinson).

2.3 Transplantation of ESC-derived mesodermal progenitors into immunodeficient mice

We carried out mouse experiments according to protocols approved by the Animal Care and Use Committee of Nagoya University Graduate School of Medicine. The PDGFR-α^+ and PDGFR-α^-ECD$^+$ populations were purified and collected by FACS (> 5x10^5 cells). Cells were resuspended at a density of 2.5x10^4 cells/µl in αMEM. For intra-muscular transplantation, a quadriceps femoris muscle of a KSN nude mouse was injured by direct cramping with a diethyl ether anesthesia. Twenty micro liters of collected cell suspension were directly injected into the injured quadriceps of each mouse. For intra-bone marrow transplantation, a hole in the tibial bone was bored through to the bone marrow at the knee joint using a 21G needle with the anesthesia diethyl ether, and twenty micro liters of collected cell suspension were directly injected into bone marrow.

3. Results

3.1 In vivo muscle regeneration by paraxial mesodermal progenitors derived from murine ES cells.

First, we simply seeded ES cells onto a 10-cm dish coated with type IV collagen in αMEM supplemented with 10% fetal calf serum. Four days later, the cells were harvested and stained with anti-PDGFR-α and anti-VEGFR-2 antibodies. PDGFR-α + cells were sorted by FACS. Cell suspensions were directly injected into the injured quadriceps femoris muscle of a KSN nude mouse. The majority of PDGFR-α mesodermal progenitors were located in the interstitial zone of muscles, especially in the area adjacent to the myofibers (Fig. 1).

Since the ES cells have a *LacZ* marker, we stained with Pax7 and CD34 antibodies, and found LacZ positive cells have these satellite cell markers. Satellite cells were isolated from the KSN nude mouse that was injected with the PDGFR-α^+ cells in the quadriceps femoris muscle. Many LacZ$^+$ cells were observed in the culture, and some of them exhibited fiber formation like other host satellite cell-derived myofibers (Fig. 1) (3).

3.2 Serum-free culture to induce mesodermal progenitors

We attempted to differentiate ES cells into mesodermal lineage cells in serum-free medium. For paraxial mesoderm differentiation, 2x10^5 ES cells were plated on a 10-cm dish coated

with type IV collagen, and were differentiated in a serum-free culture medium, which was comprised of SF-O3 supplemented with BSA, 0.1mM 2-mercaptoethanol, and 1 ng/ml BMP4. PDGFR-α^+ cells emerged after four days of differentiation, and reached a peak on day five, with almost half of the cells becoming PDGFR-α^+. The morphology of the cell aggregates changed from ES cell-like round colonies to cobblestone monolayers. The expression of T, Msgn, Tbx6 and Pax3, which play an important role in mesodermal development, was detected. We conclude that the addition of BMP4 to SF-O3 medium permits efficient induction of paraxial mesodermal progenitor cells from mouse ES cells.

3.3 BMP4-induced paraxial mesodermal progenitor cells have osteogenic and chondrogenic potentials *in vivo.*

We investigated the *in vivo* tissue differentiation potentials of paraxial mesodermal progenitor cells. At 4 to 6 days of LacZ-positive ES cell culture in serum-free medium with BMP, cells were stained with anti-PDGFR-α and anti-ECD antibodies. The PDGFR-α^+ and PDGFR-α^-ECD$^+$ populations were sorted by FACS and were directly injected into bone marrow of the tibia of a KSN nude mouse. Twenty-eight days later, β-galactosidase staining was performed to detect engrafted cells. LacZ positive cells, which have light blue nuclear staining, were observed in the trabecular bone. ES cell-derived PDGFR-α^+ cells could differentiate to osteocytes *in vivo*. These PDGFR-α^+ and PDGFR-α^-ECD$^+$ populations were also directly injected into injured quadriceps femoris muscle. The engrafted tissues were then analyzed four weeks after transplantation. Although we could not detect muscle cells derived from Lac Z ES cells, we were surprised to detect ectopic cartilage in engrafted skeletal muscle. The ectopic cartilage was derived from engrafted cells as confirmed by the expression of LacZ by fluorescent immunohistochemistry (Fig. 2).

3.4 Serum-free induction of myogenic progenitor cells by ES

In order to differentiate ES cells to myogenic progenitor cells, we exposed cells to BMP for the first three days and then replaced it with LiCl for four days. The PDGFR-α^+ population strongly expressed the dermomyotome markers Pax3 and Pax7, and the myogenic regulatory genes Myf-5 and Myo-D. Next, we asked whether this procedure could induce PDGFR-α^+ cells to form mature myofibers in vitro. In mouse skeletal muscle regeneration, many growth factors such as insulin-like growth factor-1 (IGF-1), hepatocyte growth factor (HGF), or basic fibroblast growth factor (bFGF) activate proliferation of Myf5$^+$/MyoD$^+$ myoblasts. Although IGF-1 could promote myogenin expression in this culture independently, adding both HGF and bFGF for three days enhanced myogenin expression and stimulated MRF4 expression. Myogenin expression was confirmed by immunohistochemistry. Some spindle-shape cells with mono- or multi-nuclear myogenin staining were observed. Further treatment with IGF-1 and HGF promoted mature skeletal muscle cells which expressed skeletal muscle actin (Fig.3).

4. Discussion

4.1 In vivo regeneration of muscle, cartilage and bone by ES-derived paraxial mesoderm cells

Here we described ES cell differentiation into paraxial mesoderm and paraxial mesoderm derived- tissues in vivo. First, we succeeded in regenerating muscle by direct injection of differentiated ES cells into injured muscle. In this case, we injected PDGFRα + cells, which

were differentiated from ES cells by two-dimensional culture on Type IV collagen with FCS (Fig. 1). These simple methods produced progenitors of muscle satellite cells. Because muscle satellite cells are tissue stem cells, the transplanted cells may continuously replicate in transplanted tissues. Next, we tried to differentiate ES cells to muscle cells using serum-free medium. Because BMP has been shown to be essential to mesoderm formation (4), we cultured ES cells in serum-free medium with BMP4, and transplanted PDGFRα + cells into damaged muscle. We expected to get myogenic progenitor cells by this method. However, we detected ectopic cartilage tissues in transplanted, injured muscle. These results indicate that BMP4 induces progenitors of cartilage cells. However, further experimentation revealed that BMP-4-induced progenitors also differentiated into osteogenic cells when sorted PDGFRα cells were injected into the bone marrow cavity (Fig. 2). These results indicate that, *in vivo,* BMP4-induced mesodermal progenitor cells differentiate into different cell types, depending on the surrounding tissue.
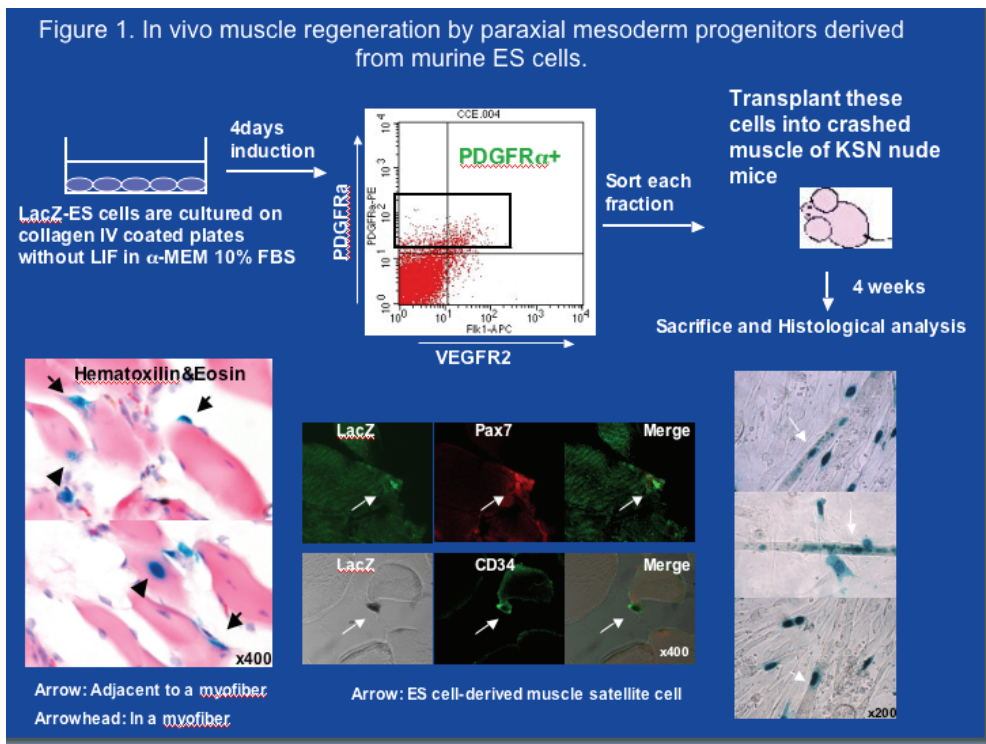
Fig. 1. ES cells were cultured on type IV collagen-coated dishes with 10%FCS. PDGFR-α + cells were sorted by FACS. Cell suspensions were directly injected into the injured quadriceps femoris muscle of a KSN nude mouse.

4.2 Serum-free induction of myogenic progenitor cells from ES cells
Early removal of BMP-4 followed by LiCl treatment promotes the differentiation of PMPs to myogenic progenitors. In mouse embryogenesis, the establishment of the myotome from the
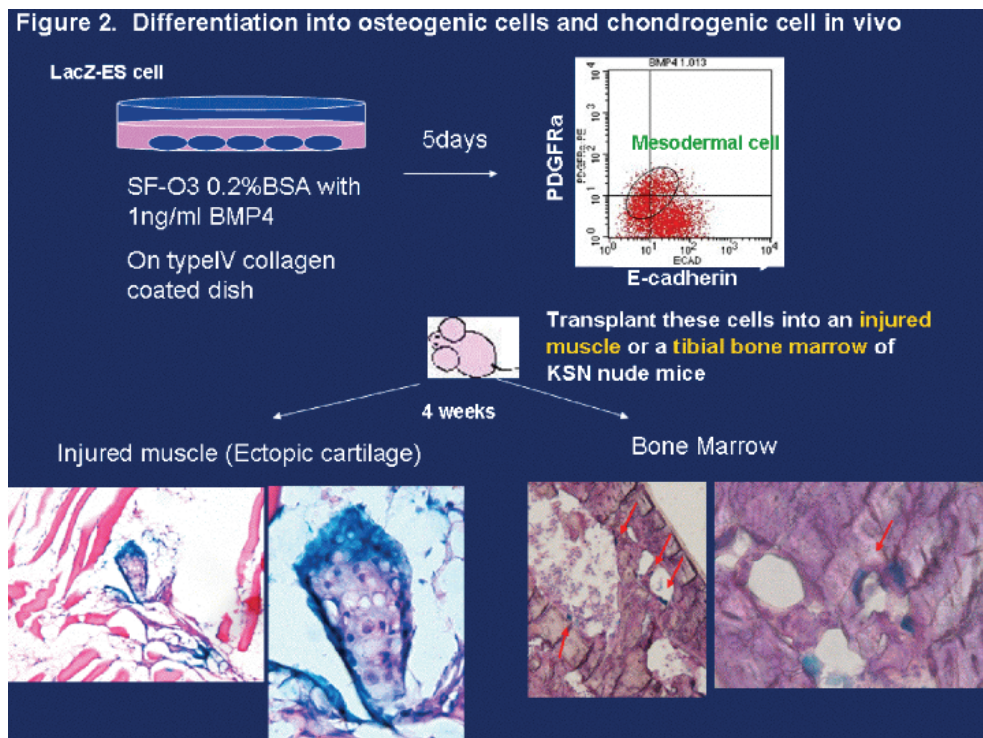
Fig. 2. BMP4-induced paraxial mesodermal progenitor cells can differentiate into chondrogenic and osteogenic cells in vivo. ES cells were plated on a 10 cm dish coated with type IV collagen and differentiated in a serum-free culture medium supplemented with 1ng/mlBMP4. At day 5 of culture, PDGFR-$\alpha^{+\sim low}$ and ECD^{low} populations were sorted (Fraction 1) and injected into bone marrow of tibia of KSN nude mouse.

dermomyotome is stimulated by Wnt signaling and by the expression of Noggin, which inhibits the BMP signaling pathway (5). Our culture system mimics these developmental events by removing BMP-4 after the induction of paraxial mesodermal progenitors and adding LiCl which inhibits GSK-3 and causes translocation of cytoplasmic β-catenin to the nucleus (6.).

In this study, we successfully induced the efficient differentiation of mouse ES cells towards myogenic cell types using chemically defined conditions. Also, simple differentiation of ES cells by BMP4 induced bone and cartilage tissues *in vivo*.

5. References

[1] Nishikawa S, Jakt LM, Era T. Embryonic stem-cell culture as a tool for developmental cell biology. Nature reviews. 2007;8:502-507.
[2] Yamanaka S. A Fresh Look at iPS Cells 2009; 137:13-17.
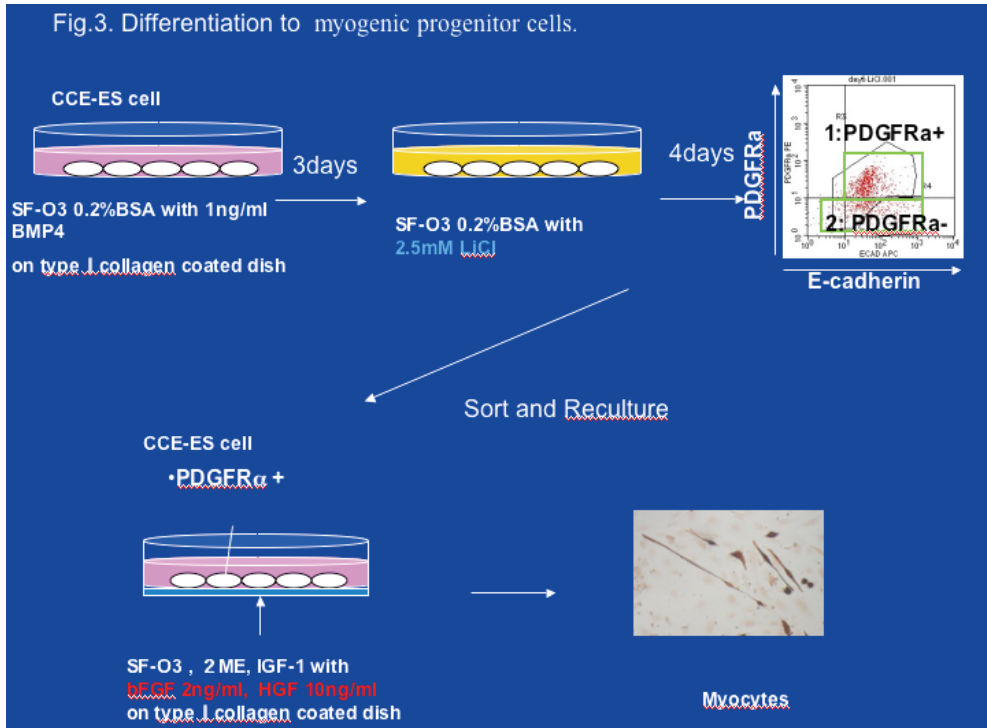
Fig. 3. Myogenic progenitor cells are induced from ES cells by transient exposure to BMP4 and subsequent LiCl treatment in chemically defined media. After sorting, fraction 1 was recultured as described.

[3] Sakurai H, Okawa Y, Inami Y, Nishio N, Isobe K. Paraxial mesodermal progenitors derived from mouse embryonic stem cells contribute to muscle regeneration via differentiation into muscle satellite cells. Stem Cells. 2008 Jul;26(7):1865-73.

[4] Winnier G, Blessing M, Labosky PA and Hogan BL. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. Genes Dev. 1995 9: 2105-2116.

[5] Parker, M.H., Seale, P. and Rudnicki, M.A. Looking back to the embryo: defining transcriptional networks in adult myogenesis. Nat. Rev. Genet. 4 (2003) 497-507.

[6] Yamamoto, H., Kishida, S., Kishida, M., Ikeda, S., Takada, S. and Kikuchi, A. Phosphorylation of axin, a Wnt signal negative regulator, by glycogen synthase kinase-3beta regulates its stability. J Biol Chem 274 (1999) 10681-4.

Micromanipulation with Haptic Interface

Shahzad Khan[1], Hans H. Langen[1] and Asif Sabanovic[2]
[1]Delft University of Technology,
[2]Sabanci University,
[1]The Netherlands
[2]Turkey

1. Introduction

As the nature has provided us with things in dimensions ranging down till micro/nanometers likewise humans also were able to fabricate components in the same scales, but the prominent challenge lies in the fact to assemble components into a single and functionalized product. Use of monolithic ways to produce complex micro/nano systems is desirable, but is not always feasible. The current state of art is to incorporate components into a single functional product and to assemble micro parts one by one (Dechev et al., 2003; Popa & Stephanou, 2004). The only solution to this problem is to develop machines capable to assist humans to assembly micro-parts and major effort is being undertaken in this direction by several research organizations (MicroNED, 2005). The first and foremost requirement for the assembly process is to "precisely manipulate" objects. Manipulation includes cutting, pushing, pulling, indenting, or any type of interaction which changes the relative position and relation of entities. This paper concentrates on manipulation by pushing as it is a useful technique for manipulating delicate, small, or slippery parts, parts with uncertain location, or parts that are otherwise difficult to grasp and carry (Lynch & Mason, 1996; Lynch, 1999; Sitti, 2003). The process of manipulation by pushing of micro-objects possesses many challenges due to the requirements of:

- Actuators with high resolution (in nanometer range), high bandwidth (up to several kilohertz), large force output (up to few newtons) and relatively large travel range (up to a few millimeters). (Khan et al., 2006)
- Robust and transparent bilateral controllers for human intervention so that high fidelity position/force interaction between the operator and the remote micro environment can be achieved. (Sitti & Hashimoto, 2003)
- Vision based algorithms to estimate the location of objects being manipulated and visual servoing procedure to position manipulators so that these objects can be manipulated (pushed) along a desired trajectory. (Torsten & Fatikow, 2006)
- Controlled pushing force to compensate the surface forces arising between the object and the environment.

Manipulating objects requires not only precise position control of actuators but also delicate control of forces involved in the manipulation process. Visual information is required for path planning whereas use of force feedback is indispensable to ensure controlled physical interactions. Thus, pushing using only visual feedback is not sufficient. It is also

indispensable to sense and control the interaction forces involved in the manipulation process with nano-newton resolution. Moreover, it is a well established fact that human operators are much more adaptable to force changes and can react much effectively under unexpected situations as compared with robotic manipulators. In other words, human operator can perform force control and motion operation much more skilfully, thus human intervention can be employed in pushing of the micro-object using bilateral control. Bilateral control is defined as the control of two systems working together on an actual or virtual task. Typically, it is used for teleoperation, in which one system is called the "master" side and the other is called the "slave" side of bilateral action. Slave subsystem tracks the positions of the master subsystem and master side provides the forces encountered by the slave side to the operator and hence, teleoperation is achieved. Nowadays, many researchers have come up with the notion of "multilateral control" consisting of more than two systems working with proper coordination to achieve a desired task. In order to perform telemicromanipulation it is indispensable to achieve robust and transparent bilateral controllers for human intervention so that high fidelity position/force interaction between the operator and the remote micro/nano environment can be achieved. As bilateral control enables skilled teleoperation on several tasks, it offers better safety, low cost and high accuracy, on the other hand it also suffers from time delay problem which affects the transparency of the systems.

The whole process of pushing of micro-object can be divided into two concurrent processes: in one pushing is performed by human operator which acts as an impedance controller and alters the velocity of the pusher while in contact with the micro-object. In another, the desired line of pushing for the micro-object is achieved through visual feedback procedures in one dimension to attain translational motion of the micro-object. In this work, vision/force hybrid feedback procedure for force controlled pushing of micro objects with human assistance is presented. The chapter is organized as follows. Section II provides the problem definition and approach and Section III explains the custom built telemicromanipulation setup. In Section IV, scaled bilateral teleoperation is demonstrated with experimental details concerning force/position tacking between the master and the slave. Finally, Section V provides the procedure for pushing micro-objects along with the experimental results and Section VI concludes the chapter and discusses future directions.

2. Problem definition and approach

The problem dealt within this work concerns utilizing semi-autonomous manipulation scheme for pushing of polygonal micro-object, by point contact to achieve pure translational motion with the aid of a human operator by employing scaled bilateral teleoperation with force-feedback. In order to achieve pure translation motion, the proper line of action of the pushing force needs to always pass through the varying center of friction of the polygonal micro-objects. Thus, while the pushing operation is in progress, it is inevitable to online estimate the center of friction and align the probe such that line of action passes through the center of friction of the micro-object.

The above mentioned problem is coped by utilizing a proposed method for pushing polygonal micro objects using semi-autonomous scheme with human assistance as shown in Figure 1. The whole process of pushing a micro-object is divided into two concurrent processes: in one pushing is performed by the human operator which can react to sudden change of force by switching between force/position controls and alters the velocity of the

pusher while in contact with the micro object. In the second part, the desired line of pushing for the micro-object is determined continuously by vision based algorithm so that it always passes through the varying center of friction. The necessary subtasks utilized to perform the whole process are as follows:

- Piezoresistive AFM microcantilever has been utilized to measure the interaction forces with the environment with nano-newton resolution.
- Human operator interacts with the micro environment using scaled bilateral teleoperation. The desired position is commanded by the human operator and transferred to the micro environment after scaling and the resultant interaction forces are felt by the human operator after performing the force scaling.
- Visual processing algorithms are employed to detect position and orientation of the micro-object for the estimation of the desired line of pushing.

Human operator utilizes the scaled bilateral control structure as in Section IV to generate the desired position which is scaled with a factor before feeding to the position controller. The position controller uses the feedback from PZT actuator to compensate the position error to achieve the desired position of the piezoresistive microcantilever. As the micro-cantilever comes in contact with the micro-object the interaction resultant forces are felt by the human operator through the force fed from the piezoresistive micro-cantilever after scaling by a factor of β. Depending upon the situation human operator which acts an impedance controller can adjust the impedance (effective muscle stiffness) to change from position control to force control to push that micro-object along X-axes with the commanded position/force. Moreover, the operator has the access to the visual information for monitoring the pushing process. Visual feedback procedures are performed automatically to estimate the correct line of pushing to achieve pure translational motion. Visual processing algorithms are employed to detect position and orientation of the micro-object for the estimation of the location of the center of friction. Finally, the velocity of the piezoresistive cantilever is varied at the contact point using visual feedback process to ensure that resultant line of pushing passes through the center of friction to achieve pure translation motion along the X-axes.
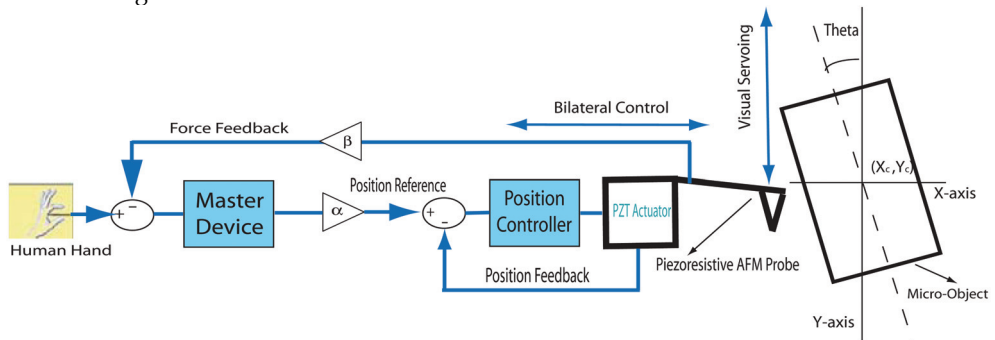
Fig. 1. Hybrid control structure for semi-automated pushing

3. Tele-micromanipulation setup

The system is composed of three parts, namely a master mechanism operated by the human operator, a slave mechanism interacting with the micro environment and human-computer

interface for visual display. For the master mechanism a DC motor is utilized, while a piezoresistive microprobe attached on PZT stacks is used for the slave. A bilateral man-machine interface is implemented for control as shown in schematics Figure 2.
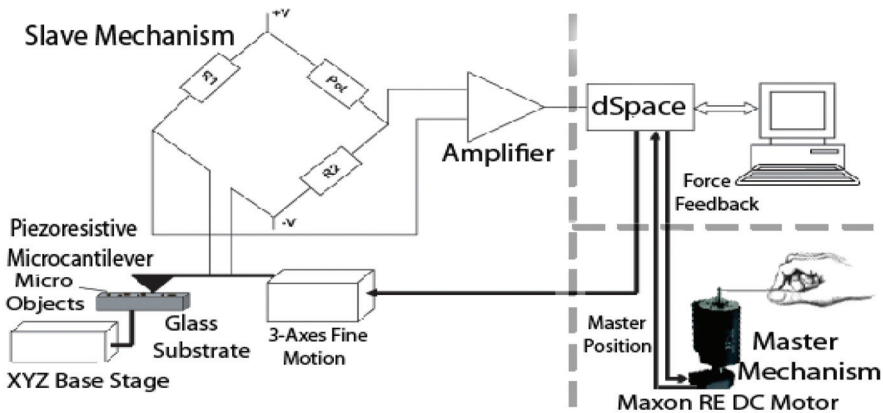
Fig. 2. Schematic of tele-micromanipulation system

The position data from the master side is scaled and transferred to slave side, while simultaneously, the force measured at the slave side is scaled and transferred back to master. XYZ base stages are used for proper alignment of micro object or in other words to bring the micro objects under the workspace. A graphical display is available to the operator through the signal processing card where the bilateral control algorithms are implemented. Figure 3 shows the experimental setup. The one degree of freedom master mechanism consists of a brushed DC servo (Maxon motors RE40) and is manually excited with the help of a light rod that is connected to the shaft. Capability to control positions with nanometer accuracy and to estimate the forces in nano-Newton scales is required. High magnification microscopy is also essential for visual feedback with acceptable resolution.

An open architecture micromanipulation system that satisfies the requirements has been developed and used as a slave mechanism. Nano scale positioning of the micro cantilever has been provided using three axes piezo stages (P-611 by Physik Instrumente) which are driven by a power amplifier (E-664) in closed loop external control mode. Potentiometers (strain gauge sensors) integrated in the amplifier, are utilized for position measurement of the closed loop stages which possess a travel range of 100 µm per axis with one nanometer theoretical resolution. Stictionless and frictionless compliant guiding systems exist in the stages. An open loop piezoelectric micrometer drive (PiezoMike PI-854 from Physik Instrumente) has been utilized as the base stage, which is equipped with integrated high resolution piezo linear drives. Manually operable linear drives are capable of 1 µm resolution and the automatic movement range of the micrometer tip with respect to the position can be set 50 µm (25 µm in/out). Nanometer range resolution is achieved for this movement by controlling the piezo voltage. As for the force feedback, a piezoresistive AFM cantilever (from AppNano) has been utilized along with an inbuilt Wheatstone bridge. A real time capable control card (dSPACE DS1103) is used as control platform and an optical microscope (Nikon MM-40) is used for visual feedback.
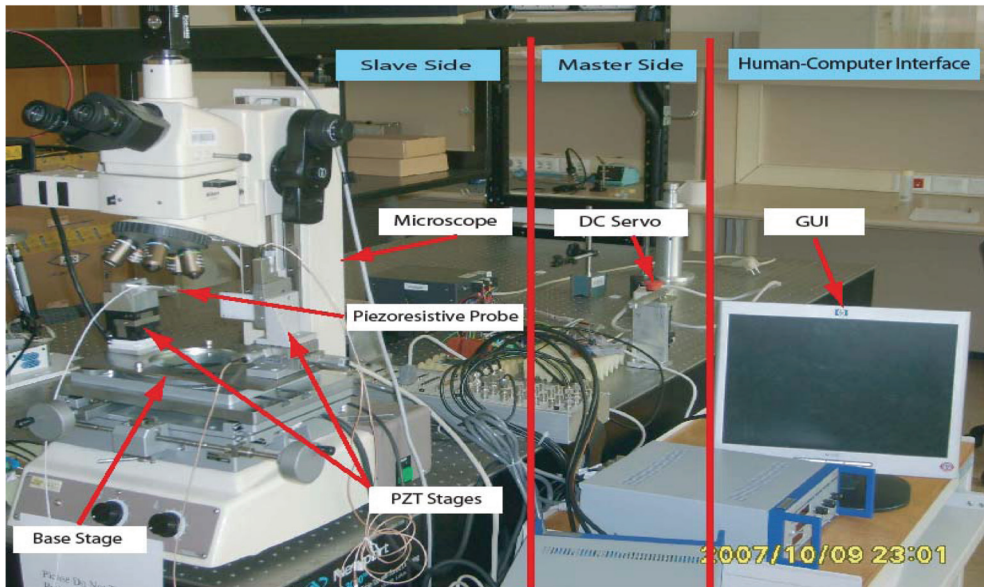
Fig. 3. Experimental setup for tele-micromanipulation

4. Scaled bilateral teleoperation

In this section implementation of scaled bilateral control in a custom built telemicromanipulation setup is presented. Force sensing with nN resolution using piezoresistive AFM (Atomic Force Microscope) micro-cantilever is demonstrated. Force/position tracking and transparency between the master and the slave is presented with varying references after necessary scaling.

4.1 Force sensing using AFM microcantilever

In order to achieve force transparency between the master and the slave, it is necessary to sense the force in nano-newton range with high accuracy. Piezoresistive AFM cantilever with inbuilt Wheatstone bridge from AppliedNanostructures is utilized as a force sensor as well as probe for pushing operation as shown in Figure 4. Piezoresistive sensors have been used for many other MEMS applications, including accelerometers, gyroscopes and AFM cantilevers. The primary advantage of this approach is that the sensor impedance is relatively low (a few $K\Omega$), and it is possible to extract small signals without interference from noise with off-chip integrated circuits.

The working principle is based on the fact that as the force is applied at the free end of the cantilever using the PZT actuator with the glass slide, the change of resistance takes place depending on deflection of the cantilever. The amount of deflection is measured by the inbuilt Wheatstone bridge providing a voltage output, which is amplified by the custom built amplifier. To compensate the offset in the resistance value, one of the active resistors in the full bridge is replaced by the potentiometer. The amplified voltage is send to the data acquisition dSpace1103 card for further processing. Figure 5 represents the attractive forces for pulling in phase between the tip and glass slide (Khan et al., 2007; Khan et al., 2009). The

decreasing distance between the tip and glass slides is represented by the increase in the position of PZT axis. As the distance between the tip and glass slide decreases the attractive forces increases. The result clearly indicates that force sensing with the resolution of nN range is achieved.
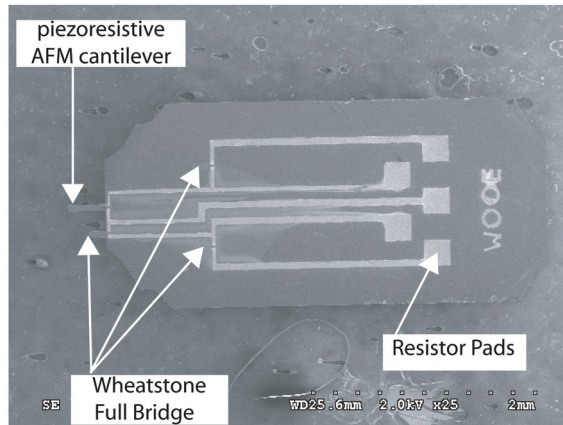
Fig. 4. Piezoresistive AFM Cantilever with inbuilt Wheatstone bridge
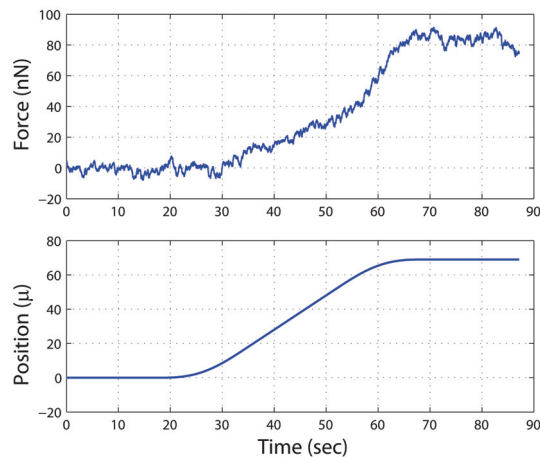
Fig. 5. Force for smooth step position reference.

4.2 Scaled bilateral control structure

Since the master and slave are working on macro and micro scales respectively, thus it's indispensable to use general bilateral controller to scale the position and forces between two sides for extensive capability. In other words, position information from the master is scaled down to slave and force information from the slave side is sent up to master as shown in Figure 6 comprising of the master and the slave side. Piezo-stage on the slave side is required to track master's position as dictated by position controller. The 1D force of interaction with environment, generated by piezoresistive cantilever, on the slave side is

transferred to the master as a force opposing its motion, therefore causing a "feeling" of the environment by the operator. The conformity of this feeling with the real forces is called the "transparency". Transparency is crucial for micro/nanomanipulation application for stability of the overall system. Furthermore, for micro system applications, position and forces should be scaled in order to adjust to operator requirements. Position of the master manipulator, scaled by a factor α, is used as a position reference for the slave manipulator, while the calculated force due to contact with environment, scaled by a factor β, is fed-back to the operator through the master manipulator. In order to eliminate oscillations both on master side because of oscillatory human hand and on the slave side due to piezoresistive cantilever dynamics, position of master manipulator and force of slave manipulator are filtered by low pass filters before scaling.
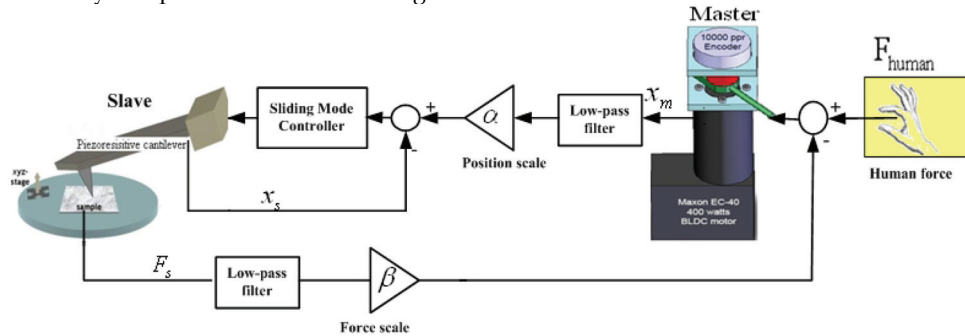
Fig. 6. Scaled bilateral teleoperation control structure

Since the master and slave side resides on macro and micro scales respectively, thus its very vital to appropriately choose the scaling factor in order to attain the optimum performance. In the ideal condition, the steady state condition of the bilateral controller should be Eqn. (1).

$$x_s = \alpha x_m$$
$$F_m = \beta F_s$$

(1)

Where α and β represents the position and force scaling respectively. x_m, x_s denotes the master and slave position respectively and F_m, F_s denotes the master and slave force respectively. To be able to meaningfully interact with the micro environment, positions and forces are scaled to match the operator requirements.

In the first and second experiments, scaling factors of α = 0.027 µm/deg and β = 0.00366 N/nN are used, that is an angular displacement of 1deg on the master side corresponds to a linear displacement of 0.027 µm on the slave side and a force of 0.00366 nN on the slave side corresponds to a force of 1N on the master side. The objective of these experiments is to provide very fine motion on the slave side for a relatively larger displacement on the master side, hence α is selected according to this objective. Then the corresponding forces/torques for each amount of displacement were compared for the selection of β , keeping in mind that the DC motor on the master side has low torques.

In order to validate the position tracking between the master and the slave, the commanded position from the master is transferred after necessary scaling to be tracked by the slave side. Figure 7 illustrates the experimental results for position tracking along with the tracking error of the bilateral controller. It can be clearly seen that the slave tracks the master

position with high accuracy. This position tracking performance is acceptable for precisely positioning the micro cantilever. In order to validate the force tracking, the slave forces encountered from the environment is being transferred to the master side after necessary scaling. Figure 8 demonstrates the force tracking between the master and slave along with the tracking error. It can be clearly observed that the master tracks the slave force precisely. The position and force result shows that it is possible to perform the pushing operation using this bilateral control structure.
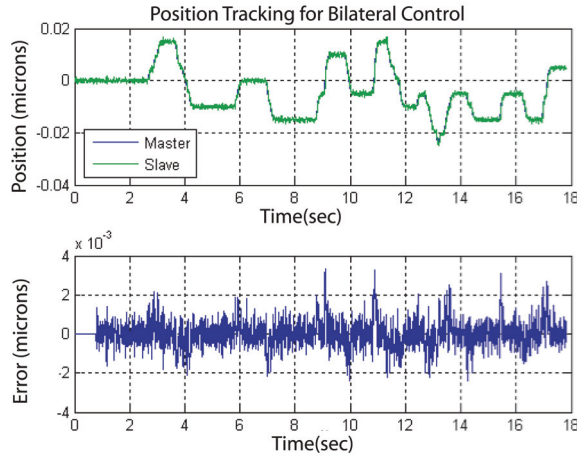
Fig. 7. Position Tracking between the master and the slave
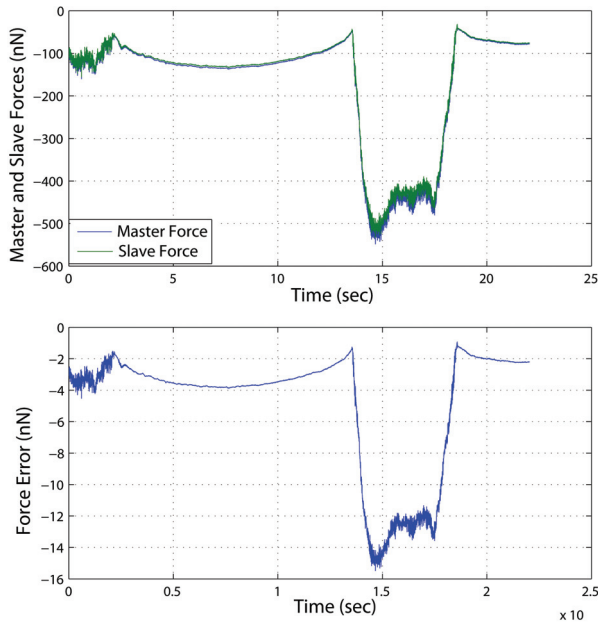
Fig. 8. Force Tracking between the master and the slave

5. Semi-autonomous pushing scheme

5.1 Point contact pushing for translational motion

Precise positioning of micro-objects lying on a substrate using a point contact pushing to track a desired trajectory poses lot of challenges. The pusher or probe needs to be controlled in such a way to reorient and transport the micro object to its final location using a stable pushing operation (The probe or pusher is always in contact with the micro object during pushing operation). Using only a point contact with a limited number of freedom the task of pushing on a horizontal plane can be realized.

In this work, pure translation of a regular object from one location to another by orientating the line of action of the pushing force to the desired direction is presented. The desired translational motion of the object cannot be achieved if the line of pushing at the contact point passes through the center of the mass of the micro-object. Due to the dominance of the frictional forces existing in between the micro-object and supporting surface, the inertial effect will be neglected and the motion will be dependent upon the motion and direction of the frictional forces. Thus, the resultant line of pushing needs to be directed through the centroid of the frictional distribution, center of friction (Center of friction is defined as single point where the frictional distribution between the interface of object/substrate can be lumped) where all the distribution of friction can be lumped into a single point to achieve pure translational motion. The pushing mechanism realized to achieve pure translational motion is explained in detail in the following subsection. Moreover, due to uncertain topography of the surfaces the frictional distribution changes with respect to time. Thus, the most important task lies in the online estimation of the center of friction using visual and force information.

5.2 Trajectory control for known center of friction

Figure 9 represents the scenario of pushing rectangular object using a point contact pushing to achieve pure translation motion. The rectangular micro-object has two points, namely COM (center of mass) and COF (center of friction). The contact point of the pusher is taken as the origin of the reference frame. The x-axis and y-axis of the frame is chosen to be parallel and perpendicular connecting to the edge of polygon. The velocity of the probe along x-axis $\vec{V}_x$ and y-axis $\vec{V}_y$ are controlled by visual feedback and human operator, respectively. The desired velocity vector $\vec{V}_{des}$ resultant of $\vec{V}_x$ and $\vec{V}_y$ needs to pass through COF, hence having an orientation angle θ_d to achieve a pure translation motion. The value of $\vec{V}_y$ cannot be controlled to achieve the desired velocity vector as it is administered by the human operator, rather it is only a measurable quantity. The variable $\vec{V}_x$ can be calculated by taking into consideration the value of $\vec{V}_y$ to achieve the desired velocity vector $\vec{V}_{des}$ making an angle θ_d as in the following equations.

The relationship between the $\vec{V}_x$ and $\vec{V}_y$ can be written as Eqn. (2) by analyzing Figure 9 and solving for $\vec{V}_{des}$ yields Eqn. (3).

$$\vec{V}_{des} \cos \theta_d = \vec{V}_x \tag{2}$$

$$\vec{V}_{des} = \frac{\vec{V}_x}{\cos \theta_d} \tag{3}$$

Similarly, the relationship between the $\vec{V}_y$ and $\vec{V}_{des}$ can be written as Eqn. (4) and inserting the Eqn. (3) into Eqn. (4) will yield Eqn. (5)

$$\vec{V}_{des} \sin \theta_d = \vec{V}_y \tag{4}$$

$$\vec{V}_y = \vec{V}_x \tan \theta_d \tag{5}$$

The Eqn. (5) indicates that it's possible to only control $\vec{V}_y$ to achieve the resultant velocity vector $\vec{V}_{des}$ to pass through COF. As it's already discussed, the location of the COF is not constant with respect to time, thus it is necessary to calculate the varying location of COF and to allow the desired line of pushing to the pass through it (Yoshikawa and Kurisu, 1991). Detail proof of the derivation for online estimation of center of friction can be found in (Khan et al., 2009).
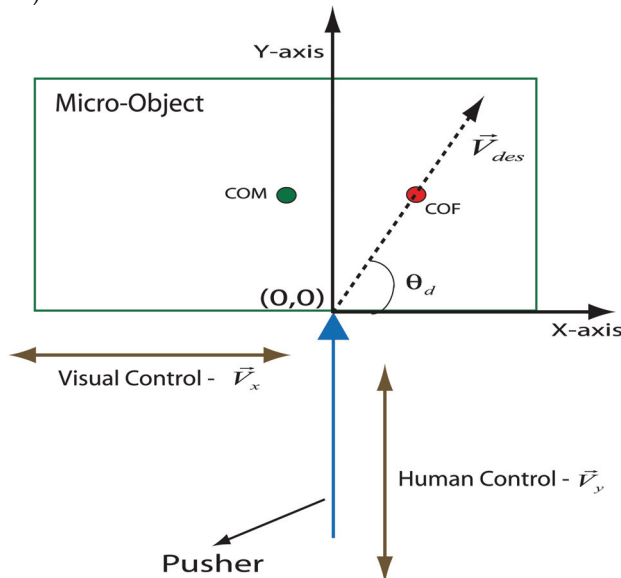
Fig. 9. Calculation of velocity vector for known center of friction

5.3 Experimental validation of pushing operation
In order validate the above mentioned pushing algorithm, several experiments were conducted by pushing a rectangular micro-object of size 200μm at the mid-point of the length of rectangle and the line of action passes through the center of mass (Khan et al., 2009). Figure 10 demonstrates the snapshot of the pushing operation and it can be clearly observed that after several steps the micro-object starts to rotate. Thus, it is unmanageable to translate a micro-object by pushing through the center of mass. The above results provide necessary arguments to conclude that to achieve pure translation motion it is necessary that the line of action passes through the center of friction to compensate the orientation angle. Figure 11 demonstrates the snapshot of pushing rectangular micro-object such that the line of action passes through the center of friction. It can be clearly seen that the proposed procedures was able to compensate the orientation effect to attain pure translational motion.
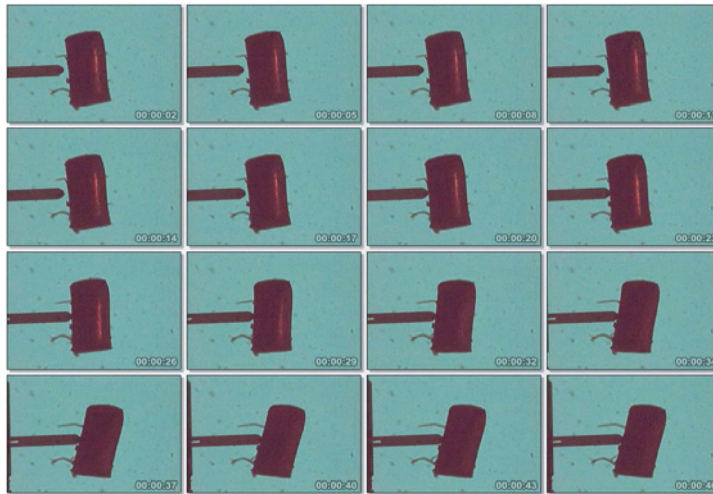
Fig. 10. Snapshot of pushing rectangular object at the mid-point of the rectangle and line of action passes through center of mass of the object.
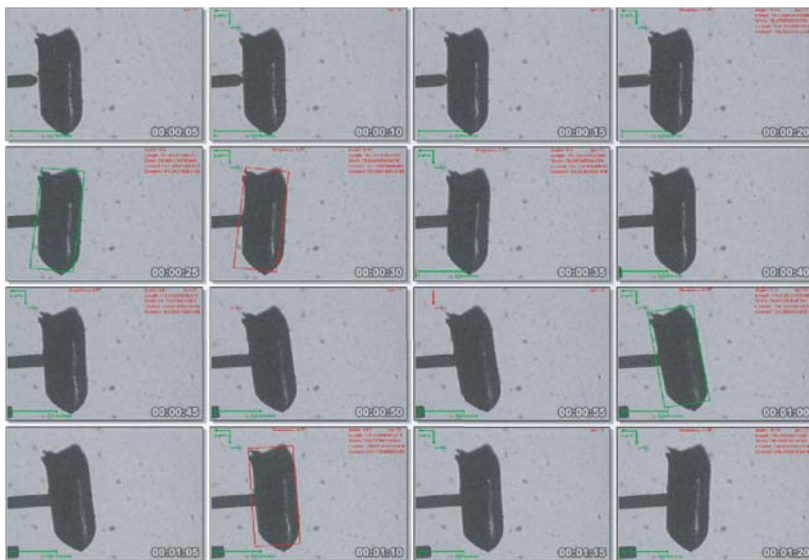
Fig. 11. Snapshot of pushing rectangular object such that the line of action passes through the center of friction

6. Conclusion

In this chapter, a semi-autonomous scheme based on hybrid vision/force feedback using a custom built tele-micromanipulation is proposed. The goal is to push the micro-object to achieve pure translation motion using semi-autonomous mechanism with the aid of human

operator. The pushing operation is undertaken by the human operator using visual display which acts an impedance controller and can switch between velocity control to force control by adjusting the stiffness (muscle stiffness) depending upon the behaviour of the motion of the micro-object. Visual module provides the information about the position and orientation of the micro-object to calculate the time-varying COF (center of friction) in recursive manner for each captured frame. The velocity at the contact point is altered using visual feedback procedures such that the resultant direction of velocity passes through the COF to achieve pure translational motion. Experimental results concerning nano-newton resolution force sensing, force/position tracking between the master and the slave is presented which is a requirement to fulfil the pushing operation.

7. References

Dechev, N.; Cleghorn W.L. & Mills J.K. (2003). "Construction of 3d mems microcoil using sequential robotic microassembly operations", *Proceedings of ASME International Mechanical Engineering Congress.*

Popa, D.O & Stephanou, H.E. (2004). "Micro and meso scale robotic assembly", *Proceedings of WTEC Workshop: Review of U.S. Research in Robotics.* MicroNED (2005), Microsystems Netherlands www.microned.nl

Lynch, K.M & Mason, M.T (1996). "Stable pushing: Mechanics, controllability, and planning", *The International Journal of Robotics Research*, Vol.15, No.6, P.P (533- 556).

Lynch, K.M. (1999). "Locally controllable manipulation by stable pushing", *IEEE Transactions on Robotics and Automation*, Vol.15, No.2, P.P (318 - 327).

Sitti, M. (2003). "Atomic force microscope probe based controlled pushing for nanotribological characterization" *IEEE/ASME Transactions on Mechatronics*, Vol.8, No.3.

Khan, S.; Elitas, M; Kunt, E.D. & Sabanovic.A (2006). "Discrete sliding mode control of piezo actuator in nano-scale range", *Proceedings of IIEE International Conference on Industrial Technology.*

Sitti, M & Hashimoto, H. (2003). "Teleoperated touch feedback from the surfaces at the nanoscale: mdeling and experiments", *IEEE/ASMETransactions on Mechatronics*, vol. 8. No. 1, pp. 287–298.

Torsten, S & Fatikow, S (2006), Real-time object tracking for the robot based nanohandling in a scanning electron microscope, *Journal of Micromechatronics*, vol. 3 of 3-4, pp. 267–284.

Khan, S; Nergiz, A.O; Sabanovic, A & Patoglu, V. (2007) Development of a micromanipulation system with force sensing, *Proceedings in IEEE/IROS International Conference on Intelligent Robots and Systems, 2007.*

Khan, S; Sabanovic, A & Nergiz, A.O. (2009). Scaled Bilateral Teleoperation Using Discrete Sliding Mode Controller, *IEEE Transactions on Industrial Electronics*, Vol.56, No.9, September 2009.

Yoshikawa, T. & Kurisu, M (1991) Identification of the center of friction from pushing an object by a mobile robot, *Proceedings in IEEE/RSJ International Workshop on Intelligent Robots and Systems – IROS*

Fabrication of High Aspect Ratio Microcoils for Electromagnetic Actuators

Daiji Noda, Masaru Setomoto and Tadashi Hattori
Laboratory of Advanced Science and Technology for Industry, University of Hyogo
Japan

1. Introduction

Actuators are finding increasing use in the various fields and many applications. Therefore, it is one of the most important components in various machines because its performance determines to operate a machine. Among the actuators, we are focusing on electromagnetic actuators that could be driven at a low voltage, high power, high efficiency, and low cost. Then, the majority of actuators used in macroscopic applications are in this type. Recently, the size reduction and sophistication are required for parts and devices. Therefore, actuators, which hold big volume and weight with a part of a product, have been required to reduce their size. Nonetheless, electromagnetic actuators are known to be unsuitable for miniaturization because the allowable current carrying capacity is very small when current paths of coil are microscopic, making it difficult to obtain high power. In addition, it is very difficult to process microscopic current paths by mean of conventional machining technology. Therefore, the key technology to realizing practical electromagnetic microactuators is micro-fabrication process.

On the other hand, LIGA (German acronym for Lithographite, Galvanoformung, and Abformung) process (Becker et al. 1986) could be used to fabricate nano- and micro-parts for devices. The LIGA is a total process and have three major steps in the process including X-ray lithography, electroforming to fabricate metallic molds, and the use of these molds to form required parts of plastic micro structure. The X-rays that are generated by synchrotron radiation have high directivity and transmission characteristics, and could be used to expose photoresist to deep depths of 1 mm or more. In this X-ray lithography, the NewSUBARU synchrotron radiation facility at our university (Ando et al. 1998) was used. This was operated at energy of 1.0 or 1.5 GeV modes. The X-ray exposure at beamline 11 (BL11) of NewSUBARU was carried out with the workpiece held in a specially manufactured nine parts operation exposure stage (Mekaru et al. 2001), in which two axes are moved by piezoelectric elements, while five of other axes are moved using stepping motor; the remaining two axes are rotated by stepping motor. Thus, this X-ray exposure stage could make it feasible to create three-dimensional (3D) structures (Mekaru et al. 2002).

In this research, we have achieved development of 3D deep X-ray lithography technique that we used to produce a high aspect ratio spiral microcoil patterns (Mochizuki et al. 2007; Matsumoto et al. 2008; Setomoto et al. 2008; Noda et al. 2007, and Noda et al. 2008a). With this technique, it is very expected to fabricate high aspect ratio of coil line structures with

narrow line widths. This electromagnetic actuator, which features high aspect ratio coil, would also be possible to provide enhanced high suction force in spite of miniature size.

2. Design and simulation of electromagnetic actuator

An electromagnetic type actuator incorporating a magnetic circuit was designed. As the design of magnetic circuit, we have used the type called "open frame solenoid", which is open on the sides (Matsumoto et al. 2008) as shown in Fig. 1. The material of magnetic core (fixed core and plunger) and shield parts (yoke) were used a Permalloy 45, which is a nickel iron alloys, because it has the largest permeability among soft magnetic metals. When a voltage is applied to the coil lines, a magnetic flux forms in a gap, which is deforming a magnetic field and producing suction force on a plunger.
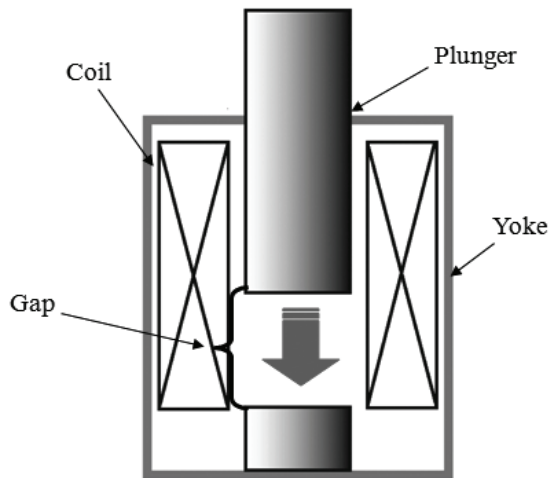
Fig. 1. Designed model of actuator operation with magnetic circuit.

An acrylic pipe with an outside diameter of 5 mm was used as a base material for coil. This material is PMMA (polymethlymethacrylate) which exhibits a positive type photoresist. Therefore, it could directly be exposed to X-ray lithography to form high aspect ratio coil line structures on pipe surface. Figure 2 shows the image diagrams of coil lines on pipe surface. Wire type coil is limited to copper wire size. However, high aspect ratio type in this research is fabricated using X-ray lithography technique. Therefore, coil line width was obtained narrow and deep depth, that is high aspect ratio lines were formed.

First, we have simulated suction force in high aspect ratio coil line structure. In this model, a magneto-motive force is proportional to squares of current path. If the aspect ratio of coil liens is increased, cross sectional area of coil lines is also increased allowing a greater current path. Figure 3 shows results of suction force and permit current path in different aspect ratio. Here, we used coil parameters as coil line width of 10 μm and coil turns of 675. And, the gap between plunger and fixed core was decided to 1 mm length. When an aspect ratio is reached 5, suction force may be about 25 times as large as aspect ratio of 1. These calculated results indicated that suction force could be increased using high aspect ratio coil line structure.
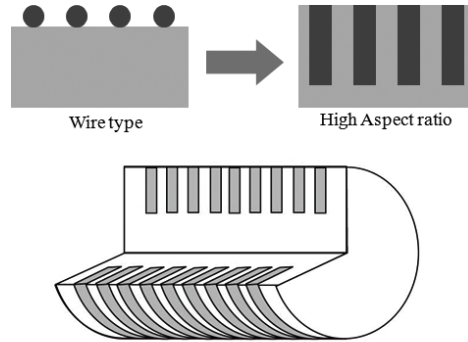
Fig. 2. Image diagram of high aspect ratio coil lines.
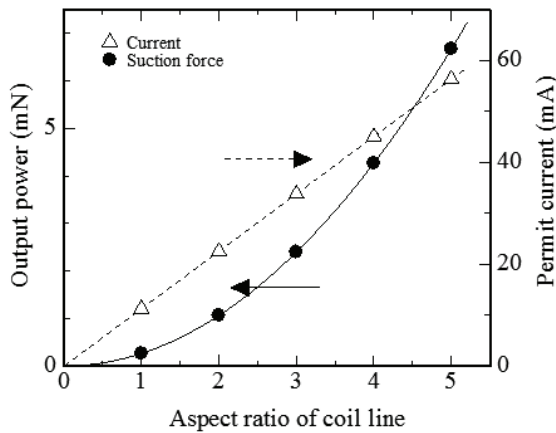
Fig. 3. Calculation of suction force and permit current in different aspect ratio.

In order to estimate of simulation results, we have also fabricated a measurement system in order to measure a suction force of designed electromagnetic actuator, as shown in Fig. 4. The measurement system was provided with a movable target, which moved with plunger. The target was suspended by means of coil spring. To determine suction force, we have measured the displacement of the target using laser displacement meter. Thus, suction force could be determined by the displacement data and spring constant of used coil spring. The gap adjustment was performed by combining a coarse-motion XY stage and fine-motion XY stage. We measured suction force generated by actuator using this measurement system. Coil is used acrylic pipe as base coil part and winding of polyurethane coated copper wire having 50 μm in wire diameter. Figure 5 shows that the theoretical values by calculation and actual measurement values of suction force were compared. The measurement results were relatively in good agreement with the theoretical values. One problem of assumed magnetic path was that magnetic flux could be extremely different from the reality in small gap region. The measurement results indicate the tendency in good agreement. Therefore, this measurement system was reliable.
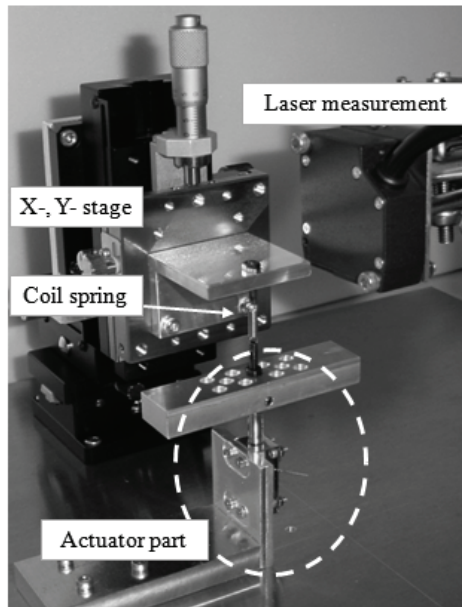
Fig. 4. Measurement system of suction force fro electromagnetic actuator.
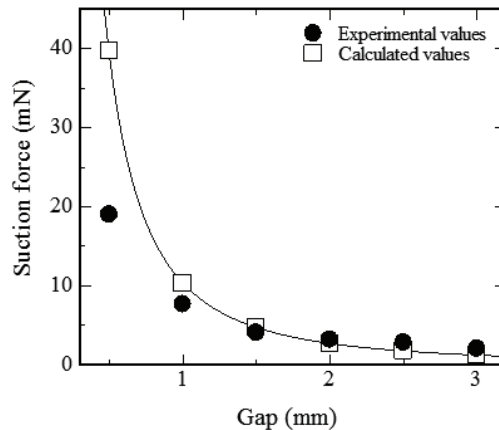
Fig. 5. Suction force comparison between measurement values and calculation values.

3. Fabrication process for coil lines

A spiral coil lines was formed on surface of acrylic pipe using X-ray lithography and metallization techniques. Fabrication processes for coil lines were shown in Fig. 6. First, a spiral coil structure of photoresist was formed on pipe surface using deep X-ray lithography. Next, copper was thinly deposited on coil line structures by sputtering to use as an electrode for electroforming. Then, the acrylic pipe was immersed into a copper plating bath for electroforming. Finally, the plated copper was dissolved on form coil lines

by leaving only deposit in the spiral structure using chemical copper etching. The following sections were discussed detailed descriptions of each process.
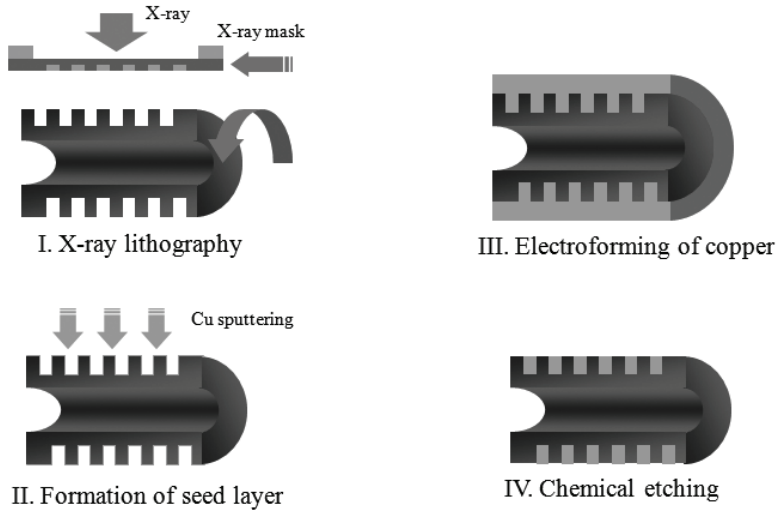
Fig. 6. Fabrication process of coil lines.

3.1 X-ray Lithography

We have mainly used X-ray mask having two line width patterns with 10 and 30 µm, respectively. Therefore, a spiral structure of coil lines was obtained 10 or 30 µm line width. To fabricate spiral micro coil lines, acrylic pipe was rotated using stepping motor and movement of X-ray mask was controlled by piezoelectric element. The X-ray exposure chamber was provided to perform 3D X-ray lithography, as shown in Fig. 7. In this study,
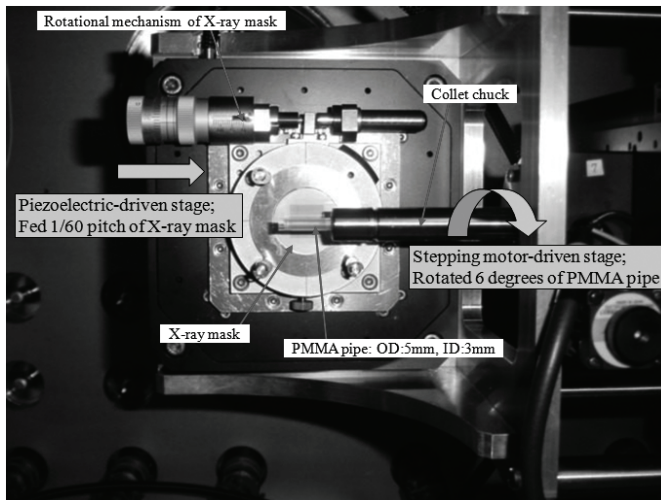
Fig. 7. Picture of multi operation X-ray exposure stage.

exposure process was divided into 60 steps in order to closely to continuity exposure. Thus, the pipe was rotated through an angle of 6 degrees while the X-ray mask was advanced by just 1/60 pitch of line pattern for each X-ray exposure cycles, as shown in Fig. 8. After X-ray lithography, spiral structure was developed on pipe surface, which was carried out using a GG developer (diethyleneglycolmonobutyether, 60 vol%; morpholine, 20 vol%; ethanoaine, 5 vol% & distilled water, 15 vol%) at room temperature. After this development, spiral structure was obtained.

The spiral structure of coil lines was observed using scanning electron microscope (SEM). The processing depth, that was determined aspect ratio of coil lines, could be controlled by the X-ray dose during X-ray exposure and the development time. Figure 9 shows relationship between processing depth and development time with three parameters of X-ray dose. Figure 10 shows the SEM image of spiral structure of coil lines with a pitch of 20 µm. In this case, an aspect ratio was achieved about 5 as coil line width of 10 µm. To use in 30 µm lines and space patterns, we obtained an aspect ratio of 2, as shown in Fig. 11. From these figures, we were able to confirm the joints between each section of spiral pattern were perfectly aligned.
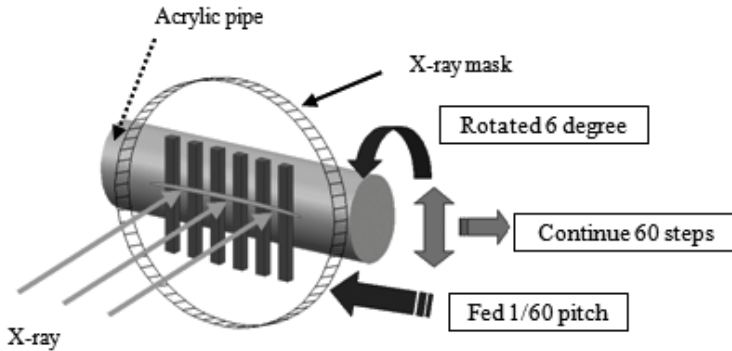
Fig. 8. Image diagram of X-ray lithography for spiral coil.
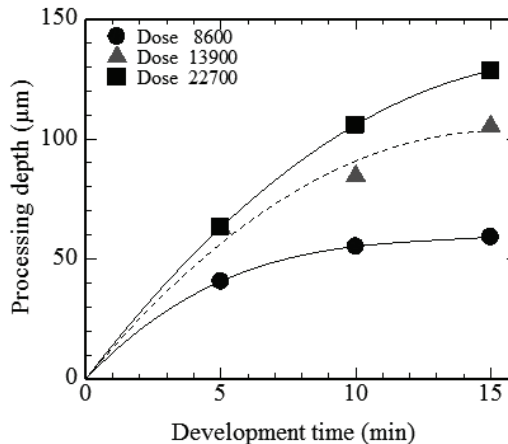
Fig. 9. Relationship between processing depth and development time.
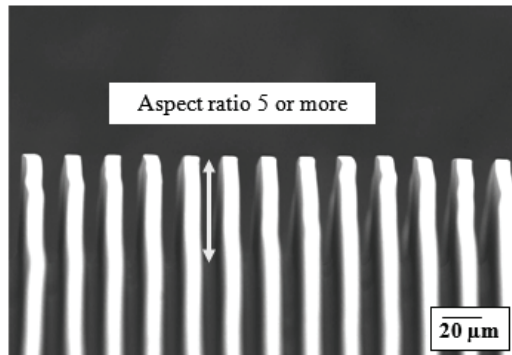
Fig. 10. SEM image of coil line width of 10 μm.
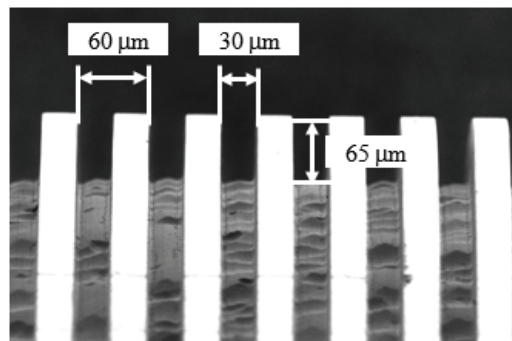
Fig. 11. SEM image of coil line width of 30 μm.

3.2 Formation of seed layer

Since an acrylic pipe is non-conductive material, a conductive seed layer is required for electroforming. We formed a seed layer on pipe surface using sputtering. In order to realize a circumferentially flat seed layer, we have performed sputtering rotating the pipe at 3 rpm and sputtering at three hours. First, it was found that a resistance of seed layer between both ends was very high. To solve this problem, we moved the position of pipe along its axis (Noda et al. 2007). As this result, we successfully produced an approximately 300 nm thick seed layer, which was level in the circumferential direction of the pip surface.

3.3 Copper electroforming

Following a seed layer deposition, we immersed an acrylic pipe in a copper sulphate solution bath including a levelling agent to uniform electroforming. To form plating film growth uniformity over the all coil line patterns, the pipe was connected to mixer and rotated in the plating solution. Figure 12 shows the plating bath.

In first plating, it is not good result to be due to voids that has generated inside the plating film. It was confirmed that these voids were produced inside coil line patterns had prevented the formation of satisfactory coil lines. Since high aspect ratio structure had been formed to be electroplated, it was estimated that air bubbles would remain inside high aspect ratio structure due to surface tension when the pipe was dipped in the plating
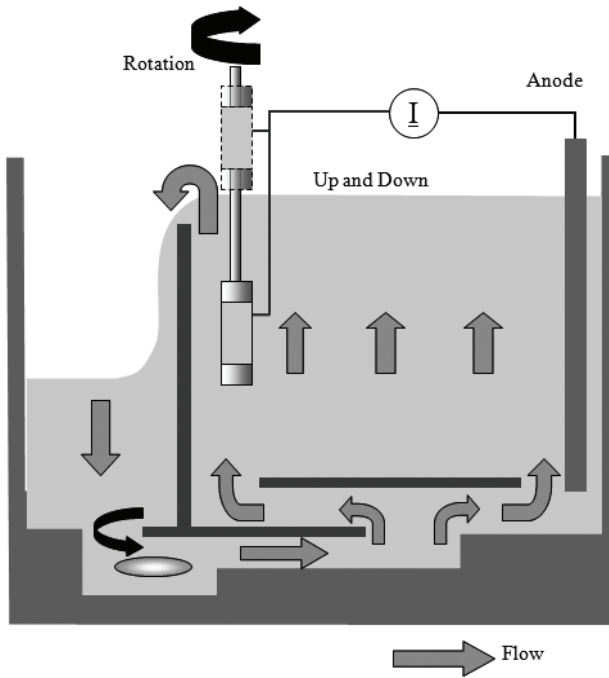
Fig. 12. Image diagram of copper electroforing bath.

solution. By cleaning air bubbles from spiral structures, a vacuum deforming was used before electroforming. In addition, the pipe was lifted out of the bath a few times, circulating the plating liquid in the patterned area during the plating. After these improvements, we confirmed to obtain a void-free copper plating film, as shown in Fig. 13.
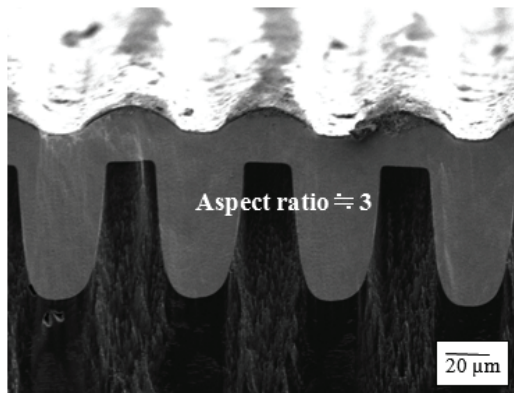
Fig. 13. SEM image of coil line width of 30 μm after copper electroforming.

However, for those whose aspect ratio was about 5, the electroforming did not even enter the grooves in the majority of area. When the aspect ratio of coil lines was higher, this tendency was more significant.

3.4 Formation of coil lines

We were investigating the formation of coil lines using etching and machining, respectively. Figure 14 shows the result by machining. It was very good result and copper coil lines were separated. However, copper and acrylic pipe of spiral structure were removed together by machining. Then, an aspect ratio of coil lines was changed and unknown in this result. On the other hand, isotropic copper etching was performed until the insulated portions of the wiring were exposed. Therefore, it was not changed an aspect ratio of coil lines. For copper etching, E-process-W etchant (made by Meltex co., ltd.) was used. In this etching process, the pipe rotation mechanism was also used to rotate the acrylic pipe in the etchant to ensure uniform pipe surface etching. As a result, we produced the coil lines by etching the copper until the protrusions of spiral structures were exposed, as shown in Fig. 15.
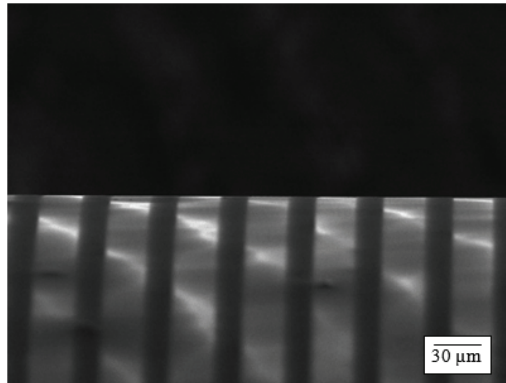
Fig. 14. SEM image of coil lines by machining.

Fig. 15. SEM image of coil lines by chemical etching.

4. Measurement of suction force

We have measured the suction force using fabricated spiral coil with narrow and high aspect ratio coil lines. Figure 16 shows the actuator part in measurement system. From this figure, this system is very easy to exchange a coil part. The gap between plunger and fixed

core was adjusted XY stages. Figure 17 shows that calculated values and actual measurement of suction force in this system were compared. In this case, we used a spiral coil having a width of 30 μm and aspect ratio of about 2. The measurement results were relatively in good agreement with calculation values. Here, result in small gap region was tended to include considerable errors. The reason is same in the result of Fig. 5.
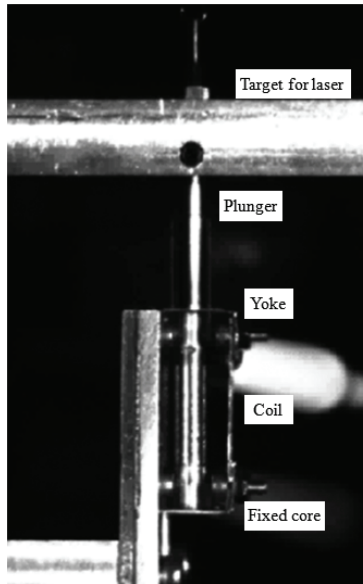
Fig. 16. Picture of actuator part in measurement system.
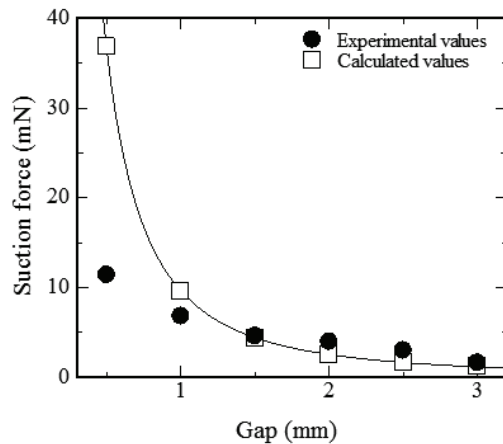
Fig. 17. Suction force comparison between calculattion and measurement values.

Figure 18 shows the suction force different aspect ratio coil lines using the gap of 1.5 mm in good agreement with calculation values. Increasing an aspect ratio of coil lines generates higher suction force. But, we have only measured under an aspect ratio of 2. Currently, we

have been proceedings to measure the suction force by providing the spiral coils with higher aspect ratio coil line structures produced by X-ray lithography and metallization techniques.
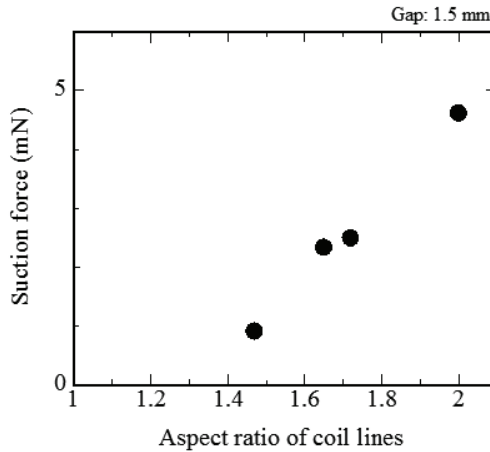
Fig. 18. Suction force different aspect ratio.

5. Development of microcoil

Acrylic pipe in outside diameter of 5 mm has been used in these experiments. Therefore, coil size is too large. Then, we were proposed and new fabrication process and tried to fabricate microcoils. First, we have calculated coil performance with high aspect ratio coil lines. Figure 19 shows calculation results of quality factor (Q factor) and resistance for different aspect ratio structures (Noda et al. 2008b). Here, we used coil parameters as coil turns of 20 and coil line width of 10 μm. And, coil diameter is 0.5 mm and coil length is 0.4 mm. From this figure, Q factor was very improved increasing the aspect ratio of coil lines. Therefore, it is also very expected that high performance microcoil with high aspect ratio could be obtained.
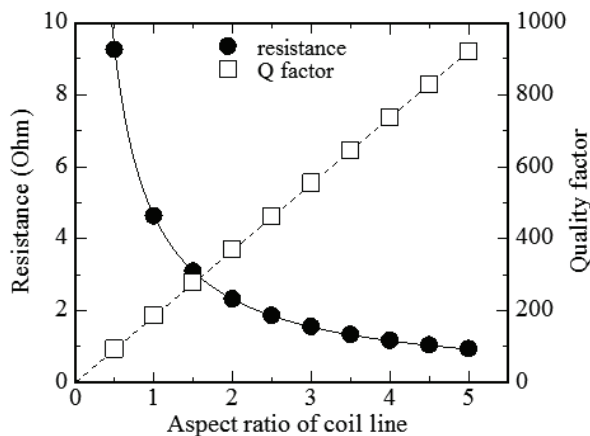
Fig. 19. Calculation of Q factor and resistance different aspect ratio.

To obtain microcoil with high aspect ratio coil lines, we used metal bar as coil base material with a diameter of 0.5 to 1 mm. However, metal bar was not directly exposed. Then, we applied a PMMA on a master metal bar using dipping method (Noda et al. 2008b; Yamashita et al. 2006 & Noda et al. 2008c). The thickness of PMMA was determined as coil line depth. Thus, it is very important factor in microcoil. The process was largely identical to that used for the acrylic pipe expect the final etching step, as shown in Fig. 20.
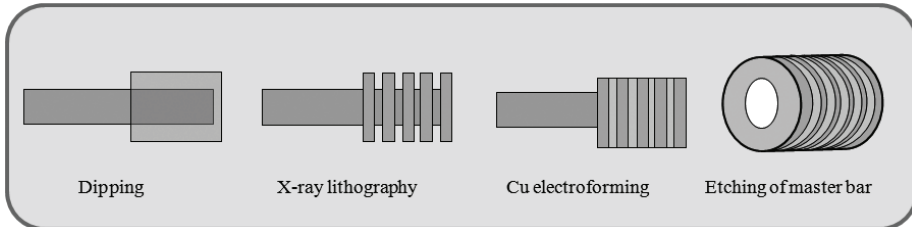
Fig. 20. Fabrication process using metal master bar.

5.1 Dipping method

We used a dipping method in order to obtain a thick layer of photoresist on metal bar. Figure 21 shows dipping method. This method comprises four steps: dipping, recovery, air drying, and baking. The master bar was dipped in a photoresist solution while rotating the cylinder shaft. Then, it was dried in the air while being rotated before the baking. In recoating, the master bar was pre-baked for short period of time before following the same steps to apply photoresist. Baking was performed in the final step. To prepare a PMMA solution, a PMMA sheet was crushed and dissolved in 2-ethoxyethyl acetate for more than 15 hours (Yamashita et al. 2006). A highly viscous photoresist solution and control over the centrifugal force are important factors to obtain a thick uniform coating, and thus enable the production of high aspect ratio coil line structures (Noda et al. 2008b).
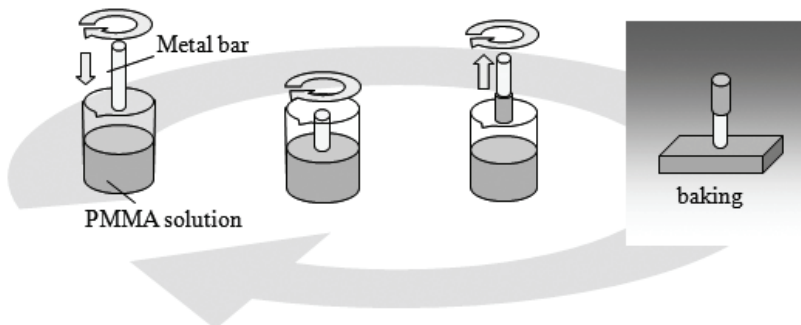
Fig. 21. Image diagram of dipping method.

5.2 Results and discussions

We were able to control the thickness of the PMMA on the metal bar by the speed of rotation and concentration of PMMA solution. PMMA thickness of more than 100 μm was obtained in a single coating. Thus, the aspect ratio achieved for 30 μm width spiral structure was greater than 3.

A spiral structure was formed in the PMMA on the metal bar using X-ray lithography technique. In the case, we used an X-ray mask with 30 μm width with a mask space ratio of 1:1. The diameter of the metal bar was used at 0.5 mm. Figure 22 shows a SEM image of a spiral structure with a pitch of 60 μm. This figure shows that the aspect ratio realized was about 6 because the spiral structure widths were narrower than the designed 30 μm width.

Next, we performed a metallization process, including electroforming and master bar etching. In this case of metal bar, the master bar acts as the seed layer for electroforming. Therefore, the plating metal grew from the surface of the master bar completely filling the
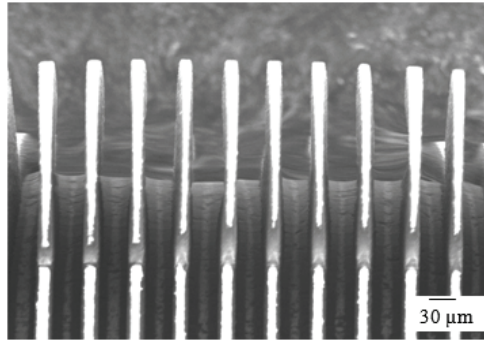
Fig. 22. SEM image of coil lines with high aspect ratio structures.
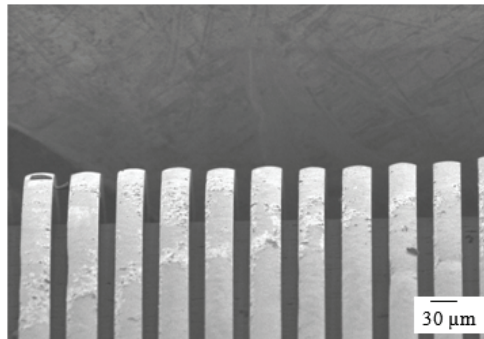
Fig. 23. SEM image of coil lines after photoresist etching.
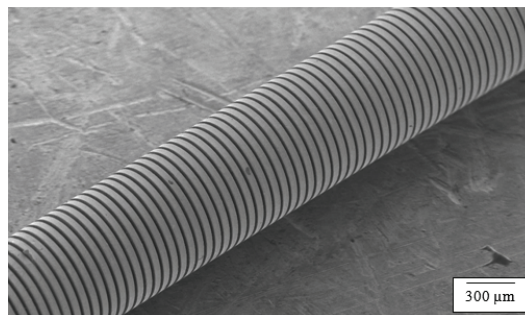
Fig. 24. SEM image of coil lines on the whole.

high aspect ratio spiral structures. Figure 23 shows a SEM image of coil lines with a pitch of 60 μm after removing the spiral photoresist patterns. This aspect ratio was obtained about 2. Then, we succeeded to fabricate a microcoil having a pitch of 60 μm on metal bar of 0.5 mm in diameter, as shown in Fig. 24.

Figure 25 shows a comparison of the size of the fabricated microcoils. On the right was a coil fabricated using acrylic pipe as the base material, and on the left was a microcoil using metal bar. We were able to fabricate a 0.5 mm diameter microcoil with high aspect ratio. It is expected to fabricate higher aspect ratio coil lines using metal bar as coil base material.
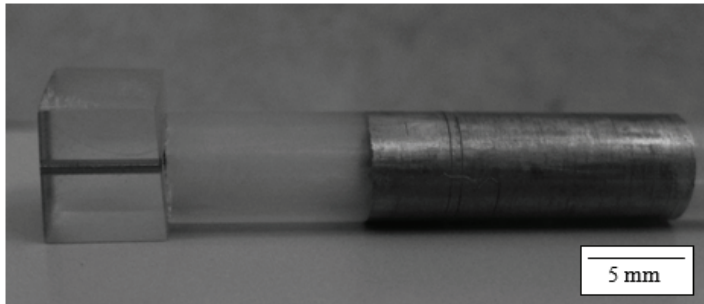
Fig. 25. Picture of fabricated microcoils with 1 mm diameter and acrylic pipe.

5.3 Fabrication of microcoil using LIGA process

In order to produce a spiral microcoil, the LIGA process was modified to the 3D processing method (Mekaru et al. 2004; Mekaru et al. 2005 & Mekaru et al. 2007). First, Au electroless plated to the surface of a brass rod with diameter of 0.46 mm, and PMMA of 10 μm thickness was applied by a dipping method. To fabricate a master of coil line strtuctures, 3D X-ray lithography was performed. In this case, 10 μm line width pattern was used for X-ray mask. Figure 26 shows the back view of X-ray exposure chamber at BL 11 used in this experiment. An X-ray exposure process was simplified only two steps (Mekaru et al. 2005), because diameter was very slim comparison with acrylic pipe diameter. Figure 27 shows the PMMA microstructure with a coil line width of 10 μm and thickness of 10 μm, which was spirally produced on the surface of brass rod with diameter of 0.46 mm.

Next, Ni electroforming was carried out to fabricate a metallic mold. Ne electroforming performed in conditions that nickel sulfamate used for electroforming solution. The brass rod and the PMMA microstructure in the inside of the Ni mold were removed by chemical etching. Then, the Ni mold wad completed having spiral structures, as shown in Fig. 28. In order to obtain this cross sectional photograph, one of the metallic mold was sectional. From this pattern, it was confirmed that the form of the master was transferred faithfully. The edge portion of the slot was also observed clearly.

For the replication process, an enscrewing injection machine (made by Juken Machine Works Co., Ltd.) was used to fabricate microcoil. In order to utilize the demolding process, it was necessary to synchronize the speed of rotation and mold opening according to the pitch of worm. Figure 29 shows the concept of worm demolding. Liquid crystal polymer (LCP) was the select as resin material, consideration for enhancing adhesion to form copper electroforming. As the result of injection molding, we succeeded in fabrication of resin core such small size worm, as shown in Fig. 30. Unlike the SEM image of the master of the
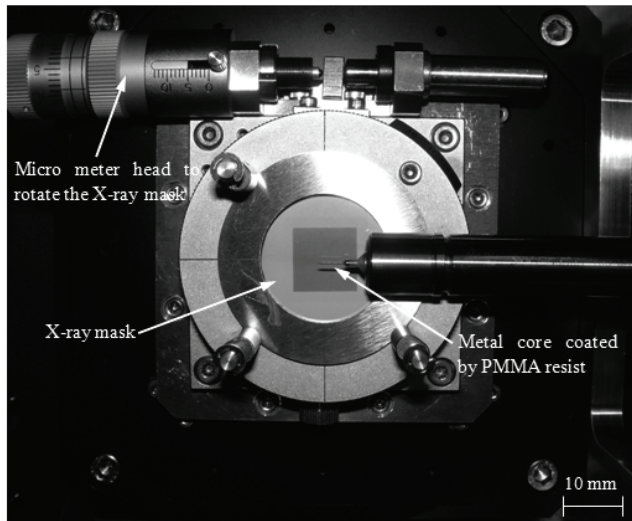
Fig. 26. Picture of back view of X-ray exposure chamber.
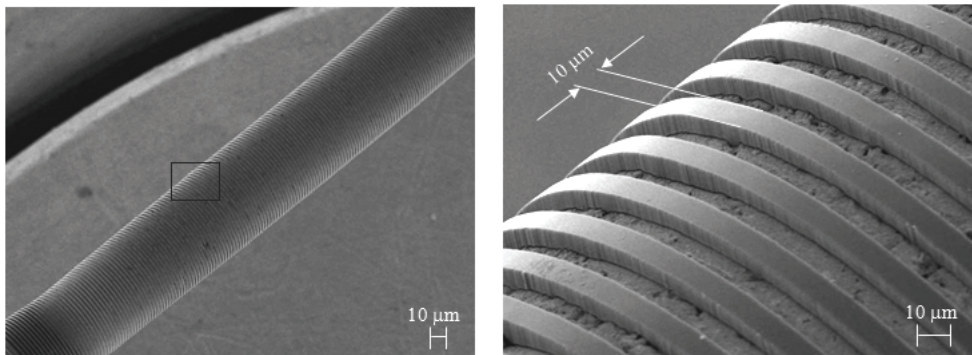
Fig. 27. SEM image of PMMA mold master.
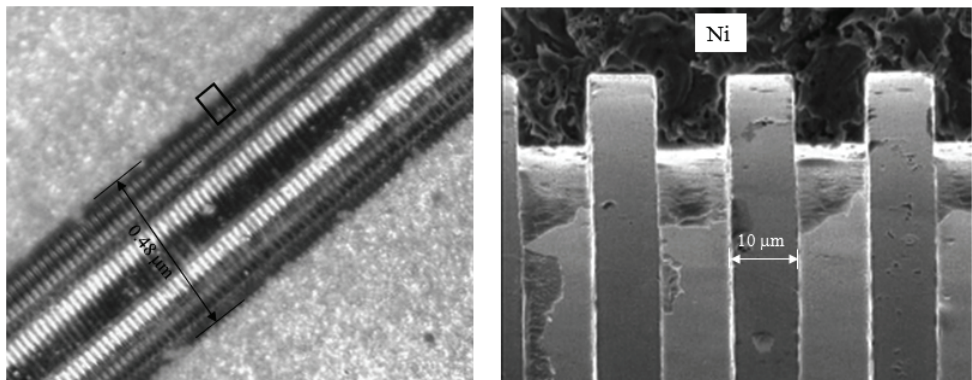
Fig. 28. SEM image of Ni mold.
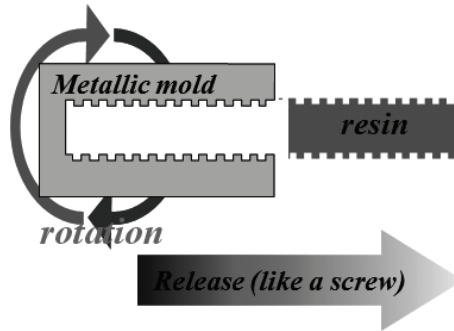
Fig. 29. Conept of worm deolding.
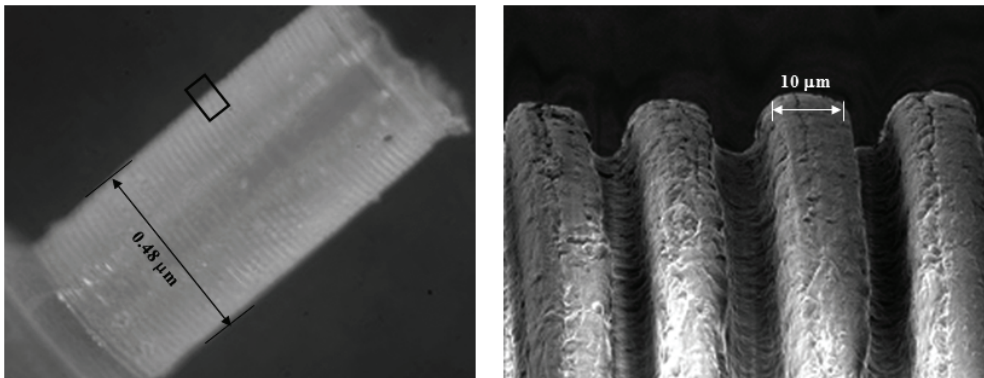
Fig. 30. SEM image of resin core.

metallic mold, and inside of the Ni mold shown in Figs. 27 and 28, the sharpness could not be checked for the edge portion of the microstructure of the spiral form. This might be a gas that evaporated the LCP. Furthermore, the edge portion might have become round by frictional heat during the unscrewing of the demolding process (Mekaru et al. 2005).

All the screw portion of the LCP core was covered by copper electroforming after pre-treating and degreasing. Silver paste was applied to the both sides of the microcoil as electrodes, and the coil line part was coated by epoxy resin for insulated protection. Figure 31 shows the completed spiral microcoil with 1 mm in length, 0.48 mm in diameter, and 15 turns. The spiral copper coil lines could be seen clearly. This inductance and the Q factor at 1GHz were 91 nH and 5.8, respectively (Mekaru et al. 2005). From these results, LIGA process has very advantage to produce high aspect ratio microstructures.

6. Conclusions

We have fabricated spiral coil with high aspect ratio coil lines for solenoid type electromagnetic actuator using 3D X-ray lithography and metallization techniques. Using these techniques, we succeeded in producing a spiral structure with 10 μm width coil line structures with maximum aspect ratio of about 5. We also succeeded in electroforming copper in the high aspect ratio structure and forming a coil by isotropic copper etching. Then, we could be obtained coil pattern on acrylic pipe with high aspect ratio.
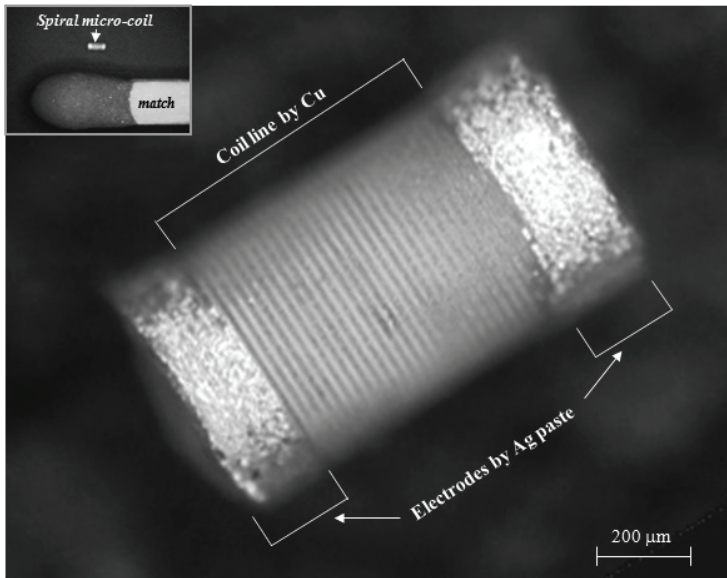
Fig. 31. SEM image of copleted spiral microcoil.

In addition, we produced a measurement system to measure the suction force produced by these electromagnetic actuators. The results of these suction force measurements enabled us to confirm the results of calculation values. The measurement results were in relatively good agreement with the calculated values.

We also attempted to fabricate microcoil with diameters of less than 1 mm. Using a dipping method, photoresist thickness of over 100 μm were achieved using highly viscous solution and controlling the centrifugal force. We succeeded in producing a spiral microcoil with 30 μm width with an aspect ratio of about 2 using same fabrication process. And, we achieved development of a method to fabricate the 3D spiral microcoil using LIGA process. The size of the microcoil was 1 mm in length and 0.48 mm in diameter.

Using these techniques, we were able to fabricate microcoils with high aspect ratio coil lines. Thus, it is very expected that electromagnetic actuators with high suction force could be manufactured despite their miniature size.

7. Acknowledgment

8. References

Becker, E. W.; Ehrfeld, W.; Hagmann, P.; Maner, A. & Munchmeyer, D. (1986). Fabrication of microstructures with high aspect ratios and great structural heights by synchrotron

radiation lithography, galvanoforming, and plastic moulding, *Microelectron. Eng.*, Vol. 4, No. 1, 35-56, ISSN: 0167-9317

Ando, A.; Amano, S.; Hashimoto, S.; Kinoshita, H.; Miyamoto, S.; Mochizuki, T.; Niibe, M.; Shoji, Y.; Terasawa, M.; Watanabe, T. & Kumagai, N. (1998). Isochronous storage ring of the New SUBARU project, *J. Synchrotron Rad.*, Vol. 5, Part 3, 342-344, ISSN: 0909-0495

Mekaru, H.; Utsumi, Y. & Hattori, T. (2001). Beam line BL11 for LIGA process at the NewSUBARU, *Nucl. Instrum. Methods A*, Vol. 467-468, Part 1, 741-744, ISSN: 0168-9002

Mekaru, H.; Utsumi, Y. & Hattori, T. (2002). Quasi-3D microstructure fabrication technique utilizing hard X-ray lithography of synchrotron radiation, *Microsyst. Technol.*, Vol. 9, No. 1-2, 36-40, ISSN: 1432-1858

Mochizuki, H.; Mekaru, H.; Kusumi, S.; Sato, N.; Yamashita, M.; Shimada, O. & Hattori, T. (2007). Design of solenoidal electromagnetic microactuator utilizing 3D X-ray lithography and metallization, *Microsyst. Technol.*, Vol. 13, No. 5-6, 547-550, ISSN: 1432-1858

Matsumoto, Y.; Setomoto, M.; Noda, D. & Hattori, T. (2008). Cylindrical coil created with 3D X-ray lithography and metallization for electromagnetic actuators, *Microsyst. Technol.*, Vol. 14, No. 9-11, 1373-1379, ISSN: 1432-1858

Setomoto, M.; Matsumoto, Y.; Yamashita, S; Noda, D. & Hattori, T. (2008). Fabrication of spiral micro coil lines for electromagnetic actuator, *J. Adv. Mech. Des. Syst. Manuf.*, Vol. 2, No. 2, 238-245, ISSN: 1881-3054

Noda, D.; Matsumoto, Y.; Setomoto, M. & Hattori, T. (2007). Fabrication of coil lines with high aspect ratio for electromagnetic actuators, *Proceedings of the 2007 IEEE International Symposium on Micro-Nano Mechatronics and Human Science*, pp. 436-441, ISBN: 978-1-4244-1858-9, Nagoya University, November 2007, Nagoya, Japan

Noda, D.; Setomoto, M. & Hattori, T. (2008a). Fabrication of microcoil with narrow and high aspect ratio lines for electromagnetic actuators, *Proceedings of the 2008 IEEE International Symposium on Micro-Nano Mechatronics and Human Science*, pp. 219-224, ISBN: 978-1-4244-2919-6, Nagoya University, November 2008, Nagoya, Japan

Yamashita, S; Matsumoto, Y.; Idei, K.; Okuda, K.; Noda, D. & Hattori, T. (2006). Fabrication of a cylindrical microcoil line with high aspect ratio for electromagnetic actuators, *Proceedings of the 2006 IEEE International Symposium on Micro-Nano Mechatronics and Human Science*, pp. 497-502, ISBN: 1-4244-0718-4, Nagoya University, November 2006, Nagoya, Japan

Noda, D.; Yamashita, S; Matsumoto, Y.; Setomoto, M. & Hattori, T. (2008b). Fabrication of high aspect ratio microcoil using dipping method, *J. Adv. Mech. Des. Syst. Manuf.*, Vol. 2, No. 2, 174-179, ISSN: 1881-3054

Noda, D.; Matsumoto, Y.; Setomoto, M. & Hattori, T. (2008c). Fabrication of microcoils using X-ray lithography and metallization, *IEEJ Transactions on Sensors and Micromachines*, Vol. 128, No. 5, 181-185, ISSN: 1341-8939

Mekaru, H.; Kusumi, S.; Sato, N.; Yamashita, M.; Shimada, O. & Hattori, T. (2004). Fabrication of mold master for spiral microcoil utilizing X-ray lithography, *Jpn. J. Appl. Phys.*, Vol. 43, No. 6B, 4036-4040, ISSN: 1347-4065

Mekaru, H.; Kusumi, S.; Sato, N.; Shimizu, M.; Yamashita, M.; Shimada, O. & Hattori, T. (2005). Development of three dimensional LIGA process to fabricate spiral microcoil, *Jpn. J. Appl. Phys.*, Vol. 44, No. 7B, 5749-5754, ISSN: 1347-4065

Mekaru, H.; Kusumi, S.; Sato, N.; Shimizu, M.; Yamashita, M.; Shimada, O. & Hattori, T. (2007). Fabrication of a spiral microcoil using 3D-LIGA process, *Microsyst. Technol.*, Vol. 13, No. 3-4, 393-402, ISSN: 1432-1858

Micro-Electro-Discharge Machining Technologies for MEMS

Kenichi Takahata
University of British Columbia, Vancouver
Canada

1. Introduction

Advances in micromachining techniques have led to the evolution of micro-electro-mechanical systems (MEMS). These techniques are typically based on semiconductor manufacturing processes, which offer various advantages such as batch manufacturing of miniaturized devices and monolithic integration of microelectronics with the devices. Surface micromachining has been used to construct complex microstructures, but since the structural geometries of these microstructures are two-dimensional, their mechanical abilities are often limited. This constraint has been addressed by the use of bulk micromachining techniques that involve etching and deposition processes. Anisotropic wet etching (Sato et al., 1998) and deep reactive ion etching (Laermer & Urban, 2005) have been widely used to create three-dimensional (3-D) geometries in MEMS. However, these processes are severely limited in their material options. As for deposition, electroplating is widely used to form 3-D metallic microstructures, but practical materials are limited to selected metals and alloys. In contrast, certain stainless steels and shape memory alloys have been commonly used for a variety of biomedical and implant devices such as stents and surgical devices. These materials have not been leveraged as much as silicon in MEMS, however, largely because they are not compatible with MEMS fabrication processes. As these examples indicate, there is an explicit gap between the diversity of engineering materials and the ability to use them in the design/fabrication of MEMS; bridging this gap is expected to create new opportunities in the field.

Micro-electro-discharge machining (μEDM) is a powerful bulk micromachining technique, as it is applicable to any type of electrical conductor, including all kinds of metals and alloys as well as doped semiconductors. μEDM is a non-contact machining technique, hence it can be easily applied to thin, fragile, and/or soft materials regardless of their mechanical properties. Complex 3-D shapes can be achieved through numerical control (NC) systems with high-precision positioning stages. These unique features and the extensive material base available to μEDM have led to the process being leveraged for industrial applications, such as ink-jet nozzle fabrication (Allen & Lecheheb, 1996), micromachining of magnetic heads for digital VCRs (Honma et al., 1999), and micromechanical tooling (Wada & Masaki, 2005). In recent years, the technique has been increasingly utilized for MEMS fabrication to exploit a broad range of engineering materials that are incompatible with standard MEMS processes, overcoming the common constraint in MEMS, i.e., lack of diversity of bulk materials available for their fabrication (Takahata & Gianchandani, 2007).

This chapter discusses the basic principles and advanced technologies of µEDM in Subsections 2 and 3, respectively. Subsection 4 describes how the technique has been used to realize various types of MEMS devices, while introducing selected applications it has enabled. The discussion involves not only how µEDM has enabled MEMS but also how MEMS have enabled advanced µEDM processes, which is included in Subsection 3.

2. Features and challenges

µEDM utilizes pulses of thermomechanical impact induced by a miniaturized electrical discharge generated between a microscopic electrode and a workpiece while both are immersed in dielectric liquid (Masaki et al., 1990). The miniaturized arc discharge, which usually involves a pulse energy of 0.01-10 µJ with duration of 10-100 ns, is reported to reach a temperature of several 1000s K (Dhanik & Joshi, 2005), locally melting and evaporating the material at the arc spot (Fig. 1). The heat generated by an arc also leads to instant evaporation of the dielectric liquid (typically kerosene-based oil or ultrapure water). This evaporation creates pressure waves that blow the melted material away, leaving a crater-like cavity on the workpiece surface. Machining is performed by repeating this unit removal by a single pulse at high frequencies. Burr-free, high-aspect-ratio (>20) micromachining achieved through this technique can produce features of a few microns with submicron tolerances. High-precision machining of high-aspect-ratio cylindrical electrodes with diameters of 3-300 µm, usually of tungsten or its alloy, is available using a µEDM technique called wire electro-discharge grinding, or WEDG (Masuzawa et al., 1985).
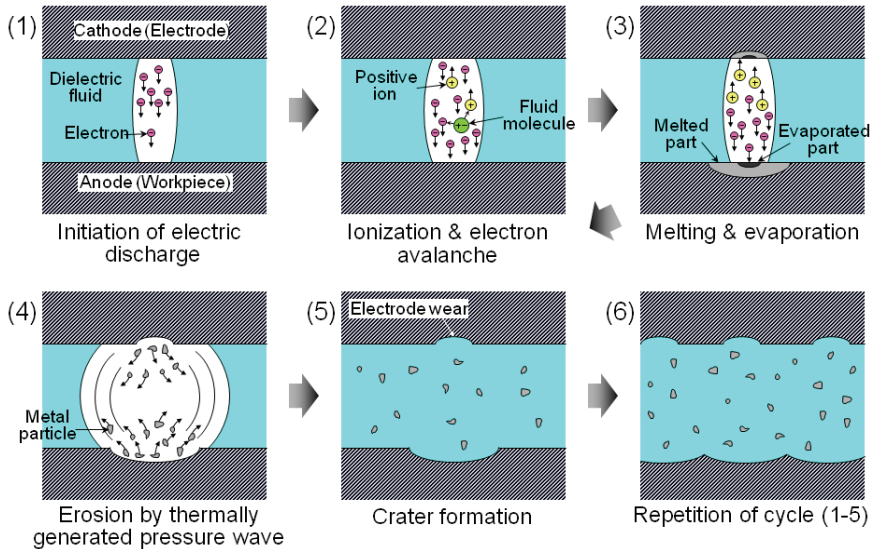
Fig. 1. µEDM principle illustrated.

3-D microstructures are machined with NC of the relative position between an electrode and a workpiece. An example of a commercial µEDM machine equipped with 3-axis stages with 100-nm resolution and a WEDG unit is shown in Fig. 2, along with a typical machining method for arbitrary shapes using a cylindrical electrode, which is usually rotated during

the process. Figure 3 shows a sample structure fabricated by a prototype 5-axis system with two rotational mechanisms (Takahata et al., 1997), demonstrating real 3-D micromachining. Another related technique is wire µEDM, where brass wire is typically used as an electrode that is scanned to cut structures out from the workpiece (Ho et al., 2004).
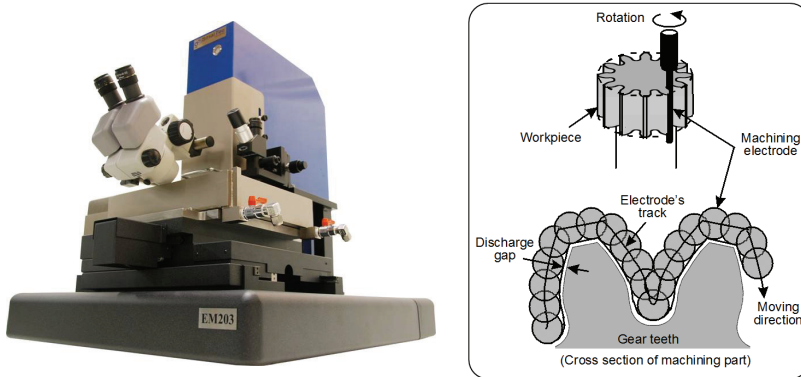
Fig. 2. (a: left) A commercially available µEDM machine (EM203, SmalTec International, USA, image courtesy SmalTec); (b: right) a method for creating arbitrary shapes in a bulk-metal workpiece using a conventional cylindrical electrode (Takahata et al., 1999) © 1999 IEEE.

Fig. 3. A sample stainless-steel 3-D structure machined using a 5-axis µEDM system.

Relaxation-type resistor-capacitor (R-C) circuitry has been predominantly employed to control pulse generation/timing in µEDM systems (Masuzawa & Sata, 1971). The discharge energy (E) of a single pulse provided through this type of circuit can be expressed as:

$$E = \frac{1}{2}(C + C_p)V^2 \qquad (1)$$

where C is the capacitance of the circuit, C_p is the lumped parasitic capacitance present in parallel to C, and V is the machining voltage. There is a trade-off between the smoothness of machined surfaces and machining speed, as a function of discharge energy; the greater the discharge energy, the larger the volume removed by a single pulse (i.e., faster machining) but the rougher the surfaces. To achieve stable discharge, the voltage is typically set above 60 V, up to 100 V. For finer machining, C is often set to be zero, using C_p only. Therefore, in order to minimize removal size and surface roughness, it is critical to reduce C_p, as it directly impacts the discharge energy and thus the unit removal volume by a single discharge. The machining systems are configured to achieve minimal C_p resulting from their

components (e.g., using bulk ceramics as mechanical parts). Another common type of pulse generator for conventional EDM is based on transistor (FET) switching circuitry. This type of generator can generate pulses at higher frequencies and hence remove material faster, but is in general limited in generating the short pulses required in µEDM. However, efforts have been made to overcome this issue (Hana et al., 2004).

Despite its excellent capability in terms of precision/tolerance, surface quality and complex 3-D formation, µEDM has not achieved widespread use in product manufacturing primarily because of its productivity drawbacks. The throughput of conventional µEDM is inherently low because it is a serial process that uses a single electrode tip to machine and produce structures individually. Another related issue is electrode wear, which tends to degrade not only machining precision but also productivity when replacement of electrodes is necessary. Applications of the technique to product manufacturing implemented in the past fall into two types as summarized in Fig. 4. One is direct machining of end products. In this case, although the wide range of material options is provided by technique, removal volume needs to be small in order to achieve an acceptable level of production throughput. A representative example of this case is trimming of the magnetic heads. Another type of application is machining of replication tools (e.g., micro molds for injection molding). µEDM allows one to fabricate robust tools using hard alloys commonly used in molding dies (e.g., tool steel and super-hard alloy). The use of such tools enables volume manufacturing of replicated products at low costs; however material options in the replication processes are limited. It is evident that there is a trade-off relationship between productivity (or removal volume) and material as shown in Fig. 4.
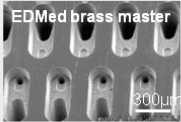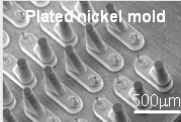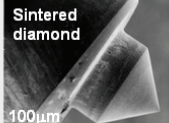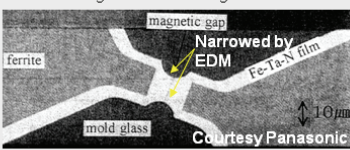
Fig. 4. Application examples of conventional µEDM: Machining of mechanical tools and direct machining of end products.

One straightforward approach to addressing these constraints is to increase removal rate. This process involves increasing pulse frequency while keeping single discharge energy low, in order to achieve faster removal without sacrificing machining quality (Hana et al., 2004). However, throughput scales down as the number of structures to be machined

increases. A conceptually different approach has been investigated to convert the machining mode of μEDM from serial to parallel or batch. This conversion has been realized by utilizing arrays of microelectrodes to implement planar processing, as will be discussed in the next section. This approach offers opportunities to achieve not only high-throughput production, due to its high parallelism, but also compatibility with other planar microfabrication techniques based on lithography processes, which are the mainstream of MEMS manufacturing. The latter feature potentially enables the integration of μEDM with standard MEMS technologies, realizing heterogeneous microstructures and devices with unique functionalities and performance.

3. Advanced μEDM enabled by photolithography and MEMS

This section discusses new types of μEDM techniques developed to enable batch micromachining and manufacturing for microdevices and their components. It has been demonstrated that the use of microelectrode arrays is a very effective route to reach this goal. This approach has been extended to a technique that leverages MEMS actuators to further advance the capability of the technique. Details follow below.

3.1 Batch-mode μEDM

The concept of parallel/batch μEDM processing enabled by electrode arrays is illustrated in Fig. 5. Photolithographic methods offer various paths to the fabrication of such arrays with arbitrary patterns on a substrate. The following are the major advantages of using lithographically fabricated electrodes over conventional single electrodes.

- Parallel machining of multiple structures for high-throughput production.
- Photo-patterned electrodes are precisely arranged on the substrate and have high structural uniformity across the arrays, offering high precision and uniformity in the machined products.
- Batch production of electrode components with high volume and low cost.
- One electrode only is used for machining one structure, in contrast to the conventional serial-processing method (i.e., one electrode for all structures), minimizing consumption/wear per electrode, and machining errors.

A common approach to electrode fabrication is to use a patterned photoresist layer as a mold for electroplating of electrode material. To deal with machining that involves deep or 3-D structures, the electrodes are often required to be high-aspect-ratio microstructures (HARMST). The process that provides HARMST with the highest precision among other techniques is deep X-ray lithography, known as LIGA (German acronym for lithography, electroforming and molding). A group from Germany and Switzerland first demonstrated μEDM using LIGA electrodes of electroplated copper with arbitrary patterns (Ehrfeld et al., 1996). In this application, the final step of molding is omitted, i.e., the electroplated structures are the end product, serving as the μEDM electrodes. This approach was advanced in the US, where LIGA fabricated electrode arrays were successfully utilized to demonstrate parallel machining of microstructures (Takahata et al., 2000; Takahata & Gianchandani, 2002).

A LIGA process used for electrode fabrication is shown in Fig. 6. This process, developed at the University of Wisconsin-Madison, utilizes thick, solid polymethylmethacrylate (PMMA) sheet as the photoresist for synchrotron X-ray lithography (Guckel, 1998). One important
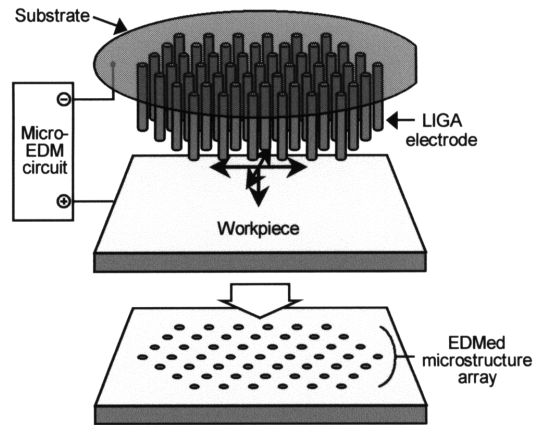
Fig. 5. Concept of batch-mode μEDM (Takahata & Gianchandani, 2002) © 2002 IEEE.

feature of μEDM that makes the use of this type of HARMST electrode feasible is that the process does not produce contact forces to the electrodes that lead to peeling of the structures from the substrate. Figure 7a shows an example of a 20×20 array of HARMST electrodes of electroplated copper with 20-μm diameter, 60-μm pitch, and 300-μm structural height. For μEDM, the electrode substrate is mounted on the X-Y stage of a μEDM machine, and a workpiece held on the vertical Z stage of the machine is advanced into the arrays along the axial direction of the electrodes to perform batch-mode machining. Results obtained with stainless-steel samples are shown in Figs. 7b and 7c. Figure 8 shows a honeycomb structure fabricated in 125-μm-thick graphite sheet by using arrayed electrodes with hexagonal pattern shape. Since graphite has high thermal conductivity, such structures may be suitable for heat exchange applications. In conventional μEDM, a
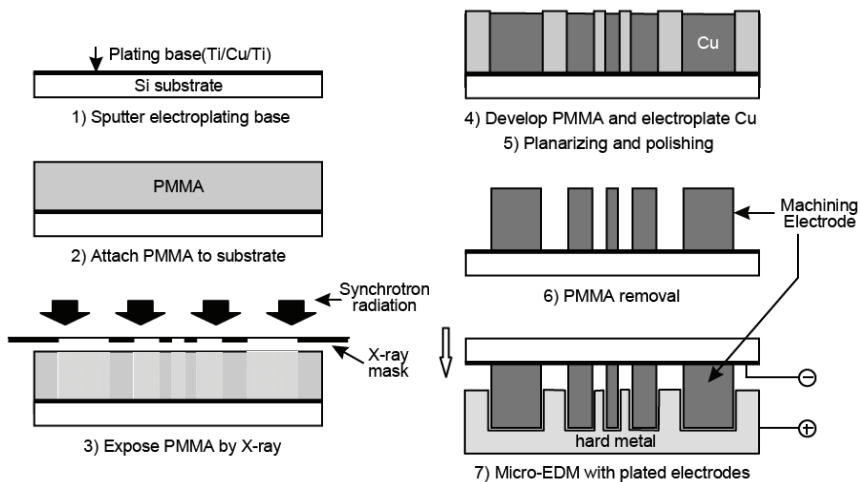
Fig. 6. A LIGA process for electrode fabrication, and subsequent μEDM using the electrodes (Takahata et al., 1999) © 1999 IEEE.
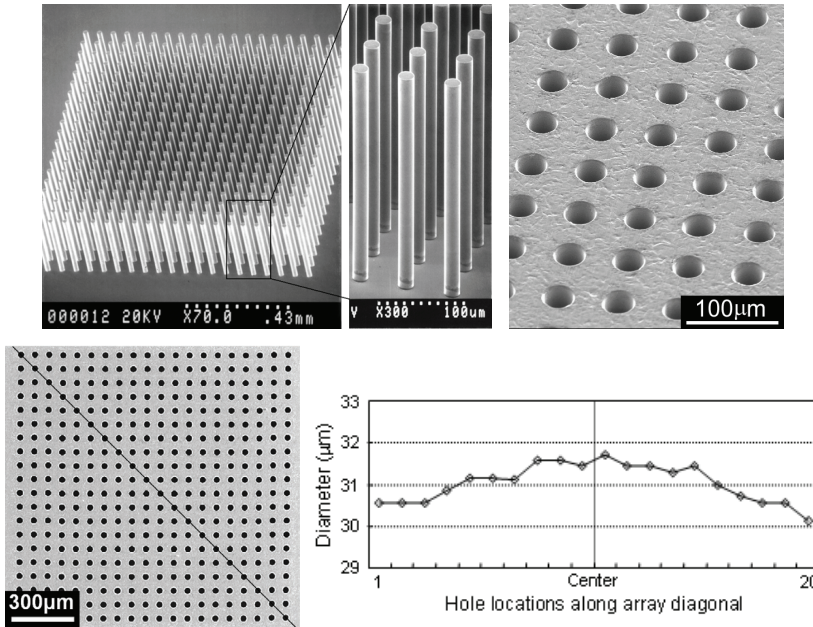
Fig. 7. (a: upper left) A 20×20 array of LIGA fabricated copper electrodes; (b: upper right) through-holes batch machined in 50-μm-thick stainless steel using the array; (c: lower) a top-view image of the machined hole array, and measured variation of hole diameter along the array diagonal shown in the image (Takahata & Gianchandani, 2002) © 2002 IEEE.
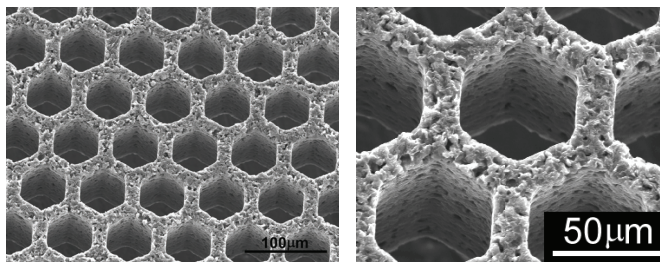
Fig. 8. Graphite honeycomb microstructures (hexagonal pitch of 70 μm and wall thickness of 16 μm) formed by batch-mode μEDM (Takahata & Gianchandani, 2002) © 2002 IEEE.

cylindrical electrode is rotated during the machining process in order to increase uniformity and prevent local welding to the workpiece. This rotation is clearly not possible when using arrayed electrodes. Instead, the electrodes are placed on a vibrator that dithers them along the axis of approach.

Figure 9 shows microchannels fabricated by the sequential application of arrayed electrodes of three different shapes, each of which contributes to a structural "layer". The layer-to-layer alignment resolution of 100 nm is afforded by the precision of workpiece movement in μEDM and the tight dimensional tolerance of LIGA. Note that each nozzle in the figure has a 40° taper at the top. This was created by a scrolling motion of the electrode array. The

result demonstrates that the combination of lithographically fabricated electrodes and µEDM can be used to create complex multi-layer structures in bulk metals.
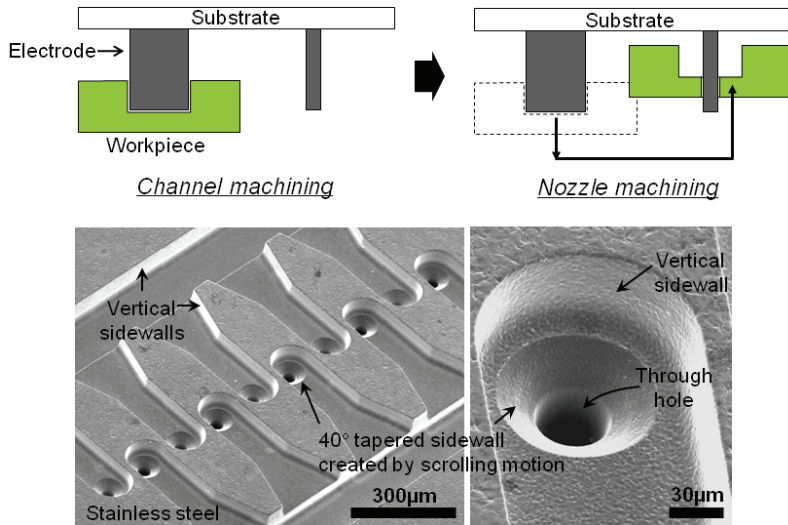
Fig. 9. (a: upper) Sequential application of electrodes formed on a substrate; (b: lower) microchannels fabricated by the technique (Takahata & Gianchandani, 2001) © 2001 IEEE.

Although the presence of multiple electrodes can increase spatial parallelism, temporal parallelism is not achieved if a single pulse discharge circuit is used, because only one electrode fires at a time. In this case, as the number of electrodes in the array increases, the pulse frequency at each of the electrodes drops. This means that the removal rate at individual electrodes will decrease, thus the throughput does not scale up as the number of electrodes increases. Further gains in throughput can be achieved by partitioning the electrode array into segments, each of which is controlled by a separate pulse generation circuit (Takahata & Gianchandani, 2002). The use of monolithically partitioned electrode arrays coupled with multiple *R-C* circuits through thin film interconnect patterned on their substrate demonstrated parallel discharging at the arrays, maximizing pulse frequencies at individual electrodes for accelerated processing. Figure 10 shows LIGA electrode arrays with interconnect to individual electrodes as well as batch-produced micro gears cut from 70-µm-thick WC-Co super-hard alloy sheet using the arrays. This experiment showed improvement in throughput by >100× compared to that in traditional serial µEDM. This arrangement permits on-chip parasitic capacitance present in each of the partitioned electrodes to be used as a capacitor of the *R-C* circuit, which is highly amenable to large-sized arrays because all the pulse control circuit elements can be integrated.

The machining process produces debris containing the particles removed from the workpiece and carbon residues produced by pyrolysis of dielectric EDM oil during the process. Debris removal from the machining region is critical to perform well-controlled µEDM, as its presence in the region tends to cause enlarged discharge gaps and tolerances as well as irregular continuous arcs, which thermally damage both the electrode and the workpiece. In traditional µEDM using a single cylindrical electrode, the rotation of the

electrode and open space around it promote the dispersion of debris from the machining region. In batch-mode μEDM using planar electrodes, however, removal of debris becomes a challenging issue as the tool movement is limited to the dither motion. In addition, a large planar form of the tool limits the flow of EDM fluid and options for flushing. It has been reported that a two-step hydrodynamic debris removal technique can address this issue effectively, resulting in improved surface and edge finish, machining time, and tool wear over the method that uses standard vertical dither flushing (Fig. 11).
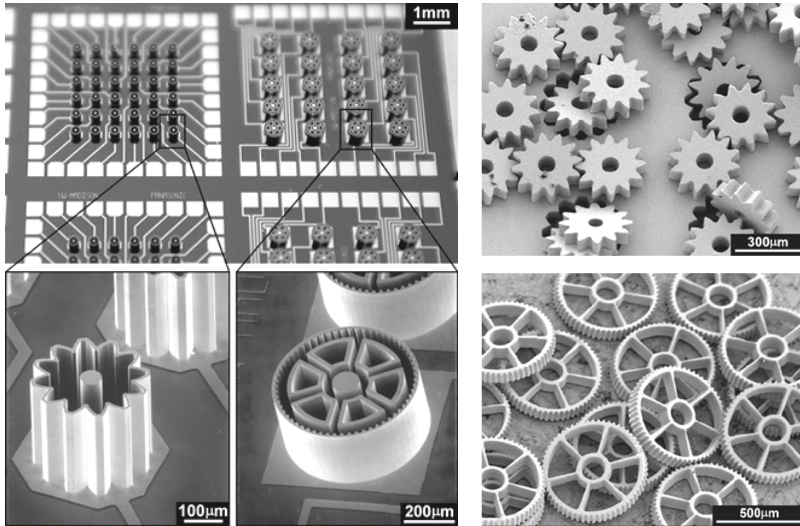
Fig. 10. (a: left) Partitioned copper electrode arrays with interconnect; (b: right) super-hard alloy gears batch produced using the arrays (Takahata & Gianchandani, 2002) © 2002 IEEE.
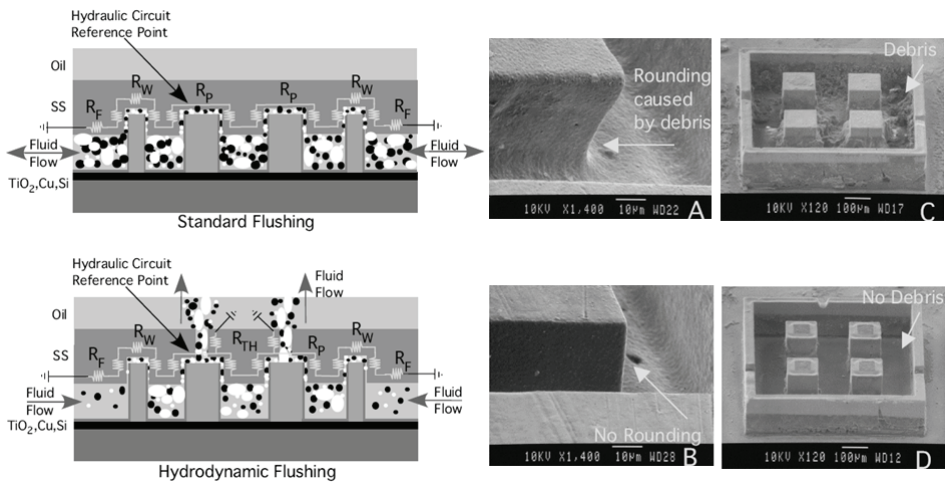
Fig. 11. Cross-section of hydraulic resistance circuit and machined result with (top) standard dither flushing and (bottom) hydrodynamic flushing (Richardson et al., 2006) © 2006 IEEE.

3.2 MEMS-based µEDM

Although the batch-mode method discussed above demonstrated improved throughputs, it is still limited in the machinable area due to the available size of the substrate holding the arrays. This implementation requires a high-cost µEDM system and an NC stage of the system to advance planar electrode arrays into the workpiece, which also limits the substrate size to be compatible with the stage. Toward µEDM of large-area samples (e.g., shape memory alloy foil for microactuator fabrication and hard-alloy plates to form microstructured molding dies), a new technique called M³EDM (MEMS-based micro-EDM) has been developed. This µEDM method uses micromachined actuators with movable planar electrodes that are fabricated directly on the workpiece material using lithography techniques and actuated to perform machining. The actuation leverages electrostatic forces generated by a machining voltage applied between the electrode and the workpiece (Alla Chaitanya & Takahata, 2008a). This approach offers an opportunity to eliminate NC machines from the process, achieving high scalability to very large areas for high-throughput, low-cost micromanufacturing.

Figure 12 illustrates the mechanical and electrical behaviors of the electrode device in the machining process. The planar electrodes are microfabricated so that they are suspended by the anchors through tethers above the surfaces of the conductive workpiece with a relatively large gap. The application of machining voltage produces electrostatic forces that drive the electrodes towards the workpiece. With properly designed structures at a selected voltage, the phenomenon known as "pull-in" takes place, when the restoring spring force through the tethers can no longer balance the electrostatic force as the gap spacing decreases. This results in a breakdown and produces a pulse current due to a discharge from the capacitors (external capacitor C and built-in capacitor C_b in Fig. 12) that removes the material at the local spot. The discharge lowers the voltage between the electrode and the workpiece, releasing the electrode. Simultaneously, the capacitors are charged through the resistor, restoring the voltage at the gap and inducing the electrostatic actuation again. This sequence of pull-in and release of the electrode is used to achieve self-regulated generation of discharge pulses that etch the material. This approach that uses electrostatic actuation is suitable for selected electrode structures and applications requiring relatively shallow machining due to the limitation of its actuation range. An M³EDM method that uses downflow of dielectric EDM fluid for electrode actuation has been demonstrated to overcome this limitation in the electrostatic actuation method (Alla Chaitanya & Takahata, 2008b).

The fabrication of movable electrodes on workpieces was implemented by a combination of film lamination, photolithography, and wet etching of 18-µm-thick copper foil used as the structural material of the electrode device. Figure 13 shows the electrode devices formed on a piece of dry-film photoresist that can be laminated on the target workpiece. The photoresist serves as the sacrificial layer, which is dissolved to release the electrode structures. This method, where electrodes are supplied with laminatable film, is potentially applicable to samples that have non-planar surfaces to be machined, or those whose sizes are incompatible with photolithography tools, making direct fabrication of the devices on them difficult. Figure 14a shows another electrode fabrication process that uses liquid photoresist. This process was developed to incorporate arbitrary features on the backside of
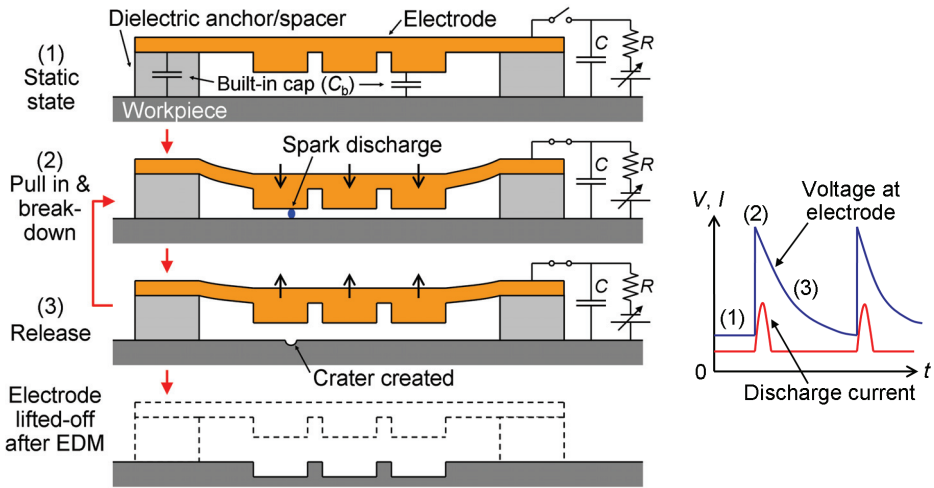
Fig. 12. (a: left) Cross-sectional view of M³EDM and its process steps; (b: right) dynamic behavior of discharge voltage and current corresponding to the steps (Alla Chaitanya & Takahata, 2008a) © 2008 IEEE.
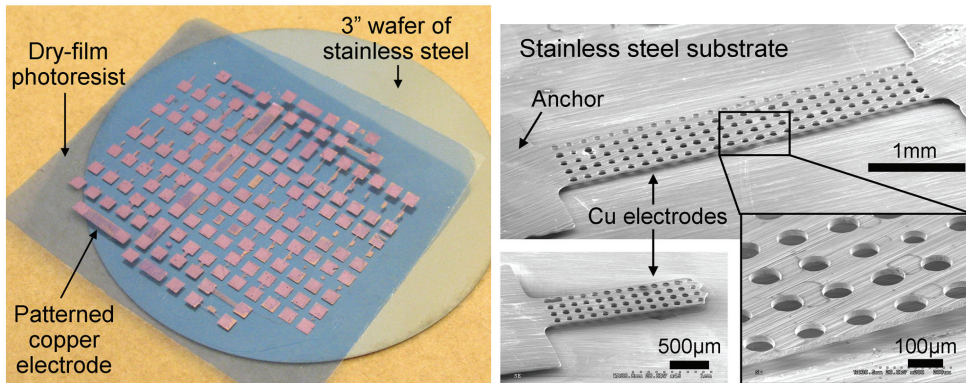
Fig. 13. (a: left) A 6×6 cm² piece of sacrificial dry-film photoresist with patterned electrode devices; (b: right) examples of fabricated electrode devices with fixed–fixed and cantilever configurations (Alla Chaitanya & Takahata, 2008b) © 2008 IOP Publishing Ltd.

the planar electrodes for μEDM of custom patterns. The fabricated electrode arrays shown in Fig. 14b were designed to have a crab-leg configuration to support planar electrodes with larger areas. Pattern transfer to stainless-steel wafers (which served as the workpieces in this experiment) was successfully demonstrated (Fig. 14c). The development of M³EDM is currently in progress toward enabling high-precision, cost-effective batch μEDM for large-area micromachining of bulk metals and alloys.
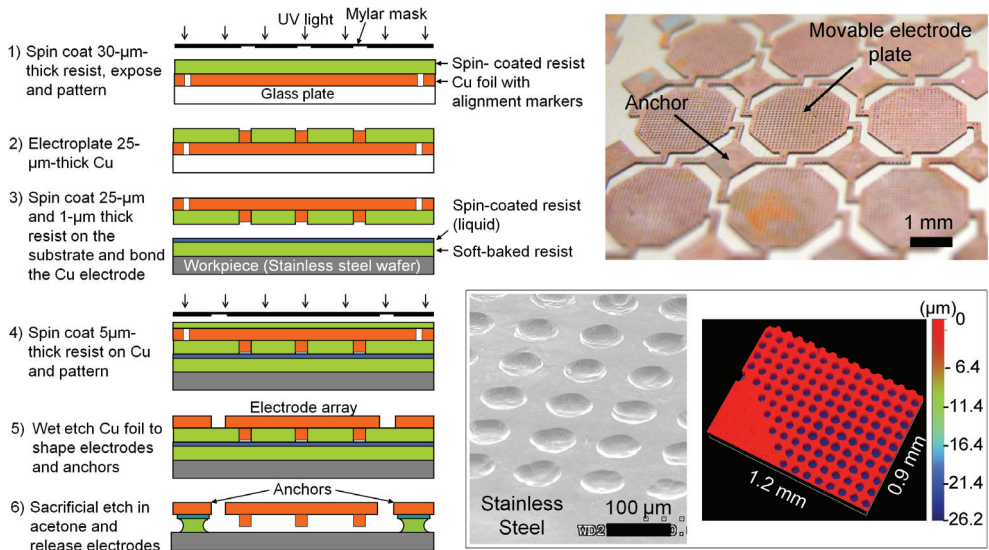
Fig. 14. (a: left) Fabrication process flow for double-layer electrode devices; (b: upper right) a fabricated array of planar electrodes that hold arrayed microstructures on the backside of the electrodes; (c: lower right) batch-machined structures created in stainless steel (Alla Chaitanya & Takahata, 2009) © 2009 IEEE.

4. Application

The application of μEDM to MEMS and micro-scale devices has been primarily driven by demand for use of bulk metals and alloys with unique performance that are difficult to be micromachined using lithography and etching processes. The most common application involves micromachining of mechanical components of the devices. For example, a rotor and bearing parts were machined using μEDM and assembled to construct a micro air turbine (Masaki et al, 1990). Another micro turbine with a more complex design was reported in (Peirs et al., 2002), where the components were fabricated by a combination of μEDM and mechanical machining. Although stainless steel was the main structural material in these particular devices, the technique allows one to select harder alloys to construct devices in order to achieve more robust mechanical systems; an example is described in Subsection 4.1. An implantable device called a stent is presented in Subsection 4.2 as another example of mechanical devices fabricated by μEDM. The technique has been used for bulk micromachining of heavily doped single crystal silicon (Reynaerts et al., 1997; Heeren et al., 1997), and it has been integrated with lithography processes to construct an inertial sensor (Reynaerts et al., 2000). Needle-shaped HARMST neural electrode arrays have been fabricated from highly doped silicon using a combination of wire μEDM and isotropic wet etching (Tathireddy et al., 2009). μEDM has also been leveraged to produce components from permanent magnet and ferromagnetic material that are incorporated in electromagnetic MEMS sensors and actuators (Grimes at al., 2001; Fischer et al., 2001). In addition to the machining (and assembly) of mechanical/magnetic components mentioned above, various efforts have utilized μEDM to construct devices with electrical

functionalities. A common need for this application is to integrate dielectric materials with μEDM structures to create electrical partitions and circuits in the devices. This is a challenging task, because μEDM by itself does not allow electrical isolation, since all the mechanically connected features are also all electrically connected. This issue has been addressed through different fabrication approaches, described in the development of the devices presented in Subsections 4.3 to 4.6, i.e., antenna stent, scanning micro Kelvin probe, electromagnetic flow sensor, and capacitive pressure sensor.

4.1 Self-propelled micromachine

A chain-type, self-propelled micromachine has been developed for the maintenance of power plants, where chained micromachines perform the inspection of the outer surfaces of tube banks (Takeda et al., 2000). The traveling device of the micromachine uses a micro reducer based on a paradox planetary gear system to achieve high torque with a micromotor, as well as magnetic wheels to achieve strong traction (Fig. 15). The gears and other mechanical components were fabricated by μEDM of hard alloys such as high-carbon tool steel and WC-Ni-Cr super-hard alloy. The fabrication precision in μEDM of the gear components was reported to be within 0.4 %, with standard deviation of 0.127. Here the machining error is largely associated with the wear of electrodes. The developed micro reducer with as-machined planetary gears (without any surface coating) and oil-lubricated rolling bearing was observed to sustain sufficient performance after 5×10^6 rotations (Takeuchi et al., 2000). These results demonstrate that μEDM is a practical fabrication technique for realizing high-precision mechanical systems with high robustness. A major drawback to the manufacturing of such systems is the need for assembly of the machined components. An example addressing this issue was reported in (Sun et al., 1996), where a micro turbine device with rotor, bearing, and base components was constructed in a pre-assembly manner using a process based on μEDM and micro ultrasonic machining.
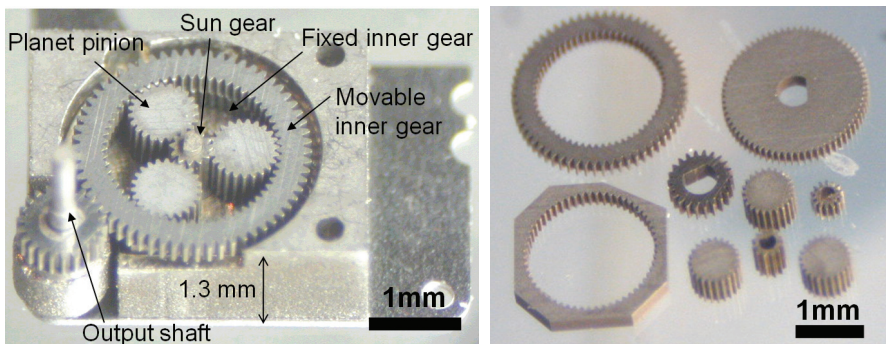
Fig. 15. (a: left) Gear reduction system developed for a self-propelled micromachine; (b: right) WC-Ni-Cr super-hard alloy gear components of the system produced by μEDM; most of the gears have a module number of 0.03 (images courtesy of Dr. Narito Shibaike).

4.2 Micromechanical stents

Stents are mechanical devices that are chronically implanted into arteries in order to physically expand and scaffold blood vessels that have been narrowed by plaque accumulation. The vast majority of stents are manufactured by laser machining of metal

tubes made of biocompatible stainless steel or shape memory alloy, creating mesh-like walls that allow the tube to be expanded radially upon the inflation of an angioplasty balloon (Kathuria, 1998). The use of μEDM is another option for cutting metal microstructures. It has been shown that tubular stents can be fabricated from planar stainless-steel foil using μEDM (Takahata & Gianchandani, 2004a). The planar design of the stent also permits the use of the batch machining method discussed in Subsection 3.1. The planar pattern has two longitudinal side-beams, connected transversely by expandable cross-bands, each of which contains identical involute loops (Fig. 16a). In a manner identical to that used with commercial stents, the stent was deployed by inflating an angioplasty balloon threaded through the planar structure such that the transverse bands alternated above and below it; the structure was plastically deformed into a cylinder shape when deployment was completed. Figure 16b shows an expanded stent with the balloon removed. Mechanical tests indicated that the developed stent had almost the same radial strength as a commercial stent (Guidant Multilink Tetra™) tested for comparison, even though the developed stent had a wall thickness (of 50 μm) that was approximately one-half that of the commercial stent. The radial stiffness was similar when the loading was applied at two extreme orientations, i.e., perpendicular to the original plane of the pre-expansion planar microstructure and parallel to the plane.
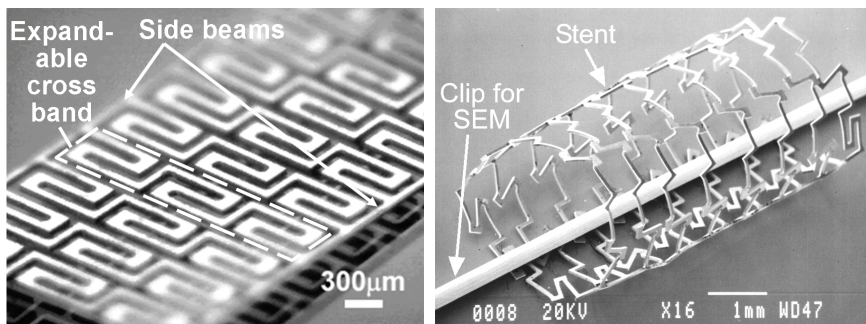
Fig. 16. (a: left) A 7-mm-long planar stent sample as cut by μEDM from 50-μm-thick stainless-steel foil; (b: right) an expanded state of the planar structure with diameter of 2.65 mm (Takahata & Gianchandani, 2004a) © 2004 IEEE.

4.3 Antenna stents

Following stent implantation, re-narrowing (restenosis) of the artery often occurs. To determine the status, patients are required to have an X-ray angiograph, potentially multiple times. Since this is an invasive procedure involving insertion of a catheter to inject contrast dye, it cannot be performed frequently. Wireless monitoring of cardiac parameters such as blood pressure and flow can provide advance notice of restenosis. Toward this end, the planar approach for stent fabrication described above was leveraged to develop a method that automatically transforms the electrical characteristics of a stent during balloon angioplasty, allowing the stent to be a helical-shaped antenna (stentenna) (Takahata et al., 2006). The planar design of the stent enables the use of lithography-based micromachining techniques for direct fabrication of sensors on the stent as well as the integration of separately fabricated microsensor chips. The planar device has a series of involute cross-

bands similar to those used in the mechanical stent described in Subsection 4.2, but designed to form dual inductors when the device is expanded to a cylindrical shape. Two micromachined capacitive pressure sensor chips were bonded to the planar stent structure and connected across the common line and the inductors, implementing a dual inductor-capacitor (*L-C*) tank configuration (Fig. 17a). In other words, integration of dielectric material (used to establish the sensing capacitor) with µEDM structures was implemented by a hybrid method, i.e., individual component assembly and packaging in this fabrication. The resonant frequency of the tank, which depends on local pressure or flow rate, was wirelessly interrogated through an external antenna magnetically coupled to the stentenna. The device was coated with Parylene-C™ for electrical protection while granting it biocompatibility. The stentenna was deployed inside a mock artery using a standard angioplasty balloon (Fig. 17a), resulting in a helical shape with inductance of ~110 nH. Wireless tests in a fluidic set-up showed that the device exhibited a frequency response of 9-31 KHz per mL/min. in the flow range over 370 mL/min (Fig. 17b).
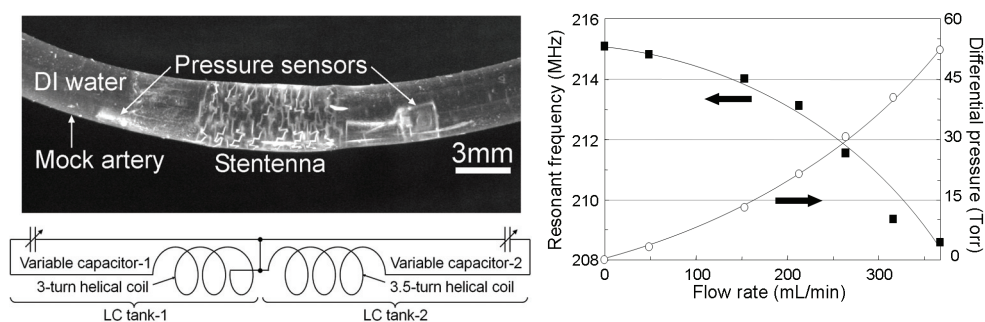
Fig. 17. (a: left) A deployed stentenna with pressure sensors and an equivalent electrical model of the deployed device; (b: right) measured resonant frequency of the stentenna as a function of flow rate (Takahata et al., 2006) © 2006 IEEE.

4.4 Microactuator-integrated scanning Kelvin probe

Kelvin probes are used to measure the contact potential difference (CPD) between materials, which cannot be measured directly using a voltmeter. One of the major applications is the characterization of solid-state devices. A probe is placed above the surface of a sample in close proximity, and an AC current is generated by dithering the gap where a CPD-induced charge is built up. The bias voltage that nulls the current indirectly determines the CPD. The micromachined probe developed by µEDM includes an actuator that provides the axial dither motion and a lead transfer beam for the probe (Fig. 18a). An electrothermal bent-beam actuator (Que et al., 2001) provides the dither motion with amplitude in the 10-µm range with drive voltages of a few volts. An isolation plug mechanically couples the probe to the actuator while electrically and thermally decoupling them from each other. A large width of isolation was desired to minimize the capacitive feedthrough of the drive signal as well as the thermal noise from the actuator. Monolithic integration of dielectric components (isolation plug) was achieved by the modified µEDM process depicted in Fig. 18b, which used a commercially available amorphous-metal foil (MetGlas 2826MB) as the conductive

material for device fabrication. The fabricated device was used for non-contact sensing of the pH of liquid inside microfluidic channels. The developed fabrication approach can potentially be applied to other devices that require mechanical structures and electrical circuits to be integrated in a monolithic manner. Further development may enable high-throughput production using batch-mode μEDM.
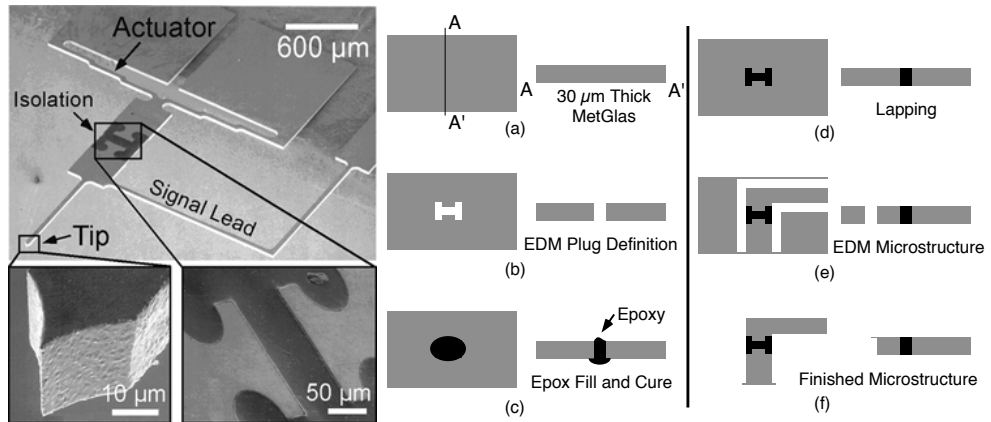
Fig. 18. (a: left) A fabricated Kelvin-probe device bonded to a glass substrate; (b: right) μEDM-based fabrication process for the device (Chu et al., 2005) © 2005 IEEE.

4.5 Intraluminal cuff for electromagnetic flow sensing

The planar-to-cylindrical reshaping technique used in stent fabrication has been applied to the development of an intraluminal ring cuff for electromagnetic (EM) sensing of flow (Takahata & Gianchandani, 2004b). EM detection offers several attractive features, such as a direct and linear relationship between output and flow, less dependence on cross-sectional flow profile, and mechanical robustness as there are no moving parts used (Yoon et al., 2000). EM flow sensors typically have two electrodes located on inner walls of the fluid channel. In the presence of a magnetic field, a voltage proportional to the flow velocity develops between the electrodes. The planar design of the ring cuff consists of a pair of meander bands comprising 50-μm-wide beams, electrode plates, and two dielectric links that mechanically tie the bands but electrically insulate them from each other (Fig. 19a). This pattern was created by μEDM in 50-μm-thick stainless-steel foil, and then all the surfaces except the front-side planes of the electrodes were coated with an insulating layer. (Without this layer, spatial averaging would reduce the voltage.) The dielectric links (of epoxy in this case) were created by a fabrication process similar to that used for the isolation plug in the Kelvin-probe device shown in Fig. 18b. The planar structure was mounted on a deflated angioplasty balloon so that one of the bands was located above the balloon and the other was below it, which allowed the structure to assume a ring shape when the balloon was inflated. In a wired set-up (Fig. 19b), the device, expanded inside a 3-mm i.d. silicon tube, showed a response that linearly and symmetrically increased or decreased depending on the orientation of an externally applied magnetic field (Fig. 19c). Signal reading for this device was also extended to a wireless implementation using the stentenna (Takahata et al., 2006).
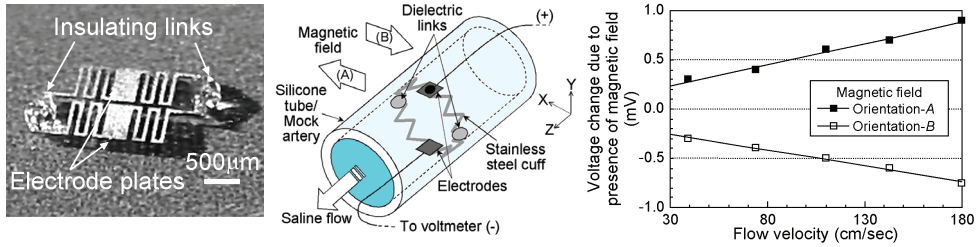
Fig. 19. (a: left) A stainless-steel cuff in the planar form; (b: middle) a fluidic measurement set-up for the expanded device; (c: right) measured responses of the device with opposing magnetic fields shown in the set-up. Figures a, b, and c are, respectively, reprints of Figs. 4, 10, and 11 in (Takahata & Gianchandani, 2004b), reprinted with the permission of the Transducer Research Foundation.

4.6 Cavity/Diaphragm-less capacitive pressure sensor

Micromachined capacitive pressure sensors typically use an elastic diaphragm with fixed edges and a sealed cavity between the diaphragm and the substrate below. Since this
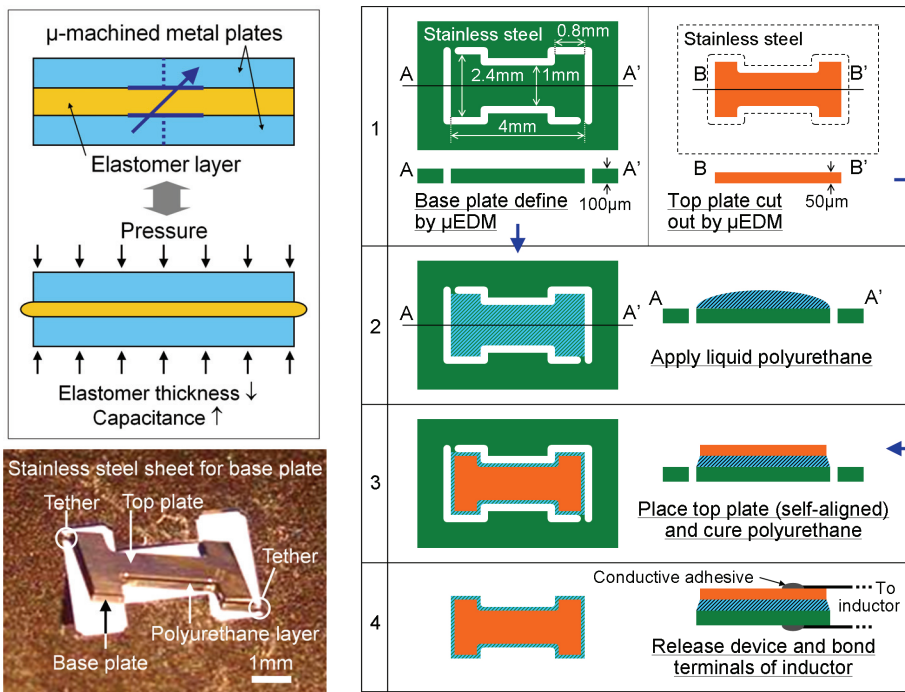
Fig. 20. (a: upper left) Cross-sectional view of the capacitive pressure sensor; (b: right) fabrication process flow; (c: lower left) a fabricated device being released from the original stainless-steel foil. Figures a, b, and c are, respectively, reprints of Figs. 1, 3, and 5a in (Takahata & Gianchandani, 2008b), reprinted with the permission of the Transducer Research Foundation.

configuration relies on the deflection of a relatively thin diaphragm against a sealed cavity, there is concern about the robustness of the diaphragm and leaks in the cavity seal in some applications. This issue has been addressed by devising a configuration consisting of two micromachined metal plates with an intermediate polymer layer, eliminating the need for diaphragms and cavities from the sensor structure (Takahata & Gianchandani, 2008a). Use of polymeric material soft enough to deform in a target pressure range allowed the thickness of the polymer, or capacitance of the parallel plate capacitor, to be dependent on the hydraulic pressure surrounding the device (Fig. 20a). The devices were constructed using micromachined stainless-steel electrodes defined by µEDM and a liquid-phase polyurethane that was applied and solidified between the electrodes. Figures 20b and 20c show the developed fabrication process and a fabricated device after Step 3 of the process, respectively. The integration of the liquid polyurethane with the electrode plates was achieved using a self-aligned assembly method. Pressure monitoring was demonstrated by measuring frequency shifts in the *L-C* tank, which was fabricated by winding a copper coil on the sensor and bonding the terminals to the electrodes. Wireless operation in liquid ambient was also demonstrated (Takahata & Gianchandani, 2008b). The bulk-metal-based cavity/diaphragm-less design of the device makes it suitable for high-pressure environments. In addition, the material combination potentially permits direct use of the device in corrosive or biological environments. These features, enabled through µEDM-based fabrication, may contribute to reducing packaging requirements for the device in selected applications.

5. Conclusion

This chapter presented conventional and advanced µEDM technologies and discussed their application to MEMS and microdevices enabled by bulk-metal micromachining. Due to its exceptional features and versatility, this technique has great potential for making broad contributions to manufacturing not only mechanical components but also the devices and systems that equip electromechanical functionalities. µEDM provides unique opportunities for R&D and manufacturing of such products, as it allows leveraging of non-traditional, high-performance engineering materials with various features such as plasticity, robustness, chemical inertness, and biocompatibility that cannot be achieved through conventional MEMS fabrication processes and their compatible materials. This ability also promotes proper choice of materials that are compatible with particular environments for MEMS fabrication, potentially allowing circumvention of constraints and problems associated with packaging and broadening application opportunities for MEMS. The batch-mode machining approach enabled by the use of lithographically formed electrodes was demonstrated to be a promising means of addressing the essential drawbacks of the traditional serial technique in terms of throughput and tolerance loss, as well as making µEDM compatible with conventional MEMS processes based on lithography techniques. In addition, the study on M[3]EDM revealed that MEMS can be utilized as means to push the limits of µEDM and advance the batch-machining technique. These new aspects of µEDM suggest that advancing the technique can be an effective way to address the constraints in the range of bulk materials available for MEMS design and manufacturing. The continued development of these new technologies will enable further breakthroughs and innovations in machining systems and MEMS manufacturing.

6. Acknowledgment

7. References

Alla Chaitanya, C.R. & Takahata, K. (2008a). Micro-electro-discharge machining by MEMS actuators with planar electrodes microfabricated on the work surfaces. *Technical Digest of the 21st IEEE International Conference on Micro Electro Mechanical Systems*, pp. 375-378.

Alla Chaitanya, C.R. & Takahata, K. (2008b). M3EDM: MEMS-enabled micro-electro-discharge machining. *Journal of Micromechanics and Microengineering*, Vol. 18, 105009 (7pp).

Alla Chaitanya, C.R. & Takahata, K. (2009). MEMS-based batch-mode micro-electro-discharge machining using microelectrode arrays actuated by hydrodynamic force. *Technical Digest of the 22nd IEEE International Conference on Micro Electro Mechanical Systems*, pp. 705-708.

Allen, D.M. & Lecheheb, A. (1996). Micro electro-discharge machining of ink jet nozzles: Optimum selection of material and machining parameters. *Journal of Materials Processing Technology*, Vol. 58, pp. 53-66.

Chu, L.L.; Takahata, K..; Selvaganapathy, P.; Gianchandani, Y.B. & Shohet, J.L. (2005). A micromachined Kelvin probe with integrated actuator for microfluidic and solid-state applications. *IEEE/ASME Journal of Microelectromechanical. Systems*, Vol. 14, No. 4, pp. 691-698.

Dhanik, S. & Joshi, S.S. (2005). Modeling of a single resistance capacitance pulse discharge in micro-electro discharge machining. *Journal of Manufacturing Science and Engineering*, Vol. 127, Iss. 4, pp. 759-767.

Ehrfeld, W.; Lehr, H.; Michel F. & Wolf, A. (1996). Micro electro discharge machining as a technology in micromachining. *Proceedings of SPIE*, Vol. 2879, pp. 332-337.

Fischer, K.; Chaudhuri, B.; McNamara, S.; Guckel, H.; Gianchandani, Y. & Novotny, D. (2001). A latching, bistable optical fiber switch combining LIGA technology with micromachined permanent magnets. *Technical Digest of the IEEE International Conference on Solid-State Sensors and Actuators*, pp. 1340–1343.

Grimes, C.A.; Jain, M.K.; Singh, R.S.; Cai, Q.; Mason, A.; Takahata, K. & Gianchandani, Y. (2001). Magnetoelastic microsensors for environmental monitoring. *Proceeding of the 14th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 278-281.

Guckel, H. (1998). High-aspect-ratio micromachining via deep x-ray lithography. *Proceedings of the IEEE*, Vol. 86, Iss. 8, pp. 1586-1593.

Hana, F.; Wachi, S. & Kunieda, M. (2004). Improvement of machining characteristics of micro-EDM using transistor type isopulse generator and servo feed control. *Precision Engineering*, Vol. 28, pp. 378-385.

Heeren, P.; Reynaerts, D.; Van Brussel, H.; Beuret, C.; Larsson, O. & Bertholds, A. (1997). Microstructuring of silicon by electro-discharge machining (EDM) - part II: Applications. *Sensors and Actuators A*, Vol. 61, Iss. 1-3, pp. 379-386.

Hiraishi, M.; Masaki, T. & Muro, M. (1999). High-speed micro EDM of micro nozzles. *Proceeding of the Annual Meeting of the Japanese Society of Electrical Machining Engineers*, pp. 45-48.

Ho, K.H.; Newman, S.T.; Rahimifard, S. & Allen, R.D. (2004). State of the art in wire electrical discharge machining (WEDM). *International Journal of Machine Tools and Manufacture*, Vol. 44, Iss. 12-13, pp. 1247-1259.

Honma, Y.; Takahashi, K. & Muro, M. (1999). Micro-machining of magnetic metal film using electro-discharge technique. *Advances in Information Storage Systems*, Vol. 10, pp. 383-399.

Kathuria, Y.P. (1998). Laser microprocessing of stent for medical therapy. *Proceedings of the IEEE International Symposium on Micromechatronics and Human Science*, pp. 111-114.

Laermer, F. & Urban, A. (2005). Milestones in deep reactive ion etching. *Digest of technical Papers, the 13th International Conference on Sold-State Sensors, Actuators and Microsystems*, Vol. 2, pp. 1118-1121.

Masaki, T.; Kawata, K. & Masuzawa, T. (1990). Micro electro-discharge machining and its applications. *Proceedings of the IEEE International Workshop on Micro Electro Mechanical Systems*, pp. 21-26.

Masuzawa, T.; Fujino, M.; Kobayashi, K.; Suzuki, T. & Kinoshita, N. (1985). Wire electro-discharge grinding for micro-machining. *Annals of the CIRP*, Vol. 34, pp. 431-434.

Masuzawa, T. & Sata, T. (1971). The occurring mechanism of the continuous arc in micro-energy EDM by RC circuit. *Journal of the Japan Society of Electrical Machining Engineers*, Vol. 5, No. 9, pp. 35–52.

Peirs, J.; Reynaerts, D.; Verplaetsen, F.; Poesen, M. & Renier, P. (2002). A microturbine made by micro-electro-discharge machining. *Proceedings of the 16th European Conference on Solid-State Transducers*, pp. 790-793.

Que, L.; Park, J.S. & Gianchandani, Y.B. (2001). Bent-beam electro-thermal actuators-I: Single beam and cascaded devices. *IEEE/ASME Journal of Microelectromechanical. Systems*, Vol. 10, No. 2, pp. 247–254.

Reynaerts, D.; Heeren, P. & Van Brussel, H. (1997). Microstructuring of silicon by electro-discharge machining (EDM) - part I: Theory. *Sensors and Actuators A*, Vol. 60, Iss. 1-3, pp. 212-218.

Reynaerts, D.; Meeusen, W.; Song, X.; Van Brussel, H.; Reyntjens, S.; De Bruyker, D. & Puers, R. (2000). Integrating electro-discharge machining and photolithography: Work in progress. *Journal of Micromechanics and Microengineering*, Vol. 10, pp. 189-195.

Richardson, M.T.; Gianchandani, Y.B. & Skala, D.S. (2006). A parametric study of dimensional tolerance and hydrodynamic debris removal in micro-electro-discharge machining. *Technical Digest of the 19th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 314-317.

Sato, K.; Shikida, M.; Matsushima, Y.; Yamashiro, T.; Asaumi, K.; Iriye, Y. & Yamamoto, M. (1998). Characterization of orientation dependent etching properties of single crystal silicon: Effects of KOH concentration. *Sensors and Actuators A*, Vol. 64, No. 1, pp. 87-93.

Sun, X.; Masuzawa, T. & Fujino, M. (1996). Micro ultrasonic machining and its applications in MEMS. *Sensors and Actuators A*, Vol. 57, Iss. 2, pp. 159-164.

Takahata, K.; Aoki, S. & Sato, T. (1997). Fine surface finishing method for 3-dimensional micro structures. *IEICE Transactions on Electronics*, Vol. E80-C, No. 2, pp. 291-296.

Takahata, K. & Gianchandani, Y.B. (2001). Batch mode micro-EDM for high-density and high-throughput micromachining. *Proceedings of the 14th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 72-75.

Takahata, K. & Gianchandani, Y.B. (2002). Batch mode micro-electro-discharge machining. *IEEE/ASME Journal of Microelectromechanical Systems*, Vol. 11, No. 2, pp. 102-110.

Takahata, K. & Gianchandani, Y.B. (2004a). A planar approach for manufacturing cardiac stents: Design, fabrication, and mechanical evaluation. *IEEE/ASME Journal of Microelectromechanical. Systems*, Vol. 13, No. 6, pp. 933-939.

Takahata, K. & Gianchandani, Y.B. (2004b). A micromachined stainless steel cuff for electromagnetic measurement of flow in blood vessels. *Proceeding of Solid-State Sensor, Actuator and Microsystems Workshop*, pp. 290-293.

Takahata, K. & Gianchandani, Y.B. (2007). Bulk-metal-based MEMS fabricated by micro-electro-discharge machining. *Proceedings of the 20th IEEE Canadian Conference on Electrical and Computer Engineering*, pp. 1-4.

Takahata, K. & Gianchandani, Y.B. (2008a). A micromachined capacitive pressure sensor using a cavity-less structure with bulk-metal/elastomer layers and its wireless telemetry application. *Sensors*, Vol. 8, pp. 2317-2330.

Takahata, K. & Gianchandani, Y.B. (2008b). A cavity-less micromachined capacitive pressure sensor for wireless operation in liquid ambient. *Proceedings of the Solid-State Sensor, Actuator and Microsystems Workshop*, pp. 300-303.

Takahata, K.; Gianchandani, Y.B. & Wise, K.D. (2006). Micromachined antenna stents and cuffs for monitoring intraluminal pressure and flow. *IEEE/ASME Journal of Microelectromechanical. Systems*, Vol. 15, No. 5, pp. 1289-1298.

Takahata, K. & Masaki, T. (1999). High precision machining of micro V-grooves. *Proceedings of the Japan Society for Precision Engineering Kansai-Region Annual Meeting*, pp. 35-36.

Takahata, K.; Shibaike, N. & Guckel, H. (1999). A novel micro electro-discharge machining method using electrodes fabricated by the LIGA process. *Proceedings of the 12th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 238-243.

Takahata, K.; Shibaike, N. & Guckel, H. (2000). High-aspect-ratio WC-Co microstructure produced by the combination of LIGA and micro-EDM. *Microsystem Technologies*, Vol. 6, No. 5, pp. 175-178.

Takeuchi, H.; Nakamura, K.; Shimizu, N. & Shibaike, N. (2000). Optimization of mechanical interface for a practical micro-reducer. *Proceedings of the 13th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 170-175.

Takeda, M.; Namura, K.; Nakamura, K.; Shibaike, N.; Haga, T. & Takada, H. (2000). Development of chain-type micromachine for inspection of outer tube surfaces (basic performance of the 1st prototype). *Proceedings of the 13th IEEE International Conference on Micro Electro Mechanical Systems*, pp. 805-810.

Tathireddy, P.; Rakwal, D.; Bamberg, E. & Solzbacher, F. (2009). Fabrication of 3-dimensional silicon microelectrode arrays using micro electro discharge machining for neural applications. *Technical Digest of the 15th IEEE International Conference on Solid-State Sensors, Actuators and Microsystems*, pp. 1206-1209.

Wada, T. & Masaki, T. (2005) Machining of micro PCD tool using micro EDM process and machining example. *Journal of the Japan Society for Abrasive Technology*, Vol. 49, No. 10, pp. 546-549.

Yoon, H.J.; Kim, S.Y.; Lee, S.W. & Yang, S.S. (2000). Fabrication of a micro electromagnetic flow sensor for micro flow rate measurement. *Proceedings of SPIE*, Vol. 3990, pp. 264-271.

Mechanical Properties of MEMS Materials

Zdravko Stanimirović and Ivanka Stanimirović
IRITEL A.D.
Republic of Serbia

1. Introduction

The performance of micro electronic and mechanical systems (MEMS) strongly depends on the mechanical properties of materials used. The evaluation of the mechanical properties of MEMS materials is indispensable for designing MEMS devices. Accurate values of mechanical properties (elastic properties, internal stress, strength, fatigue) are necessary for obtaining the optimum performances. For an example, elastic properties are necessary in prediction of the amount of deflection from an applied force and material strength sets device operational limits. Also, in view of reliability and life time requirements, mechanical characterization of MEMS materials becomes increasingly important. Small size of MEMS devices often leads to their usage in harsh environments, and good knowledge of mechanical properties may lead to elimination of some of the mechanical failure modes through proper material selection, design, fabrication and packaging processes. As the interest in MEMS grows, the demand for applicable data increases. Reliability, accuracy and repeatability of evaluation methods also became an issue. However, MEMS use materials such as silicon and many other thin films that are not fully characterized regarding their mechanical properties because they had not previously been considered as mechanical materials. The properties of thin films have so far been evaluated mostly to satisfy demands in semiconductor device research, but evaluations were mainly focused on the electrical properties, while investigations of mechanical properties were limited mainly to internal stresses. For that reason, the bulk properties were adopted whenever mechanical properties were needed, but with the growing application of thin films in various mechanical structures grew the need for better understanding of their mechanical and electromechanical properties. Therefore mechanical properties of thin films used in MEMS need to be accurately evaluated – they should be measured at the same scale as micro- and nano-devices since they differ from bulk material properties. Thin-film and bulk materials usually have different compositions, phase and microstructure and the formation process for thin films must be taken into account (deposition, thermal treatment, implantation and oxidation). Mechanical processing as the processing method for most bulk structures is in case of thin films substituted with photolithography and etching. Also, bulk and thin film have different surface finishing of processed structures. When size effect is concerned, one must have in mind that the ratio of surface area to the volume increases as dimensions of a device decrease. The dimensions of structures in MEMS devices range from sub-micrometers to millimetres and therefore the size effect in thin films is more sensitive than in bulk materials. Many measurement methods have been developed for evaluation of

mechanical properties of thin films and various values have been measured (Ammaleh, 2003; Dual, 2004; Yi, 1999). Reported variations in measured values were large requiring extensive research in order to evaluate repeatability, accuracy and data reliability of various measurement methods for mechanical properties of MEMS materials. Therefore, development of international standards on MEMS materials and their properties measurement methods is one of the primary tasks when MEMS technology is in question. For that reason, this chapter intends to give an overview of basic test methods and mechanical properties of MEMS materials. Definitions of mechanical properties of interest are presented along with current test methods for MEMS materials. Also, a summary of mechanical properties of various MEMS materials is given. Measured material data for MEMS structural materials is obtained from the literature. Finally, the brief overview of the topic is presented in the last section, pointing out the necessity of standardization of testing procedures that would accelerate advances in MEMS technology.

2. Mechanical properties

MEMS devices use materials such as silicon and many other thin films. These materials had not previously been considered mechanical materials and for that reason are not fully characterized regarding their mechanical properties. The evaluation of the mechanical properties of electrical materials forming MEMS devices is needed to provide the engineering base for full exploitation of the MEMS technology. It is essential both from the aspect of MEMS device performances, as well as from the reliability aspect. Mechanical properties of interest fall into three general categories: elastic, inelastic, and strength. In order to predict the amount of deflection from the applied force, or vice versa, the elastic properties of MEMS materials must be known. Inelastic material properties are important for ductile materials, when deformed structure does not return to its initial state. When defining operational limits of MEMS device, the strength of the material must be known. The key factor in manufacturing reliable MEMS devices is good understanding of the relation between the material properties and its processing. When studying material properties, measured values should be independent of test method and the size of the specimen. However, when MEMS devices are in question, the size of the specimen may affect the measurements. For that reason an extensive process should be initiated in defining test methods with adequate sensibility and repeatability that would provide accurate values of mechanical properties.

2.1 Elastic properties

Elastic properties are directly related to the device performance. Young's modulus and Poisons ratio are basic elastic properties that govern the mechanical behavior. Since two independent mechanical properties are necessary for full definition of mechanical properties of MEMS materials, their properties can be accurately determined by measuring Young's modulus and Poisson's ratio. Young's modulus (E) is a measure of a material stiffness. It is the slope of the linear part of stress-strain (ε-σ) curve of a material. Poisson's ratio is a measure of lateral expansion or contraction of a material when subjected to an axial stress within the elastic region. Load-deflection technique enables measuring E together with σ. The concept of this technique is shown in figure 1 using a circular membrane. The load-deflection technique is easy to apply because the membrane is flat without load enabling

easy load-deflection relationship measurement. The deflection of the membrane center (d) is measured with the applied pressure (P) across the membrane. Then, the pressure-deflection behavior of a circular membrane (Tsuchiya, 2008) is expressed by

$$P = \frac{4\sigma_0 t}{a^2}d + \frac{8Et}{3(1-v)a^4}d^3 , \tag{1}$$

where P is the applied pressure, d is the center deflection, a, t, E, σ_0 and v are the radius, thickness, Young's modulus and Poisons ratio of the circular membrane, respectively. As the equation shows, the range of Poison's ratio of materials is not wide and rough estimation of the ratio is acceptable using the bulk properties.
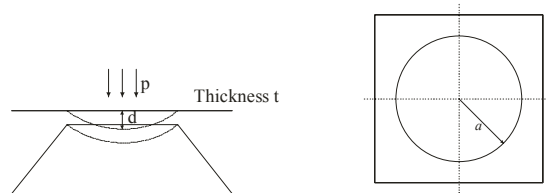
Fig. 1. The load-deflection technique for simultaneous E and σ measurement

2.2 Internal stress
Internal stress (σ), the strain generated in thin films on thick substrates, causes the deformation of the microstructure and occasionally destruction of the structure. It has two sources:
-   thermal mismatch between a substrate and a thin film – extrinsic stress,
-   microscopic structural change of a thin film (caused by chemical reactions, ion bombardment, absorption, adsorption etc.) – intrinsic stress.
In case of thin film compression the compressive stress is in question. Compressive stress is expressed as a negative value and it may cause buckling. In case of thin film expansion the tensile stress is in question. Tensile stress is expressed as a positive value and if excessive may lead to fracture of structures. According to Hooke's law, for isotropic materials under biaxial stress (such as thin films on substrates), internal stress is described by

$$\sigma = \varepsilon E/(1-v) , \tag{2}$$

where ε, E and v are the strain, Young's modulus and Poisson's ratio of the thin film, respectively.
As a micro fabricated test for strain measurement the beam buckling method is often used. In order to measure ε of thin films the doubly supported beam shown in figure 2 is loaded by the internal stress. The preparation of pattern with incrementally increasing size enables determination of the critical length of the beam which causes buckling.
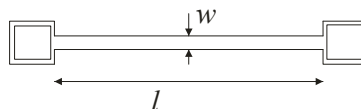
Fig. 2. Doubly supported beam structure

The strain deduced from the buckling length of the beam (Tabata, 2006) is given as:

$$\varepsilon = \frac{\pi^2}{3}\left(\frac{t}{l_c}\right)^2 , \tag{3}$$

where ε, t and l_c are the strain, thickness of the thin film and the buckling length, respectively. In this case, the internal stress is assumed to be uniform along the thickness direction. In case of the stress distribution along the thickness direction, variation of ε may cause vertical deflection of the cantilever beam.

2.3 Strength
The strength of a material determines how much force can be applied to a MEMS device. It needs to be evaluated in order to assure reliability of MEMS devices. Strength depends on the geometry, loading conditions as well as on material properties. As the useful measure for brittle materials, the fracture strength is defined as the normal stress at the beginning of fracture. The flexural strength is a measure of the ultimate strength of a specified beam in bending and it is related to specimen's size and shape. For inelastic materials, the yield strength is defined as a specific limiting deviation from initial linearity. The tensile strength is defined as a maximum stress the material can withstand before complete failure while the compressive strength is usually related to brittle materials.

2.4 Fatigue
MEMS devices are often exposed to cyclic or constant stress for a long time during operation. Such operational conditions may induce fatigue. Fatigue may be observed as change in elastic constants and plastic deformation leading to sensitivity changes and offset drift in MEMS devices. It may also be observed as the strength decrease that may lead to fracture and consequentially failure of the device. Fatigue behavior of a MEMS device also depends on its size, surface effects, effect of the environment such as humidity and temperature, resonant frequencies etc. In order to realize highly reliable MEMS device a detailed analysis of the fatigue behavior must be performed using accelerated life test method as well as life prediction method.

3. Testing methods

Minimum features in MEMS are usually of the order of 1µm. Measuring mechanical properties of small MEMS specimens is difficult from the aspects of reliability, repeatability and accuracy of measurements. In order to measure mechanical properties of the MEMS device a specimen must be obtained and mounted. Since the microdevices are produced using deposition and etching processes a specimen must be produced by the same process used in device production. The following step is dimension measurement. The thicknesses of layers are controlled and measured by the manufacturer and lengths are sufficiently large to be measured by an optical microscope with required accuracy. However, the width of the specimen may cause the problem due to its small dimensions as well as imperfect definition of cross section that may cause uncertainty in the area. Therefore, possible measurement techniques include optical or scanning electron microscopy, interferometry, mechanical or optical profilometry. The next step in measuring mechanical properties of MEMS is the

application of force/displacement resulting in deformation. This step is followed by force, displacement or strain measurements. Force and displacement measurements are based on tensile and bending tests or on usage of commercially available force and displacement transducers. When strain measurements are in question, it is preferable to measure strain directly on tensile specimens. It enables determination of the entire strain-stress curve from which the properties are obtained. The strain measurement technique known as interferometric strain/displacement gage is usually used apart from variety of other techniques that have not yet been applied to extensive studies of mechanical properties of MEMS. However, in most cases when MEMS materials are in question, direct methods for mechanical properties determination are not suitable. Instead, inverse methods are being used: a model is constructed of the test structure. After the force application and displacement measurements, elastic, inelastic or strength properties can be extracted from the model. Nevertheless, the variations in measured properties are large for both types of testing methods: direct and inverse. The source of variations is not established since there are too many differences among the properties measured by different methods. Obviously, the development of international standards for measuring the mechanical properties of MEMS materials will result in more accurate properties and reliable measurements.

3.1 Tensile testing methods

When tensile testing methods are concerned there are three arrangements that can be used. The first of them is specimen in a supporting frame. The tensile specimen is patterned onto the wafer surface and the gage section is exposed by etching the window in the back of the wafer. The specimen suspended across a rectangular frame enables convenient handling and testing. An example of specimen in a supporting frame is shown in figure 3.

Fig. 3. Schematic of a silicon carbide tensile specimen in a silicon support frame

The second arrangement used in tensile testing is a specimen fixed at one end. At one end the test specimens fixed to the die while the other is connected to the test system. There is a variety of ways a specimen fixed at one end may be connected to a test system. A free end may be gripped by the electrostatic probe, glued to the force/displacement transducer, connected to the test system by the pin in case of ring shaped grip end, etc. An example of specimen fixed at one end is shown in figure 4.
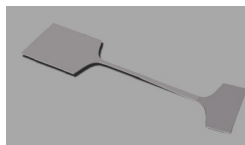
Fig. 4. Schematic of a tensile specimen fixed to the die at one end

The third arrangement used in tensile tests of MEMS materials is the freestanding specimen. This arrangement applies to small tensile specimens with submillimeter dimensions. The

geometry commonly used in these tests is shown in figure 5. Microspecimens have grip ends that can be fitted into inserts in the grips of the test machine.
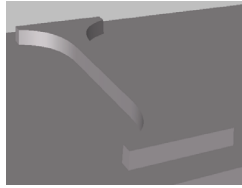
Fig. 5. Schematic of nickel free standing microspecimen on a silicon substrate

3.2 Bend tests

Similar to tensile testing methods, bend tests also use three arrangements. The first of them is out-of plane bending. Long, narrow and thin beams of the test material are being patterned on the substrate. The material under the cantilever beam is being etched away leaving the beam hanging freely over the edge. By applying the force as shown in figure 6 and measuring the force vs. deflection at the end or near the end of the beam, Young's modulus can be extracted.
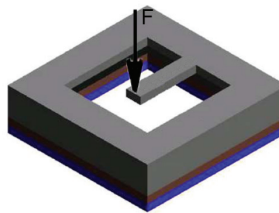
Fig. 6. Shematic of crystal silicon cantilever microbeam that can be used in out-of-plane bending test

The second arrangement used in bend tests is the beam with fixed ends – so called fixed-fixed beam. The schematic of the most usually used on-chip structure is shown in figure 7. Between the silicon substrate and polysilicon beam with clamped ends a voltage is applied pulling the beam down. The voltage that causes the beam to make contact is a measure of beam's stiffness.

The third arrangement used in bend testing of MEMS materials is in-plane bending (fig. 8). Test structure consisting of cantilever beams subjecting to in-plane bending may be used in fracture strain determination, crack growth and fracture toughness measurements, etc.
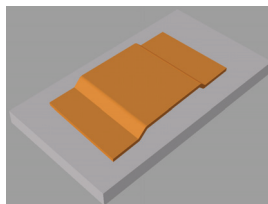
Fig. 7. Schematic of a polysilicon fixed-fixed beam on a silicon substrate
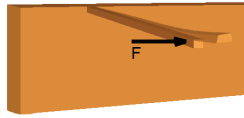
Fig. 8. Schematic of polysilicon cantilever beam subjected to in-plane bending

3.3 Resonant structure tests

Resonant structure tests are being used for determination of elastic properties of MEMS devices. Very small test structures used in these tests can be excited by capacitive comb drives which require only electrical contact making this approach suitable for on-chip testing. The most often used resonant structure concepts also include different in-plane resonant structures with a variety of easily modeled geometries as well as test structures based on arrays of cantilever beams fixed at one or both ends excited in different manners. As an illustration, in figure 9 a schematic of a in-plane resonant structure is shown.
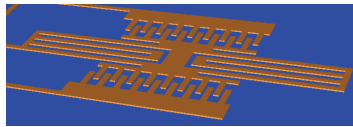
Fig. 9. Schematic of the in-plane resonant structure

3.4 Bulge testing

Bulge testing is also often called membrane testing. By etching the substrate material a thin membrane of test material is formed. The ideal architecture to achieve a direct tensile testing scheme involves a free standing membrane fixed at both ends (Espinosa, 2003) as shown in figure 10. When load is applied at the center of the membrane (usually using nanoindenter), a uniform stretch on the two halves of the thin membrane is achieved. In this manner the specimen's structural response is obtained as well as elastic behavior and residual stress state.
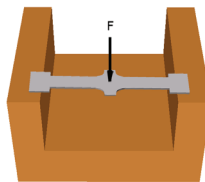
Fig. 10. Shematic of an Au membrane used in bulge testing

3.5 Indentation tests

In indentation tests a miniature and highly sensitive hardness tester (nanoindenter) is being used allowing force and displacement measurements. Penetration depths can be a few nanometers deep and automation permits multiple measurements and thus provides more reliable results. In such a manner Young's modulus and strength of various thin films can be obtained. As an illustration, a schematic of an indentation test is given in figure 11.

Fig. 11. Schematic of indentation test

3.6 Other tests

In order to measure forces in specimens the buckling test method can be used and if the specimen under pressure breaks the estimate of fracture strength can be obtained. This test method applied to test structures with different geometries and based on different MEMS materials can be used for determination of the Poisson's ratio, strain at fracture, residual strain in film, etc.

Another test method is the creep test. Creep tests are usually performed in cases when possibility of creep failure exists such as in thermally actuated MEMS devices resulting in a strain vs. time creep curve.

When torsion, one of important modes of deformation in some of MEMS devices, is concerned a few torsion tests have been developed enabling force and deflection measurements.

Fracture tests are of interest when brittle materials are in question. Fracture toughness is being measured using crack formation with a tip radius small relative to the specimen dimensions. Different positions and shapes of cracks are being used formed using different means such as etching, various types of indenters, etc.

When mechanical testing of MEMS materials is in question, standardization of test methods is a challenging task. A step forward in the direction of standardization may be implementation of "round robin" tests that should involve all relevant MEMS researchers in an effort to test common materials used in MEMS at their premises using the method of their choice. First such tests resulted in significant variation of results suggesting that further efforts should be made by involving more scientific resources.

4. Data

Polysilicon is the most frequently used MEMS material. In table 1 polysilicon mechanical properties data is given obtained by three types of tests: bulge, bend and tensile tests. Presented results show that polysilicon has Young's modulus mostly in the range between 160 and 180 GPa. Fracture strength depends on flaws in the material and performed tests do not necessarily lead to failure of the specimen. For that reason there are fewer entries for fracture strength and obtained results vary.

Mechanical properties of single-crystal silicon are given in table 2. Presented data is obtained using bending, tensile and indentation tests. The average values for the Young modulus ranged between 160 an 190 GPa.

In table 3 silicon-carbide mechanical properties data is presented. It is a promising MEMS material because of its superior properties (strength, stability, stiffness) and because of the current work on thin-film manufacturing processes few results are available obtained using bulge, indentation and bending tests.

Silicon nitride and silicon oxide mechanical properties data is presented in tables 4 and 5, respectively. Silicon nitride is used as an insulating layer in MEMS devices but it also has a potential as a structural material. On the other hand, silicon oxide because of its properties

(low stiffness and strength) although included in MEMS devices does not have a potential of becoming a MEMS structural material.

There are few reports regarding the mechanical properties of metal thin films. Table 6 lists measured values of mechanical properties of metal materials commonly used in MEMS devices: gold, copper, aluminum and titanium. Metal films are tested using tensile testing in a free-standing manner. Results for electroplated nickel and nickel-iron MEMS materials are given in table 7. Presented results are obtained using tensile testing methods. Electroplated nickel and nickel-iron MEMS are usually manufactured by LIGA process. The microstructure and electrical properties of electroplated nickel are highly dependent on electroplating conditions while the properties of nickel-iron alloy depend on its

Methods	Young's Modulus [GPa]	Fracture Strength [GPa]
Bulge test	160	-
Bulge test	190-240	-
Bulge test	151-162	-
Bulge test	162±4	-
Bending test	-	2.11-2.77
Bending test	170	-
Bending test	174±20	2.8±0.5
Bending test	135±10	-
Bending test	198	-
Bending test	-	3.2±0.3
Bending test	-	3.4±0.5
Tensile test	164-176	2.86-3.37
Tensile test	-	0.57-0.77
Tensile test	140	0.7
Tensile test	160-167	1.08-1.25
Tensile test	169±6	1.20±0.15
Tensile test	132	-
Tensile test	140±14	1.3±0.1
Tensile test	172±7	1.76
Tensile test	167	2.0-2.7
Tensile test	163	2.0-2.8
Tensile test	-	1.8-3.7
Tensile test	166±5	1.0±0.1
Tensile test	-	4.27±0.61
Tensile test	-	2.85±0.40
Tensile test	-	3.23±0.25
Tensile test	158±8	1.56±0.25
Tensile test	-	2.9±0.5
Fixed ends test	123	-
Fixed ends test	171-176	-
Fixed ends test	149±10	-
Fixed ends test	178±3	-

Table 1. Polysilicon mechanical properties data (Sharpe, 2001)

composition. Presented results show that these materials have high strength values (especially Ni-Fe) and therefore are suitable for application in actuators.

In table 8 diamond-like carbon mechanical properties data is presented. Diamond-like carbon is a MEMS material with excellent properties such as high stiffness and strength and low coefficient of friction. Presented results are obtained using three types of test methods: bending, buckling and tensile tests.

Methods	Young's Modulus [GPa]	Fracture Strength [GPa]
Bending test	177±18	2.0-4.3
Bending test	163	>3.4
Bending test	122±2	-
Bending test	173±13	-
Bending test	-	0.7-3.0
Bending test	165±20	2-8
Bending test	-	2-6
Bending test	169.9	0.5-17
Tensile test	147	0.26-0.82
Tensile test	125-180	1.3-2.1
Tensile test	142±9	1.73
Tensile test	-	0.59±0.02
Tensile test	169.2±3.5	0.6-1.2
Tensile test	164.9±4	-
Indentation test	60-200	-
Indentation test	168	-

Table 2. Single-crystal silicon mechanical properties data (Sharpe, 2001)

Methods	Young's Modulus [GPa]
Bulge test	394
Bulge test	88±10 - 242±30
Bulge test	331
Indentation test	395
Bending test	470±10

Table 3. Silicon-carbide mechanical properties data (Sharpe, 2001)

Methods	Young's Modulus [GPa]	Fracture Strength [GPa]
Resonant test	130 - 146±20%	-
Resonant test	192	-
Resonant test	194.25±1%	-
Bulge test	230 & 330	-
Bulge test	110 & 160	0.39-0.42
Bulge test	222±3	-
Indentation test	101-251	-
Indentation test	216±10	-

Table 4. Silicon-nitride mechanical properties data (Sharpe, 2001)

Methods	Young's Modulus [GPa]	Fracture Strength [GPa]
Indentation test	64	>0.6
Bending test	83	-
Tensile test	-	0.6-1.9

Table 5. Silicon-oxide mechanical properties data (Sharpe, 2001)

	E$_{bulk}$ [GPa]	Young's Modulus [GPa]	Yield Strength [GPa]	Ultimate Strength [GPa]
Gold	74	98±4	-	-
Gold	74	82	-	0.33-0.36
Copper	117	86-137	0.12-0.24	0.33-0.38
Aluminum	69	8-38	-	0.04-0.31
Aluminum	69	40	-	0.15
Titanium	110	96±12	-	0.95±0.15

Table 6. Metal films mechanical properties data (tensile test) (Sharpe, 2001; Tabata, 2006)

	Young's Modulus [GPa]	Yield Strength [GPa]	Ultimate Strength [GPa]
Ni	202	0.4	0.78
Ni	176±30	0.32±0.03	0.55
Ni	131-160	0.28-0.44	0.46-0.76
Ni	231±12	1.55±0.05	2.47±0.07
Ni	181±36	0.33±0.03	0.44±0.04
Ni	158±22	0.32±0.02	0.52±0.02
Ni	182±22	0.42±0.02	0.60±0.01
Ni	156±9	0.44±0.03	-
Ni	160±1	0.28	-
Ni	194	-	-
Ni-Fe	119	0.73	1.62
Ni-Fe	155	-	2.26
Ni-Fe	-	1.83-2.20	2.26-2.49

Table 7. Electroplated nickel and nickel-iron mechanical properties data (tensile test) (Sharpe, 2001; Tabata, 2006)

Methods	Young's Modulus [GPa]	Fracture Strength [GPa]
Bending test	600-1100	0.8-1.8
Buckling test	94-128	-
Tensile test	-	8.5±1.4

Table 8. Diamond-like carbon mechanical properties data (Sharpe, 2001)

5. Summary

The measurement of MEMS materials mechanical properties is crucial for the design and evaluation of MEMS devices. Even though a lot of research has been carried out to evaluate

the repeatability, accuracy and data reliability of various measurement methods for mechanical properties of MEMS materials, the manufacturing and testing technology for materials used in MEMS is not fully developed. In this chapter an overview of basic test methods and mechanical properties of MEMS materials is given along with definitions of mechanical properties of interest. Also, a summary of the mechanical properties of various MEMS materials is given. Variation of obtained results for common materials may be attributed to the lack of international standards on MEMS materials and their properties measurement methods. It must be pointed out that although MEMS is an area of technology of rapidly increasing economic importance with anticipated significant growth, the ability to develop viable MEMS is to a large degree constrained by the lack of international standards on MEMS materials and their properties measurement methods that would establish fundamentals of reliability evaluation, especially on MEMS material properties.

6. Acknowledgement

7. References

Allameh, S.M. (2003). An intorduction to mechanical-properties-related issues in MEMS structures. *Journal of materials science*, 38, (2003) 4115-4123, ISSN: 1573-4803

Dual, J.; Simons, G.; Villain, J.; Vollmann, J. & Weippert, C. (2004). Mechanical properties of MEMS structures, Proceedings of ICEM12, ISBN: 88-386-6273-8, Bari, Italy, August-September 2004, McGraw-Hill

Espinosa, H.D.; Prorok, B.C. & Fischer, M. (2003). A methodology for determining mechanical properties of freestanding thin films and MEMS materials. *Journal of the Mechanics and Physics of Solids*, 51, (2003) 47-67, ISSN: 0022-5096

Sharpe, W.N.Jr. (2001). Mechanical Properties of MEMS materials, In: *The MEMS Handbook*, Mohamed Gad-el-Hak, 3/1– 3/33, CRC Press, ISBN: 978-0849300776, USA

Tabata, O. & Tsuchiya, T. (2006). Material Properties: Measurement and Data, In: *MEMS, A Practical Guide of Design, Analysis, and Applications*, Jan Korvink, 53–92, Springer, ISBN: 978-3540211174

Tsuchiya, T. (2008). Evaluation of Mechanical Properties of MEMS Materials and Their Standardization. In: *Advanced Micro and Nanosystems*, Tabata, O. & Tsuchiya, T., 1-25, Wiley-VCH Verlag GmbH & Co. KgaA, ISBN: 978-3-527-31494-2, Weinheim

Yi, T. & Kim, C-J. (1999). Measurement of mechanical properties for MEMS materials. *Measurement Science and Technology*, 10, (1999) 706-716, ISSN: 1361-6501

Reliability of MEMS

Ivanka Stanimirović and Zdravko Stanimirović
IRITEL A.D.
Republic of Serbia

1. Introduction

Reliability is a key factor for successful commercialization of micro electronic and mechanical systems (MEMS). MEMS devices are becoming essential components of modern engineering systems and their reliability is of particular importance in applications where their failure can be catastrophic and devastating (surgical devices, implantable biosensors, navigation in aerospace, sensors in automotive industry, etc.). However, although MEMS devices are made of minute delicate components realized primarily using physical-chemical processes, the main reason for the lack of success in commercialization of MEMS cannot be attributed to the advance of micro technology but to packaging techniques used in production of MEMS devices. When MEMS packaging is in question, it is of the greatest importance that design and realization of MEMS device must include all levels of reliability issues from the onset of the project. For that reason, this chapter is intended to be a general overview focusing on mechanisms that cause failure of MEMS devices. An insight in reliability of MEMS packaging (types of MEMS packaging, material requirements and package reliability) is given. Also, the reliability of MEMS in view of materials, structural and process reliability and associated failure mechanisms is presented. As the closing subsection the brief summary of the topic will be presented with an emphasis on the importance of the further R&D work on MEMS reliability testing and development of industrial standard for assembly, packaging and testing.

2. Reliability in MEMS packaging

Although the most silicon-based MEMS are produced using the same microfabrication processes developed for integrated circuits (ICs), these two technologies are significantly different and MEMS are not evolved integrated circuits. There are several principal differences between silicon-based MEMS and integrated circuits (Hsu, 2006):

- Silicon–based MEMS are complex 3D structures while integrated circuits are primarily 2D structures.
- Many MEMS devices involve precision movement of solid components and fluids in sealed enclosures, and integrated circuits are stationary encapsulated electric circuits.
- While MEMS perform a great variety of specific functions of biological, chemical, electromechanical and optical nature, integrated circuits transmit electricity for specific electrical functions.

- MEMS as delicate moving or stationary components are interfaced with working media while IC dies are isolated from contacting media.
- MEMS are using silicon and silicon compounds plus variety of other industrial materials, while integrated circuits are limited to single crystal silicon and silicon compounds, ceramic and plastic.
- In MEMS there are many components to be assembled and in integrated circuits there are fewer components to be assembled.
- MEMS packaging technology is far from being developed while IC packaging techniques are relatively well developed.
- For MEMS there are no available industrial standards regarding design, materials selections, fabrication processes and assembly-packaging-testing while integrated circuits have available industrial standards in all these areas.
- Most MEMS are custom built and assembled on batch production lines in contrast to mass production of ICs.
- MEMS have limited sources of commercialization while integrated circuits are fully commercialized.

MEMS packaging is more complex than packaging of integrated circuits because of their complex structures and specific performances. MEMS packaging must provide support and protection to ICs, associate wire bonds and the printed circuit board (PCB) from mechanical or environmentally induced damages and protect elements that require interface with working media which can be environmentally hostile. The fact that many MEMS require non-standard packages is one of the reasons why they have not made their way to the market. There are three basic types of packages used in MEMS technology: ceramic, metal and plastic. Some of the features of these three types of packages are given in table 1.

2.1 Materials selection for MEMS packaging

When MEMS devices are in question, materials selection should be done carefully. Similar to IC packaging, most of the MEMS devices are diced from a wafer and mounted on a substrate inside a package and therefore a careful attention must be paid to die attachment materials selection. Die attach material should firmly bond die to the substrate eliminating any possibility of motion. Die movement may cause various problems especially in optoelectronic devices where alignment is important. Fracture toughness is very important for brittle attachment materials because it determines material resistance to fracture. Mismatch of the coefficient of thermal expansion (CTE) between die attach material, silicon and substrate may lead to undesirable stress. Another important factor in attachment materials selection is thermal conductivity because die attachment material conducts heat from the die to the substrate. Moisture adsorption is critical because it causes degradation of die attach bonding properties. In order to minimize stress induced to the die, organic materials (epoxies, silicones, polyamides) are often used as die attach materials. These low cost materials are also convenient because of the ease of rework. However, in unpassivated MEMS devices outgassing of organic material may cause contamination. Organic materials are usually not used for ceramic packages. Temperature needed to produce frit seal after die attachment may lead to the degradation of the adhesive. Inorganic materials are also being used as die attach materials. These materials exhibit excellent fatigue resistance and provide lowest levels of contaminant gasses, but due to the lack of plastic flow may cause mismatch between substrate and die.

Ceramic packaging	• Commonly used in MEMS packaging • Usually consist of a base and a header • Die attachment by solder or adhesives • Generally electrically insulating • Hermetic • The match between coefficient of linear thermal expansion between ceramic and Si is fairly good • High mechanical strength • Resistant to chemicals
Metal packaging	• Robust • Easy to assemble • Allow prototyping in small volumes with short turnaround periods • Hermetic when sealed
Plastic packaging	• Cost effective • Small weight • Allow moist absorption

Table 1. Main features of three basic types of packages used in MEMS technology

Substrates for MEMS packaging must meet different electrical, thermal, physical and chemical requirements. One of the most important factors is dielectric constant of the substrate whose high value may cause crosstalk between wires. Another important factor is CTE. In order to minimize the thermal-mechanical stress in the package that may cause cracks or errors (in piezo-resistive sensor elements), CTE values of the substrate, die and die attach material must be matched. Another substrate property that must be taken into consideration is the loss tangent. If the substrate with high loss tangent is used, performances of MEMS devices sensitive to the frequency of applied signals may be reduced significantly. That also may result in low quality factor (Q) which measures performances of MEMS devices. Thermal conductivity of the substrate is important from the aspect of the heat transfer and porosity. Also, porosity and purity of the substrate must be evaluated because of the possibility of moisture penetration through the substrate. Properties of most commonly used substrates for MEMS packaging are given in table 2. It should be pointed out that Low Temperature Cofired Ceramic (LTCC) being the multilayer substrate allows implementation of cavities and allows movement in Z direction.

2.2 MEMS package reliability

When MEMS package reliability is in question, basic issues that should be taken into consideration are issues related to reliability of die attachments, ceramic substrates and released MEMS structures.

Mechanical connection between the substrate and MEMS structure is provided by die attach materials. CTE mismatch between used materials induces stress on the MEMS structure that may lead to formation of cracks on silicon MEMS structure. Cracks can appear at the centre or at the corners of the die usually when hard adhesives are used as die attach materials. In that case, CTE mismatch stress is transferred to the die causing cracks. Die attach can also crack if soft adhesives are used because it acts as a strain buffer at the die-substrate interface.

Single Layer Substrates	Tensile Strength (MPa)	Elastic Modulus (GPa)	Flexural Strength (MPa)	Dielectric Strength (kV/mm)
BeO	230	345	250	0.78
Si	-	310-343	360	0.55
AlN	-	190	580	-
Al_2O_3 (96%)	127.4	310.3	317	0.33
Al_2O_3 (96%)	206.9	345	345	0.33
Steatite	55.2-69	90-103	110	7.9-15.7
Fosforite	55.2-69	90-103	124	7.9-11.8
Quartz	48.3	71.7	-	-

Single Layer Substrates	Dielectric Constant @ 1MHz	Thermal Conductivity (W/m°C)	CTE (ppm/°C)
BeO	6.7-8.9	150-300	6.3-7.5
Si	8.5-10	82-320	4.3-4.7
AlN	11.9	125-148	2.33
Al_2O_3 (96%)	4.5-10	15-33	4.3-7.4
Al_2O_3 (99%)	4.5-10	15-33	4.3-7.4
Steatite	5.5-7.5	2.1-2.5	8.6-10.5
Fosforite	6.2	2.1-4.2	11
Quartz	4.6	43	1.0-5.5
Multilayer Substrate			
LTCC	6-9	2-4	5-7

Table 2. Properties of commonly used substrates for MEMS packaging (Pecht, 1998)

When organic die attach materials are being used, outgassing becomes an issue. In that case vacuum packaging is recommended. It protects MEMS devices from damage and contamination. Besides outgassing, if the organic die attach material is being used, moist absorption may cause failure. In hermetically sealed packages moisture trapping may occur causing delamination.

When ceramic substrate reliability is in question, CTE mismatch between substrate and silicon die may induce stress on the die causing cracking or bending. This can be avoided by careful evaluation of material properties. Matching CTE values of the substrate and the die lead to elimination of this problem.

Another reliability issue when MEMS package reliability is in question is packaging of released MEMS structures. Since they are susceptible to contamination, excessive handling, mechanical shock and stiction caused by the presence of moisture, the wafer level vacuum packaging is recommended.

3. Reliability of MEMS

Variety of applications may lead to misconception that amount of different structural parts of MEMS devices is large. However, there are a number of basic parts that are being used: cantilever beams, membranes, hinges, etc. The most common generic MEMS elements are listed in table 3.

- Structural beams
  - rigid
  - flexible
  - one side clamped
  - two sides clamped
- Structural thin membranes
  - rigid
  - flexible
  - with holes
- Flat layers (usually adhered to substrate)
  - conductive
  - insulating
- Hinges
  - substrate hinge
  - scissors hinge
- Cavities
  - sealed
  - open
- Gears
  - teeth
  - hubs
- Tunnelling tips
- Reflective layers

Table 3. Generic MEMS elements (Merlijn van Spengen, 2003)

MEMS devices are usually batch fabricated using silicon wafers as the material and etching techniques to build components. Fabrication process is more complex than fabrication process of ICs because of mechanical parts and electromechanical parts that are being integrated with electronic parts on the same substrate. MEMS have more complex shapes, have moving parts and need more material strength. Mechanical parts need special attention throughout the production cycle: from material deposition to material removal. These parts may have complex shapes, may require material with special strength and may have moving parts. Therefore, deposited film must be thick enough to form the mechanical layer. Moving parts are released after etching away the SiO_2 layer. Common processing techniques include bulk micromachining, wafer-to wafer bonding, surface micromachining and high-aspect ratio micromachining. Many MEMS failure modes are introduced in the fabrication process. Also, many failure modes in operation are related to fabrication process. MEMS common failure modes are fracture, creep, stiction, electromigration, wear, degradation of dielectrics, delamination, contamination, pitting of contacting surfaces, electrostatic discharge (ESD), etc.

One of the most important failure modes is *stiction*. Due to small sizes of MEMS structures surface forces dominate all others. The most important surface forces in MEMS are electrostatic force, capillary force and molecular van der Waals force (Tadigadapa, 2001). They cause stiction between microscopic structures when their surfaces come into contact. It can affect even elements that are not powered. Illustration of this failure mode is given in figure 1.
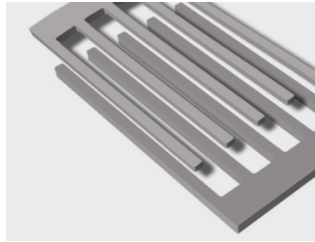
Fig. 1. Illustration of stiction failure mode

Creep is important issue for reliability of metal MEMS. High stresses and stress gradients introduce possibility of time-dependent mass transfer through glide and diffusion mechanisms. The creep is much more severe in MEMS structures than expected from macroscopically known behaviour. Macroscopically, creep is negligible. MEMS manufacturers should pay special attention when using metal as a structural material in MEMS where room temperature creep exists.

Most metals and alloys are degraded by material *fatigue* when subjected to a large repetitive mechanical stress. Cyclic loading of MEMS couples with other failure mechanisms associated with static loading, creep and environmental effects. Any process that results in an irreversible repositioning of atoms within a material can contribute to fatigue. Brittle materials like ceramics and silicon do not have a significant cyclic fatigue effect. Poly and possibly mono-crystalline silicon seem to suffer from a stress corrosion cracking mechanism (Muhlstein, 1997). In a not completely water free environment, small cracks propagate under tensile stress, due to hydrolysis of the native oxide layer (fig.2).
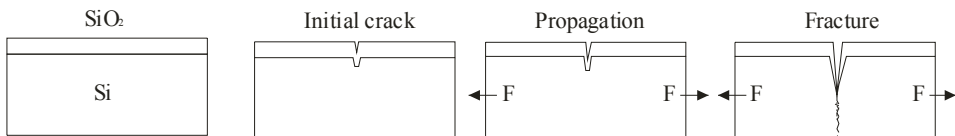
Fig. 2. Stress corrosion cracking failure mechanism

Friction and wear are of interest when sliding/rotating MEMS are in question. The wear mechanism in silicon is adhesive wear (Merlijn van Spengen, 2003). Rough contacting surfaces adhere to each other at their highest points. These points are broken and stay attached to other surfaces. Material is then transferred between surfaces and when asperities grow to a certain size they break off leaving worn surface and causing the accumulation of debris. Illustration of adhesive wear is shown in figure 3.
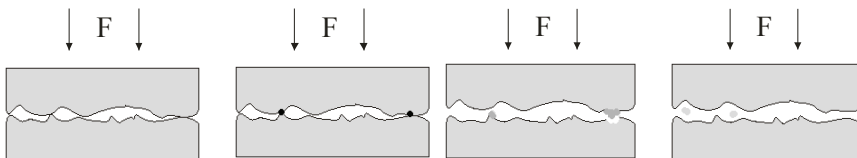
Fig. 3. Illustration of adhesive wear

Dielectric charging is an important issue for reliability of MEMS that contain dielectric layers. Parasitic charge accumulating in MEMS may alter actuation voltages and affect mechanical

behavior of the device. Also, a common problem is charging due to high field strengths required for actuation of electrostatically actuated MEMS.

Delamination is associated with multilayer films. High stress can be introduced by processing, thermal mismatch or epitaxial mismatch. The adhesion between layers depends strongly on their chemical and mechanical compatibility.

Electrostatic clamping of gears may prevent gears from moving due to presence of electrical charges at certain energy levels.

Particles have a damaging effect on devices where small gaps exist between bearing surfaces or elements with large potential difference. *Particulate contamination* is important when contaminating particles are internally generated or present in spite of a clean room environment.

Environmental effects may be important for design of variety of MEMS applications. *Environmental Attack* is of interest in case of valves, sensors and pumps where contacting fluids may be corrosive resulting in crack growth.

MEMS failure mechanisms are numerous and list of possible failure modes does not end here. Root causes of MEMS failure modes are different from the common causes on macroscopic level. Some of them are capillary forces, operational methods, mechanical and electrical instabilities. MEMS failure analysis techniques are similar to techniques that are being applied in failure analysis of ICs: optical microscopy, scanning laser microscopy, scanning electron microscopy, focused ion beam, atomic force microscopy, light emission, acoustic microscopy, acoustic emission, laser cutting, lift-off technique, etc. Failure models for MEMS are scarce because a failure model should be able to describe physic of failure and allow failure prediction. It is obvious that current knowledge of MEMS reliability is insufficient in comparison with the amount of MEMS devices that are already available.

4. Conclusion

An insight in reliability of MEMS packaging and the reliability of MEMS has been presented in this chapter. Since there is a common misconception that silicon MEMS devices and ICs are similar because they use the same microfabrication techniques, principal differences between silicon-based MEMS and ICs have been outlined. Materials selection for MEMS packaging has been considered as well as MEMS package reliability. As far as MEMS generic elements are concerned, a number of common failure modes have been presented. Reliability of MEMS devices requires better understanding of mechanisms that cause failure in MEMS devices. Production of reliable MEMS device requires sophisticated design considerations and better control of microfabrication processes that are used in production and packaging of a MEMS device. Reliable MEMS package should isolate non-sensing areas from sensing ones what is of extreme importance in harsh, corrosive or mechanically demanding environments. Also, it must not prevent mechanical action of moving parts of the structure or disable transfer of fluids from one region to another. Coupling of energy, motion or momentum from one region to another should be allowed. Finally, reliable MEMS package should prevent transfer of heat, mechanical strain, outgassing, pressure, moisture, etc. Reliability of MEMS generic elements is also of utmost importance. Knowledge of physics of degradation and failure mechanisms in the microdomain is still very limited. Another important issue is the need for credible testing techniques to be used during fabrication, assembly and packaging as well as during operation of the device. Device with self-testing capability will insure the reliability of the device during service.

It should be pointed out that little research and development efforts have been made in the area of testing. MEMS reliability studies lack dedicated equipment and the development of new and the upgrading of existing equipment is highly desirable. Also, highly diversified functions and materials involved make industrial standard for MEMS packaging almost impossible task. Projected timeline for standardization of MEMS technology is at least five years away (Hsu, 2006). Till then MEMS devices will be custom made according to customer's requirements and the lack of information flow as well as the reluctance in sharing experience and knowledge will keep MEMS still far away from the full commercialization.

5. Acknowledgement

6. References

Hsu, T-R. (2006). Reliability in MEMS packaging, *Proceedings of 44th International Reliability Physics Symposium*, ISBN: 0-7803-9498-4, San Jose, CA, March 26-30, 2006, IEEE International

Merlijn van Spengen, W. (2003). MEMS reliability from a failure mechanisms perspective. *Microelectronics Reliability*, 43, 7, (2003) 1049-1060, ISSN: 0026-2714

Muhlstein, C. & Brown, S. (1997). Reliability and Fatigue testing of MEMS, In: *NSF/AFOSR/ASME Workshop, Tribology Issues and Opportunities in MEMS*, Bhushan, B. pp. 519- 528, Springer, ISBN: 0792350243

Pecht, M.G.; Agarwal, R.; McCluskey, P.; Dishongh,T.; Javadpour, S. & Mahajan, R.(1998). Electronic Packaging Materials and Their Properties, CRC Press, ISBN: 978-0849396250, USA

Tadigadapa, S. & Najafi, N. (2001). Reliability of Microelectromechanical Systems (MEMS), *Proceedings of Reliability, Testing, and Characterisation of MEMS/MOEMS Conference*, pp. 197-205, ISBN: 0-8194-4286-0, San Francisco, CA, October 22-24 2001, SPIE, Bellingham, USA

Numerical Simulation of Plasma-Chemical Processing Semiconductors

Yurii N. Grigoryev and Aleksey G. Gorobchuk

Institute of Computational Technologies Russian Academy of Sciences, Siberian Branch
Russia

1. Introduction

The growing rates of microchip world production during the last two decades exceed essentially the corresponding indexes of any others production branches. Between 1997 and 2003 the consumer and communication electronics sales have grown from USD 744 billion to about USD 165 trillion. The present-day electronics is based on a silicon technology and such a state will be conserve at least during the nearest ten years. The low temperature plasma facilities - so called plasma reactors or glow discharge reactors, play an important role in technological processes of chip production.

Such reactors are widely used for etching and deposition of semiconductor films, for taking off photoresist and some other operations. Very often they enter the complex cluster equipment for making chips.

Some characteristic schemes of these glow discharge reactors are presented in Fig. 1. The typical reactor consists of two parallel plate electrodes forming an axisymmetrical cylinder chamber, in which the high-frequency discharge is appeared. The processing wafer is placed on one of the electrodes. The originally inert feed gas enters the discharge zone where an active etchant species is produced by the electron - impact dissociation. The active species transfers to the wafer and reacts with it forming the volatile products. The unreacted feed gas and the products of physical-chemical processes and reactions are pumped outwards from the reactor.

From the presented schemes one can see that these reactors are not very complicated and expensive apparatus. But yet in 1995 a world volume of sales of the reactors have made up USD 2 billion and it keeps on growing. This numeral can give us a rough idea about the quantity of operative reactors in modern industry.

Despite of the relative simple construction the etching process in a reactor is a very complicative one. For silicon wafer operating the complex molecular gases such as CF_4, SF_6 and their mixtures with oxygen O_2 and hydrogen H_2 are used. Under a microwave discharge and ion current in etching chamber a reacting medium appears which is characterized by simultaneously proceeding processes of ionization, dissociation, heat and mass transfer with complex chemical reactions. A similar processes take place on the surfaces of the chamber and wafer under operating.

The quality and manufacturing rate of producing chips depend strongly on a large number of process variables in a reactor including parent gas composition, pressure, temperature, frequency and power of a discharge, flow rate and configuration, etc. Because of numerous

and complex interconnections of the factors which defines the qualities of etching wafer, the opportunities of experimental studies and optimization of reactor process are very restricted.
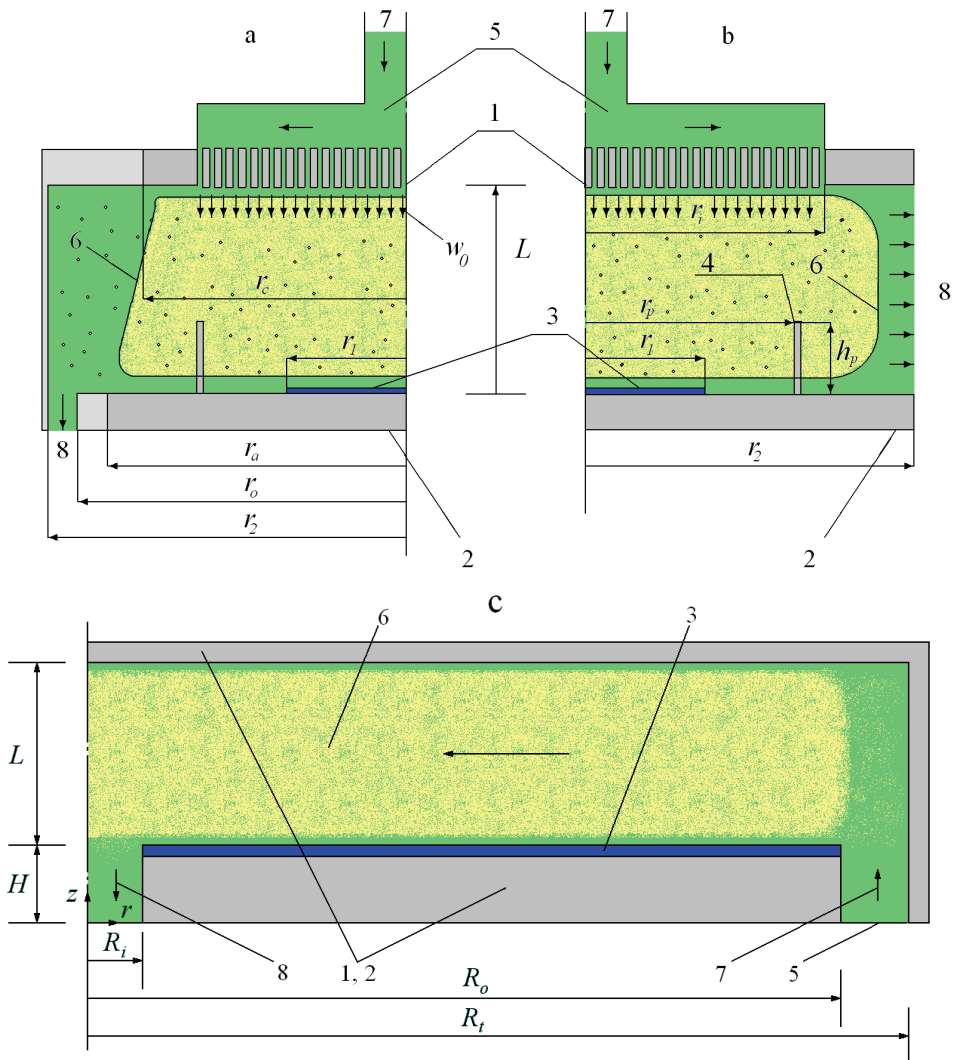
Fig. 1. The schemes of plasma - chemical etching reactors: a – "pedestal", b – "stadium", c - radial flow reactor. 1, 2 - RF - electrodes, 3 - processing wafer, 4 - protector, 5 - feed gas, 6 - RF - discharge zone, 7 - inlet, 8 - outlet. The arrowed lines show the direction of the gas flow in the reactor.

A natural alternative here is the mathematical modelling. It is especially necessary in respect to insufficient understanding of many real plasma physics and chemistry governing mechanisms which take place in this apparatus.

By such a way there are economical, technical and scientific preconditions for the well-directed efforts in the development of mathematical modelling of plasma etching reactors.

2. Numerical model formulation

In first turn some characteristic features of authors' numerical model of plasma etching process will be described. The numerical model was developed during several years with successive improving its adequacy and prognostic abilities step by step (Grigoryev & Gorobchuk, 1996; Grigoryev & Gorobchuk, 1997; Grigoryev & Gorobchuk, 1998; Shokin et al., 1999; Grigoryev & Gorobchuk, 2004; Grigoryev & Gorobchuk, 2007; Grigoryev & Gorobchuk, 2008). Today the created model corresponds completely to the world standards in mathematical modelling of plasma reactors and includes some novel elements.

2.1 Gas flow and temperature distribution
Under the typical operating conditions in plasma reactors the continuum approach is valid, and gas flow is laminar, viscous and incompressible. Therefore, the steady Navier - Stokes equations with heat transfer in standard Boussinesq approximation were used for the flow description (Grigoryev & Gorobchuk, 1997; Grigoryev & Gorobchuk, 1998; Shokin at al., 1999). The axisymmetric statement of a problem is considered. The conservation equation of total mass (continuity equation) was written as follows:

$$\nabla \cdot \mathbf{v} = 0 \tag{1}$$

The conservation equation of momentum had the form:

$$\rho_0 \mathbf{v} \cdot \nabla \mathbf{v} = \nabla \cdot \tau - \rho_0 \mathbf{g}\beta(T - T_0),$$
$$\tau = -p\mathbf{I} + \eta[\nabla \mathbf{v} + (\nabla \mathbf{v})^*] \tag{2}$$

where ρ is a density of gas mixture, $\mathbf{v}$ is a fluid velocity vector, p is a pressure, τ is a stress tensor, $\mathbf{I}$ is a identity matrix, T is a local temperature of gas mixture, T_0 is a temperature of the feed gas at the inlet of the reactor, $\mathbf{g}$ is a gravitational acceleration vector, β is a thermal expansion coefficient, η is a shear viscosity. The density ρ_0 corresponds to the gas temperature T_0. For the velocity components on impenetrable walls the nonslip boundary conditions were used in range of operating pressures p = 0.1 – 1.0 torr and slip conditions for low pressures p = 0.01 – 0.1 torr correspondingly.

The temperature distribution was obtained by solving the energy balance equation with heat transfer at the surfaces of reactors (Grigoryev & Gorobchuk, 1997; Grigoryev & Gorobchuk, 1998; Shokin et al., 1999):

$$\rho c_p(\mathbf{v} \cdot \nabla T) = \nabla \cdot (\lambda \nabla T) - \nabla \cdot \mathbf{q}_r, \tag{3}$$

where c_p is a constant - pressure heat capacity, λ is a gas thermal conductivity, $\mathbf{q}_r$ is a radiation flow rate. The radiation flow $\mathbf{q}_r$ under operating pressures p = 0.1-1.0 torr was calculated in thin optical layer approximation.

The dynamical and energy balance equations were coupled through the temperature dependence of gas viscosity and the buoyancy term. The gas viscosity, thermal conductivity and heat capacity were considered as functions of temperature. The boundary conditions on the temperature expressed a balance of convective flow, heat conduction and radiation heat

flows at the solid walls. At the axis of symmetry the no flux boundary condition was used. The gas temperature at the inlet of reactor is equal to the wall temperature. Under low pressures a "temperature jump" condition was used.

2.2 Physical-chemical kinetics and species concentration distribution

In general case a binary mixture CF_4/O_2 was considered as a parent gas because it is widely spread in silicon technology (Grigoryev & Gorobchuk, 2004; Grigoryev & Gorobchuk, 2007). An important difficulty for CF_4/O_2 system is a simulation of plasma-chemical kinetics which is extraordinarily complicated. Generally the governing set of chemical reactions and corresponding number of reagents essential for given chemical system are chosen using real experimental data. For Si - CF_4/O_2 parent system a subset of 14 gas-phase reactions were derived which describes adequately the experimental observations (Plumb & Ryan, 1986). This improved chemical kinetic model was added by several heterogeneous reactions with CF_2, CF_3 radicals (Venkatesan at al., 1990; Sang-Kyu Park & Economou, 1991).

The kinetic model included the following processes: electron-impact dissociation of binary gas mixture, volume recombination of reactive atoms and radicals, silicon etching, chemisorption of fluorine and oxygen atoms on Si surface, recombination and adsorption of CF_2, CF_3 at wafer. The complete set of reactions used in the paper looks as follows:

$$CF_4 + e^- \xrightarrow{k_{e1}} CF_3 + F + e^-, \tag{4}$$

$$CF_4 + e^- \xrightarrow{k_{e2}} CF_2 + 2F + e^-, \tag{5}$$

$$O_2 + e^- \xrightarrow{k_{e3}} O + O + e^-, \tag{6}$$

$$COF_2 + e^- \xrightarrow{k_{e4}} COF + F + e^-, \tag{7}$$

$$CO_2 + e^- \xrightarrow{k_{e5}} CO + O + e^-, \tag{8}$$

$$CF_3 + CF_3 + M \xrightarrow{k_{v1}} C_2F_6 + M, \tag{9}$$

$$F + CF_3 + M \xrightarrow{k_{v2}} CF_4 + M, \tag{10}$$

$$F + CF_2 + M \xrightarrow{k_{v3}} CF_3 + M, \tag{11}$$

$$O + CF_3 \xrightarrow{k_{v4}} COF_2 + F, \tag{12}$$

$$O + CF_2 \xrightarrow{k_{v5}} COF + F, \tag{13}$$

$$O + CF_2 \xrightarrow{k_{v6}} CO + 2F, \tag{14}$$

$$O + COF \xrightarrow{k_{v7}} CO_2 + F, \tag{15}$$

$$F + COF + M \xrightarrow{k_{v8}} COF_2 + M, \tag{16}$$

$$F + CO + M \xrightarrow{k_{v9}} COF + M, \tag{17}$$

$$F + F + M \xrightarrow{k_{v10}} F_2 + M, \tag{18}$$

$$F_2 + M \xrightarrow{k_{v11}} F + F + M, \tag{19}$$

$$CF_3 \xrightarrow{k_{s1}} CF_3\,(s), \tag{20}$$

$$CF_2 \xrightarrow{k_{s2}} CF_2\,(s), \tag{21}$$

$$F + CF_2\,(s) \xrightarrow{k_{s3}} CF_3, \tag{22}$$

$$F + CF_3\,(s) \xrightarrow{k_{s4}} CF_4, \tag{23}$$

$$CF_3 + CF_3\,(s) \xrightarrow{k_{s5}} C_2F_6, \tag{24}$$

$$CF_2\,(s) + O \xrightarrow{k_{s6}} CO + 2F, \tag{25}$$

$$CF_3\,(s) + O \xrightarrow{k_{s7}} CO + 3F, \tag{26}$$

$$O \xrightarrow{k_{s8}} O\,(s), \tag{27}$$

$$O\,(s) + F \xrightarrow{k_{s9}} O + F, \tag{28}$$

$$4F + Si \xrightarrow{k_s} SiF_4 \uparrow, \tag{29}$$

$$4F + Si + I^+ \xrightarrow{k_i} SiF_4 \uparrow. \tag{30}$$

where k_{e1}-k_{e5} are the rate constants of electron-impact dissociation of parent gas; k_{v1}-k_{v11} are the rate constants of volume recombination; k_{s1}-k_{s7} are the rate constants of heterogeneous reactions. The designation (s) marks the species adsorbed on the wafer surface. The values of these constants were taken from (Plumb & Ryan, 1986; Venkatesan at al., 1990; Sang-Kyu Park & Economou, 1991).

The model contains 16 gas-phase reactions and 8 heterogeneous reactions on the wafer. Reactions Eqs. (4)-(8) represent the electron-impact dissociation of binary gas mixture; Eqs. (9)-(19) are the reactions of volume recombination of reactive atoms and radicals; Eqs. (20)-(26) are the reactions of recombination and adsorption of CF_2, CF_3 at wafer; Eqs. (27), (28) are the chemisorption processes of fluorine and oxygen atoms on Si surface; Eqs. (29), (30) are the reactions of spontaneous and ion-induced silicon etching correspondingly. The

twelve products of dissociation and recombination processes - F, F_2, CF_2, CF_3, CF_4, C_2F_6, O, O_2, CO, CO_2, COF, COF_2 are taken into account.

Accordingly to multicomponent chemical kinetic model the distribution of species concentration for each component was derived from the system of convective-diffusion equations:

$$\mathbf{v} \cdot \nabla C_i = \nabla \cdot (C_t D_i (\nabla x_i + k_T \nabla \ln T)) + G_i(C_i, C_j), \quad i, j = 1, \ldots, 12, \tag{31}$$

where C_i, x_i are the molar concentration and molar fraction of species i correspondingly, C_t is the molar gas concentration, D_i is the multicomponent diffusion coefficient of species i, k_T is the thermal diffusion relation, G_i is the rate of formation of species i in gas-phase reactions. The gas phase reactions are incorporated in right-hand side of this system and define a complex interconnection between all species generation processes. The surface and silicon etching reactions entered the boundary conditions at the wafer. The latter were written as a balance of mass flows for each component. At the reactor inlet the Danckwert's type boundary conditions were stated (Sang-Kyu Park & Economou, 1991). The feed species concentrations at the inlet are fixed. No radial gradients of species concentrations are considered at the reactor centerline. At the reactor outlet, zero axial gradients of species concentrations are also used.

2.3 Glow discharge structure and electron concentration

The exact calculation of glow discharge structure demands a solving the Boltzmann kinetic equation for the electrons in a mixture multiatomic gases and radicals. From both physical and computational points of view this is a formidable task (Aydil at al., 1993). Therefore in the parametric calculations some simplest model distributions of electron density in reactor were used. Depending on the pressure and gas medium under consideration, the dominant electron loss mechanism can be diffusion, recombination or attachment. In calculations usually it was assumed that the electron density distribution corresponded to a "diffusion-dominated" discharge (Dalvie at al. 1986).

2.4 Numerical method

The presence of two-order elliptic operators in all equations of the mathematical model allows us to approximate each equation by implicit iterative finite difference splitting-up scheme with stabilizing correction (Grigoryev & Gorobchuk, 1996). The scheme in general form looks as follows:

$$\frac{\phi^{k+1/2} - \phi^k}{\tau} = L_r^\phi \phi^{k+1} + L_z^\phi \phi^k + F(\phi^k),$$

$$\frac{\phi^{k+1} - \phi^{k+1/2}}{\tau} = L_z^\phi (\phi^{k+1} - \phi^k).$$

The scheme has $O(\tau + h_1^2 + h_2^2)$ approximation order where h_1, h_2 are the mesh sizes along r and z coordinates, τ is the iterative parameter. The solution of the original steady state problem was derived by the relaxation method. The iterative process was terminated after achieving the relative error $\varepsilon_\phi = 10^{-10} - 10^{-4}$ in the uniform norm

$$\max_{\Omega_h} \left| \frac{\phi^{k+1} - \phi^k}{\tau \phi^{k+1}} \right| < \varepsilon_\phi.$$

The equations (1), (2) were solved together with the heat transport equation (3). Calculations were worked out in the "stream function – vorticity" variables. The second-order Thom's vorticity boundary condition was applied to the flow problem. The stream function and gas temperature were found for each iteration of vorticity. The species concentrations were then calculated from the convective - diffusion equations (31) using the resulting velocity and temperature distributions.

3. Main results of plasma-chemical reactor modelling

Here some new effects obtained in our studies will be commented. These effects were not considered before in the literature.

3.1 Optimization of reactor design with respect to etching uniformity
Firstly it will be demonstrated that the mathematical modelling even in frameworks of simplified model can give the results useful for technical applications. In (Grigoryev & Gorobchuk, 1996) we considered two most spread plasma-chemical etching reactor schemes - "pedestal" and "stadium" ones, which are used for individual etching of wafer with diameters up to 500 mm. They are shown in Fig. 1.
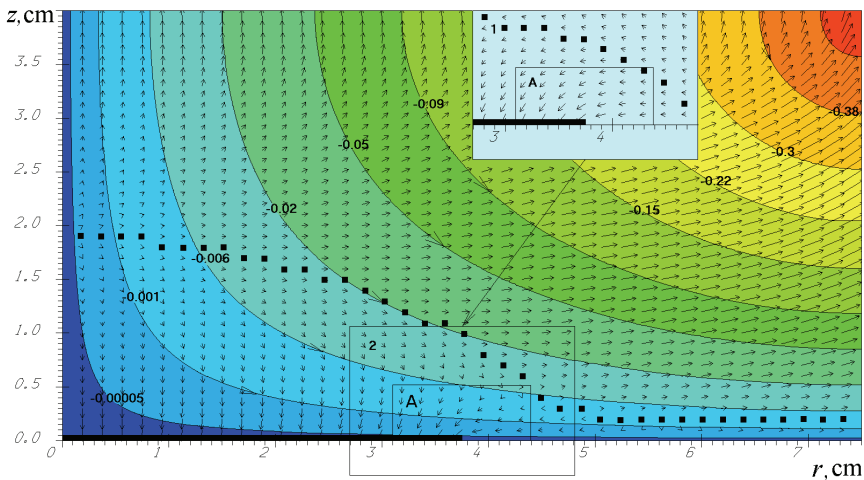
Fig. 2. Isolines of stream function and full flow of etchant reactant in "stadium" type reactor without the protector. 1, 2 - distributions of diffusion and full flows of etchant in zone **A**. Processing regimes: p = 0.2 torr, Q = 30 cm³/min, I^+ = 0 mkA/cm².

The ends of cylinder chamber are employed as the electrodes between which the plasma RF-discharge is exited. The parent gas enters uniformly through upper electrode porous wall. The residual feed gas and products of dissociation, recombination and etching are pumped outwards either in radial ("stadium") or in axial ("pedestal") direction through circle gap on periphery of lower electrode.

It was noted that under operating in these industrial reactors an essential nonuniformity appeared at the outer edge of the patterns. As a consequence up to 30% of the initial wafer square went into a defective part. To minimize such an edge nonuniformity it was suggested to surround a pattern by a cylindrical protector with low etching reactivity. The problem was to obtain an optimal protector geometry.

For this purpose the dimensions and operating parameters were taken in the range which is characteristic for industrial reactors. The etching process of silicon Si in a tetrafluoromethane plasma CF$_4$ was chosen as a basic one. Because of relatively small temperature gradients during the etching process it was used the isothermal approach. In parametric calculations the multicomponent gas medium in a reactor was considered as binary gas mixture consisted of the fluorine atoms F supporting the etching reaction on the wafer and the feed gas CF$_4$. In this case the active species concentration distribution was derived by solving a single convection-diffusion equation with the generation term describing the generation (Eqs. (4), (5)) and depletion of active component in reactor volume (Eqs. (10), (11)).

To understand a mechanism of edge defect appearing and effect of protector the vector fields of fluorine flow densities were calculated:

$$\mathbf{Q}_e = \mathbf{Q}_c + \mathbf{Q}_d$$

where $\mathbf{Q}_e$ is the full flow density, $\mathbf{Q}_c = \mathbf{v}C_F$ and $\mathbf{Q}_d = -D_F C_t \nabla x_F$ are the convective and diffusion ones correspondingly, D_F is the binary diffusivity of fluorine atoms in CF$_4$. The typical distribution of the full flow density $\mathbf{Q}_e$ in "stadium" reactor without protector is presented in Fig. 2.
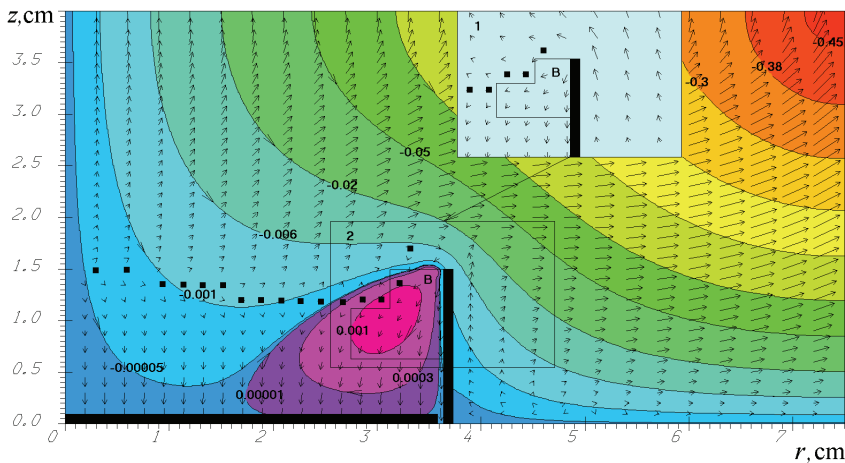
Fig. 3. The stream function isolines and distribution of full flow of etchant in "stadium" type reactor with optimal protector r_p = 38 mm, h_p = 15 mm. 1, 2 - distributions of diffusion and full flows of etchant in zone **B**. Processing regime: p = 0.2 torr, Q = 30 cm^3/min, I^+ = 0 mkA/cm^2.

One can see that near the outer edge the characteristic zone **A** exists where the relatively intensive diffusion of fluorine to the wafer takes place. It is connected with the large difference in etching reactivities of wafer and anode surfaces. The black markers single out

the layer where $\partial x_F / \partial z = 0$ and the diffusion flow changes a sign. Consequently the etching nonuniformity in this case is defined by the nonuniformity of diffusion flow near the wafer edge.

The parametric calculations for different values of height and diameter of a protector were fulfilled. The results allowed us to choose the optimal sizes of protector. Fig. 3 presents the full flows in "stadium" reactor with optimal protector. Thereat material of protector has a low reactivity. One can see from Fig. 3, that circle protector which has the same radius as the wafer interrupts completely the local diffusion flow arising from difference in reactivities of wafer and anode. Although near the top edge of protector zone **B** is appeared where the influence of low anode reactivity is preserved.

Despite of these results, obtained for the very simple plasma chemical kinetics, the further investigations have shown that the optimal protector provides high etching uniformity for enough complicated models also.

3.2 Heat radiation transfer and thermodiffusion

Plasma - chemical etching is usually related to the category of low temperature processes. Therefore, at the mathematical simulation of etching reactors one can limit himself by isothermal approach giving often the satisfactory results (Grigoryev & Gorobchuk, 1996). However the employment of low heatproof resists and thermosensitive polymers for the wafers requires thorough investigation of the heating the processing chip and the elements of reactor construction. Also it is necessary to determine the heating influence on the processing quality. The main sources of heating effects in reactor are the heat generation on the wafer and surrounded electrode by the energetic ion bombardment, plasma radiation, heat effects of the exothermic reactions and glow discharge. The heat removal is realized by the complex heat transfer in reactor chamber and by the cooling system if such a system exists. Some nonisothermal effects in the reactor as a function of the electrode and wafer temperature were investigated. The temperature was specified and varied in characteristic limits T_{w2} = 300 – 500 K (Grigoryev & Gorobchuk, 1997; Grigoryev & Gorobchuk, 1998). The heat radiation transfer in the gas was determined using the optical thin layer approximation. According to such an approach the source of heat radiation in (3) have the following form:

$$\nabla \cdot \mathbf{q}_r = -2\sigma \left[\kappa_p(T_{w1})T_{w1}^4 - \kappa_p(T)T^4 + \kappa_p(T_{w2})T_{w2}^4 \right],$$

where κ_p is the Plank average absorption coefficient in the parent gas, σ is the Stefan-Boltzmann constant, T_{w1} is the temperature of cathode, T_{w2} is the temperature of anode. The problem which we have obtained consisted in the absence of the necessary data about the emissivity of CF$_4$. Because the value κ_p was estimated by the emissivity of methane CH$_4$ having the same structure and the similar main vibration modes as the molecule CF$_4$. For calculation of the spectral absorption coefficient κ_v the exponential model of spectral band was used (Edwards & Menard, 1964). The boundary conditions on the temperature described the balance of heat flows, namely convective, heat conductive and radiative ones. For example, at the wafer surface:

$$-\lambda \frac{\partial T}{\partial z} = \alpha(T_{w2} - T) + \sigma \varepsilon \varepsilon_w (T_{w2}^4 - T^4), \quad 0 \le r \le r_1, z = 0,$$

where α is the heat transfer coefficient, ε_w is the emissivity of the wafer, ε is the emissivity of the gas.

The boundary conditions for fluorine species took into account the diffusion and thermodiffusion of active species, its heterogeneous recombination and consumption in the chemical reactions of spontaneous and ion - induced etching. For instance, the boundary conditions at the wafer surface had the following form (in binary approach):

$$D_F\left(\frac{\partial x_F}{\partial z} + k_T \frac{\partial \ln T}{\partial z}\right) = k_s x_F (1 - \mu x_F) + k_i I^+ (1 - \mu x_F)/C_t,$$

$$\mu = 1 - m_F / m_{CF_4}, \quad 0 \leq r \leq r_1, z = 0,$$

where k_T is the thermodiffusion ratio; k_s, k_i are the constants of spontaneous (Eq.(29)) and ion - induced etching (Eq.(30)); m_F, m_{CF_4} are the molecular masses of active species and parent gas, respectively.
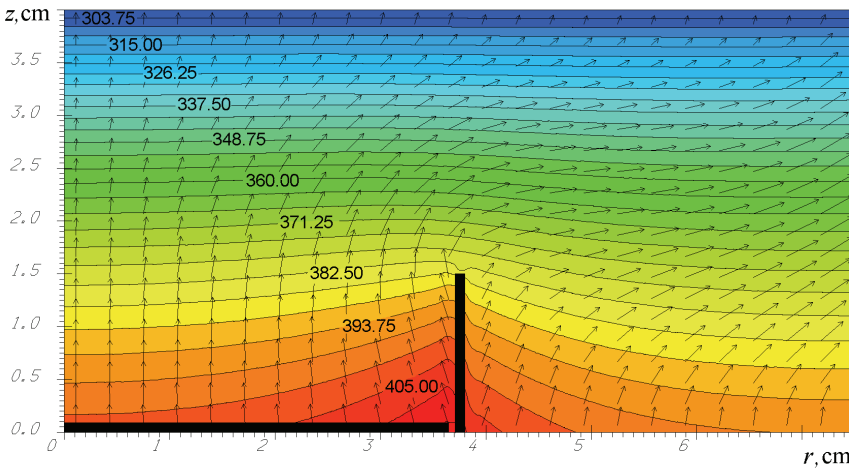
Fig. 4. The distribution of isotherms and full heat flow $\mathbf{q}_h$ in "stadium" type reactor. Processing regime: p = 1.0 torr, Q = 50 cm³/min, T_{w2} = 500 K, I+ = 0 mkA/cm².

In calculations the vector fields of local density heat flow were analyzed. It is consisted of three components:

$$\mathbf{q}_h = \mathbf{q}_a + \mathbf{q}_c + \mathbf{q}_r,$$

where

$$\mathbf{q}_a = \rho c_p T \mathbf{v}, \quad \mathbf{q}_c = -\lambda \nabla T$$

are the densities of heat convective and heat conductive flows. The characteristic picture of full heat flow $\mathbf{q}_h$ distribution is shown in Fig. 4. It was obtained that the main contribution in $\mathbf{q}_h$ is given by the heat conduction and heat radiation flows in cylindrical volume over the substrate. In particular, near the wafer $|\mathbf{q}_c| \approx |\mathbf{q}_r|$, and they exceed convective flow $\mathbf{q}_a$ over two order. Simultaneously $\mathbf{q}_c$ and $\mathbf{q}_r$ have the same order at the outlet in the middle part of reactor, but $\mathbf{q}_a$ exceeds them by factor of 1.5 - 2. It is worth to note that if the heat radiation transfer is not taken into consideration the values of temperature on the isolines in Fig. 4 reduce approximately on 60K. This allows one to conclude that despite of the

temperature nonuniformities characteristic for etching reactors are relatively not large, the main heat transfer in the reactor realizes by the heat conduction and radiation. Therefore, a radiation heat transfer is necessary to take into account in numerical modelling.

Under the thermal nonuniformity conditions the flow density of active species may be written as follows:

$$\mathbf{Q}_e = \mathbf{Q}_c + \mathbf{Q}_d + \mathbf{Q}_t$$

where there are the convective flow $\mathbf{Q}_c = \mathbf{v}C_F$, diffusion one $\mathbf{Q}_d \sim -D_F\nabla x_F$ and thermodiffusion one $\mathbf{Q}_d \sim -D_F k_t \nabla \ln T$. The distribution of full flows $\mathbf{Q}_e$ of etchant is shown in Fig. 5.
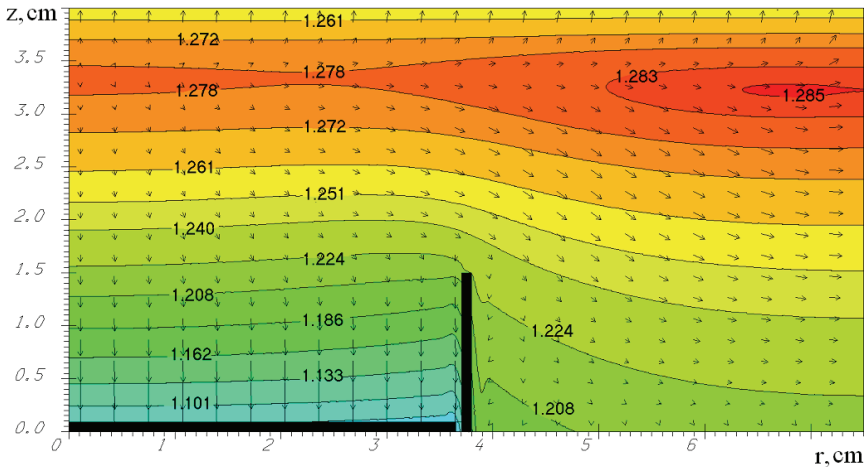
Fig. 5. The distribution of full flow of active species $\mathbf{Q}_e$ in "stadium" type reactor and isolines of concentration (C x 10^{-10}, mol/cm³). Processing regime: p = 0.2 torr, Q = 50 cm³/min, T_{w2} = 500 K, I^+ = 0 mkA/cm².

Since the thermodiffusion ratio $k_T < 0$, the vectors $\mathbf{Q}_t$ direct along the gradients of temperature. The calculations show that $|\mathbf{Q}_d| \approx |\mathbf{Q}_t|$ at the temperatures of lower electrode T_{w2} = 400 – 500 K outside of zone limited by the protector, and they determine the value of full flow $|\mathbf{Q}_d| \approx |\mathbf{Q}_t|$ substantially exceeding $|\mathbf{Q}_c|$. However immediately on the substrate $|\mathbf{Q}_t| \approx (0.1 \div 0.2)|\mathbf{Q}_d|$. It means that the direct contribution of the thermodiffusion $\mathbf{Q}_t$ to the etching processes gives 10 - 20% and under the local temperature gradients it may negatively affect the etching uniformity.

These conclusions about the significant role of the heat radiation and thermodiffusion have a general character for typical conditions of plasma etching process that was supported by our calculations of other reactor schemes.

3.3 Effect of choice of plasma etching kinetics

In foregoing decade the process of plasma chemical etching of silicon in CF$_4$ was studied by many authors for different reactors and some variants of chemical kinetics. Unfortunately,

different original data, geometrical configurations of reactors, numerous additional assumptions did not allow one to compare obtained results for choosing an adequate kinetic model. To do such a choice we have carried out a series of calculations of radial flow etching reactor that is a necessary aggregate in VLSI industry. The results obtained in the frameworks of one numerical model give us reliable data for comparison of wide used variants of Si - CF_4 plasma kinetics with respect to etching rate (Shokin at al., 1999).

The scheme of radial flow plasma - chemical etching reactor is shown in Fig. 1. The feed gas enters the reacting chamber at the outer edge of lower electrode. RF - discharge is appeared between the electrodes. Products of physical - chemical reactions are pumped from outlet in the centre of lower electrode. The processing wafer is placed on the lower electrode. The dimensions of reactor were taken from (Dalvie at al., 1986).
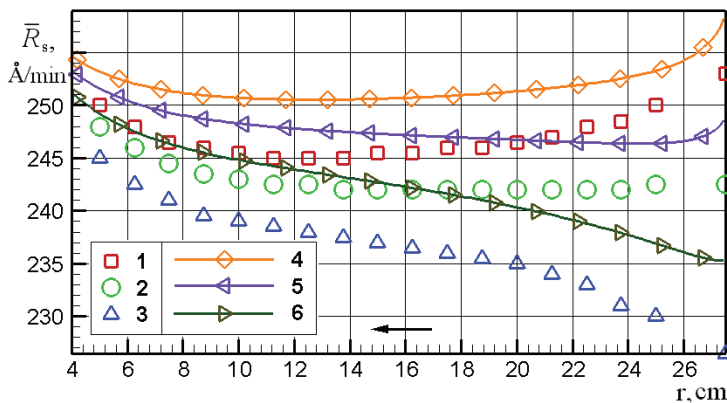
Fig. 6. Distributions of etching rate as a function of radial position along the wafer for different gas flow rate (inflow feed gas structure). Processing regimes: p = 0.525 torr, T_{w2} = 300 K. Designations: 1, 4 - Q = 300 cm^3/min; 2, 5 - Q = 340 cm^3/min; 3, 6 - Q = 400 cm^3/min. Markers 1, 2, 3 - data of article (Dalvie at al., 1986). Markers 4, 5, 6 - author data.

The operating conditions of reactor and process parameters were varied in the range that is characteristic for industrial reactors. In particular, it was chosen: the pressure p = 0.5 torr, the gas flow rate Q = 300 – 400 cm^3/min, the average electron density $\bar{n}_e$ =10^{10} cm^{-3}, the temperature of electrodes T_{w1} = 300 K, the temperature of wafer T_{w2} = 300 – 500 K. Since only a small fraction of tetrafluoromethane is dissociated under RF-discharge (<10%), all macroscopic characteristics of the medium - $\eta, \lambda, c_p, \kappa_p$ were chosen for the pure tetrafluoromethane CF_4. The dependencies of gas viscosity, thermal conductivity, heat capacity and absorption coefficient on the gas temperature and pressure were taken into account.

The homogeneous and heterogeneous reactions included in the basic kinetic model in these calculations was following: the reactions of fluorine production by electron - impact dissociation of parent gas Eqs. (4), (5); the reactions of volume recombination of reactive atoms and radicals Eqs. (9)-(11); the reactions of recombination of fluorine atoms, the recombination of CF_3 and the adsorption of CF_3 at wafer surface Eqs. (20), (23), (24); the reaction of silicon etching Eq. (29).

Before the comparison of kinetic models the test calculations with kinetics and data used in (Dalvie at al., 1986) have been fulfilled. At these conditions our simulations of isothermal reactor showed the satisfactory agreement with the results in (Dalvie at al., 1986). The comparison of the obtained profiles of etching rates with those reported in (Dalvie at al., 1986) are presented in Fig. 6.

The maximum relative and maximum root-mean-square errors did not exceed 0.05799 and 0.02166, correspondingly. Thus, it was shown that the results obtained by two different numerical methods coincide with a good accuracy. It allowed one to be sure in further results of comparison of the kinetic models.

A special interest of our research was a comparison of two-, three- and four- component kinetic models considered in (Grigoryev & Gorobchuk, 1996; Dalvie at al., 1986; Sang-Kyu Park & Economou, 1991) correspondingly. Two- and three-component kinetics were obtained as some simplifications of the basic kinetic model. Thereat it was essentially used that the concentrations of products of gas-phase reactions C_F, C_{CF_2}, C_{CF_3} are significantly smaller then the total gas concentration C_t and the unreacted feed gas concentration C_{CF_4}. For two-component kinetic model it was assumed that $C_{CF_2} \sim C_{CF_3} \sim C_F$. In such a case only one component convection-diffusion equation was solved to calculate a fluorine concentration C_F. The concentrations C_{CF_2}, C_{CF_3} were taken into account as parameters. Therefore, the net generation term in equation (31) describing the processes (9)-(11) introduced by some authors as follows:

$$G_F = (k_{e1} + k_{e2})\overline{n}_e(1 - 3C_F) - (k_{v2} + k_{v3})C_F^2 \tag{32}$$

In (Edelson & Flamm, 1984; Plumb & Ryan, 1986) it was shown that the most important reactions are (4) and (9). The three-component kinetics incorporates two equations on F and CF_3 concentrations. The C_{CF_2} was regarded as the parameter ($C_{CF_2} \sim C_F$). Eventually the four-component kinetic model contains all above-mentioned equations for the species concentrations of F, CF_2, CF_3.

The spontaneous etching rate was calculated as follows:

$$R_s = 1.75 \cdot 10^{12}(1 - \vartheta_{CF_3})C_F T^{1/2}\exp\left[-\frac{E_e}{kT_{w2}}\right],$$

where R_s in   Å /min; E_e = (0.108 ± 0.005) eV; T is the gas temperature near the wafer, K; ϑ_{CF_3} is the parameter of silicon coverage by CF_3. The average etching rate and fluorine concentration were determined by the formulas:

$$\overline{R}_s = \frac{2}{R_o^2 - R_i^2}\int_{R_i}^{R_o} R_s r dr, \qquad \overline{C}_F = \frac{2}{R_o^2 - R_i^2}\int_{R_i}^{R_o} C_F r dr, \qquad I_n = \frac{R_s^{max} - R_s^{min}}{2R_s}.$$

Their variation along the wafer as a function of temperature for two-, three- and four-component kinetics is presented in Table 1.

The obtained data show that the values of average etching rate and fluorine concentration obtained for two-, three- and four- component kinetic models are strongly differ from each other for any fixed reactor regime (see Table 1). These distinctions are connected with unsuccessful parametrization of species generation terms (32) in two- and three-component

models. The obtained results show that there is a problem of choosing of adequate kinetics for CF_4 - kinetics that up to day was not discussed in the corresponding studies. In order to solve this problem it is necessary to compare these results with the reliable experimental data. Unfortunately the latter are absent in the present-day existent literature.

T_{w2}, K	$\bar{R}_s$, Å/min			$\bar{C}_F \cdot 10^{-9}$, mol/cm³			I_n		
	Number of components			Number of components			Number of components		
	2	3	4	2	3	4	2	3	4
Inflow									
300	83.05	373.50	624.94	0.17844	0.80833	1.34619	0.0888	0.1106	0.0528
373	189.17	823.89	968.95	0.16869	0.74013	0.86954	0.1017	0.1178	0.1049
473	389.56	1591.27	1377.89	0.15556	0.64741	0.56379	0.1168	0.1207	0.1513
573	608.61	2375.13	1707.10	0.14276	0.56176	0.40903	0.1286	0.1176	0.1846
Outflow									
300	83.02	373.30	599.01	0.17838	0.80789	1.29064	0.1762	0.2491	0.3009
373	189.10	823.48	930.54	0.16861	0.73966	0.83491	0.1943	0.2469	0.2869
473	385.09	1590.56	1314.05	0.15545	0.64704	0.53747	0.2156	0.2440	0.2982
573	607.99	2374.27	1610.66	0.14265	0.56168	0.38604	0.2340	0.2401	0.3088

Table 1. Average etching rate, average fluorine concentration and uniformity index: two-, three- and four- components kinetic models. Processing regime: p = 0.512 torr, Q = 340 cm³/min.

3.4 Numerical optimization of silicon plasma etching in CF_4/O_2 mixture

The binary mixture CF_4/O_2 is a perspective parent gas for industrial plasma etching process that allows to increase essentially an etching rate of silicon wafers. Such mixtures are available commercially, but the number of papers devoted to mathematical modelling of this process are relatively small. Moreover in most studies the simplified models are used and the obtained results are inconsistent with experimental date in many effects. It is clear that the obtained numerical results depend on the choice of chemical kinetics and adequate modelling of complex heat-mass transfer.

In present calculations the improved chemical kinetic model (Venkatesan at al., 1990) was used including 12 all above mentioned components which are F, F_2, CF_2, CF_3, CF_4, C_2F_6, O, O_2, CO, CO_2, COF, COF_2. The heat-mass transfer in the reactor was calculated on the basis of two-dimensional mathematical model of nonisothermal reactor described above taking into account the chosen multicomponent chemical kinetics. The heat radiation in the gas mixture was determined in optical thin layer approximation. To evaluate the emissivity of gas mixture the exponential model of band adsorption (Edwards & Menard, 1964) describing the spectral adsorption of multiatomic molecules was employed.

Accordingly to existent representation of chemisorption process some parts of silicon surface are covered by various adsorbed atoms and radicals. Let $\vartheta_{CF_2}, \vartheta_{CF_3}$ are the fractions of silicon surface covered by adsorbed radicals CF_2 and CF_3 correspondingly; ϑ_O is the fraction of silicon surface covered by oxygen atoms, x_i is the molar fraction of i-th particles. Writing down the balance of mass flows for CF_2, CF_3 and O components on silicon surface at equilibrium as in (Venkatesan at al., 1990; Sang-Kyu Park & Economou, 1991) a system of linear equations for unknown parameters $\vartheta_{CF_2}, \vartheta_{CF_3}, \vartheta_O$ is obtained in the form:

$$\begin{cases} k_{s2}x_{CF_2} / (k_{s3}x_F + k_{s6}x_O) = \vartheta_{CF_2} / (1 - \vartheta_{CF_2} - \vartheta_{CF_3} - \vartheta_O) \\ k_{s1}x_{CF_3} / (k_{s4}x_F + k_{s5}x_{CF_3} + k_{s7}x_O) = \vartheta_{CF_3} / (1 - \vartheta_{CF_2} - \vartheta_{CF_3} - \vartheta_O) \\ \alpha_s x_O / x_F = \vartheta_O / (1 - \vartheta_{CF_2} - \vartheta_{CF_3}), \end{cases}$$

Solving this system allows to derive the next formulas for $\vartheta_{CF_2}, \vartheta_{CF_3}, \vartheta_O$:

$$\begin{aligned} \vartheta_{CF_2} &= k_{s2}x_{CF_2} / (k_{s2}x_{CF_2} + k_{s3}x_F + k_{s6}x_O + \Delta_{CF_2}), \\ \Delta_{CF_2} &= \Delta_{1,CF_2} + \Delta_{2,CF_2}, \\ \Delta_{1,CF_2} &= k_{s1}x_{CF_3}(k_{s3}x_F + k_{s6}x_O) / (k_{s4}x_F + k_{s5}x_{CF_3} + k_{s7}x_O), \\ \Delta_{2,CF_2} &= \alpha_s(k_{s3}x_F + k_{s6}x_O)x_O / x_F; \end{aligned}$$

$$\vartheta_{CF_3} = k_{s1}x_{CF_3} / (k_{s1}x_{CF_3} + k_{s4}x_F + k_{s5}x_{CF_3} + \Delta_{CF_3}), \tag{33}$$

$$\Delta_{CF_3} = k_{s7}x_O + \Delta_{1,CF_3} + \Delta_{2,CF_3},$$

$$\begin{aligned} \Delta_{1,CF_3} &= k_{s2}x_{CF_2}(k_{s4}x_F + k_{s5}x_{CF_3} + k_{s7}x_O) / (k_{s3}x_F + k_{s6}x_O), \\ \Delta_{2,CF_3} &= \alpha_s(k_{s4}x_F + k_{s5}x_{CF_3} + k_{s7}x_O)x_O / x_F; \end{aligned}$$

$$\vartheta_O = \alpha_s x_O / (\alpha_s x_O + x_F + \Delta_{1,O} + \Delta_{2,O}), \tag{34}$$

$$\begin{aligned} \Delta_{1,O} &= x_F k_{s2}x_{CF_2} / (k_{s3}x_F + k_{s6}x_O), \\ \Delta_{2,O} &= x_F k_{s1}x_{CF_3} / (k_{s4}x_F + k_{s5}x_{CF_3} + k_{s7}x_O), \end{aligned}$$

where the parameter $\alpha_s = k_{s8} / k_{s9}$ is introduced analogically with (Schoenborn at al., 1989; Kopalidis & Jorine, 1993). Such a presentation expresses these fractions through the rate constants of the processes.

A special case can be considered for ϑ_{CF_3}. By substitution to the Eq. (33) $x_{CF_2} = x_O = 0$ we derive the formula $\vartheta_{CF_3} \approx k_{s1}x_{CF_3} / ((k_{s1} + k_{s5})x_{CF_3} + k_{s4}x_F)$, which coincides with three component chemical kinetic model (Sang-Kyu Park & Economou, 1991).

In the second particular case supposing in formula Eq. (34) $x_F \gg x_{CF_2}, x_{CF_3}$ one can carry out an expression

$$\vartheta_O \approx \alpha_s x_O / (\alpha_s x_O + x_F). \tag{35}$$

which coincides with formula for ϑ_O presented in (Kopalidis & Jorine, 1993). In this simplified expression the denominator reflects a competition of chemisorption processes of fluorine and oxygen atoms on silicon which leads to hysteresis effect on the diagram of etching rate with respect to fluorine concentration (Mogab at al., 1978).

Under assumption that $\vartheta_{CF_2}, \vartheta_{CF_3} \ll 1$ (Grigoryev & Gorobchuk, 2007), the parameter α_s in these formulas which contains unknown rate constants of Eqs. (27), (28) can be defined by the next ratio of fluorine and oxygen atoms adsorbed on the wafer:

$$\alpha_s \approx \frac{F_O S_O}{F_F S_F} = \frac{S_O}{S_F}\sqrt{\frac{m_F}{m_O}}, \quad F_j = \sqrt{\frac{kT}{2\pi m_j}}, \quad j = \text{F,O,} \tag{36}$$

where m_j, F_j are the molecular mass and intensity of molecular flow of the j-th particles to the wafer; S_F, S_O are the sticking coefficients of oxygen and fluorine atoms on silicon (Kopalidis & Jorine, 1993).

The parameter α_s indicates the intensity of oxygen chemisorption on silicon and essentially influence the location and amplitude of maximum etching rate (Grigoryev & Gorobchuk, 2007). The nonzero fraction ϑ_O leads to hysteresis effect on the diagram of etching rate versus the fluorine concentration.

Accordingly to an improved surface kinetic model which describes the competing processes of etching, chemisorption of O and adsorption of CF_2, CF_3 on silicon the local spontaneous etching rate in Å /min was defined by the formula:

$$R_s = 1.81 \cdot 10^{10}(1 - \vartheta_O - \vartheta_{CF_2} - \vartheta_{CF_3})k_s x_F C_t,$$

where k_s is the etching rate constant, cm/s; C_t is the molar concentration of gas mixture, Mol/cm³.

The radial flow reactor was considered in the calculations (see Fig. 1). The constructive dimensions and operating regimes were taken from (Venkatesan at al., 1990). The calculations have been done for several values of gas flow rate under normal conditions Q = 100, 200, 400, 800 cm³/min. The wafer temperatures T_{w2} = 300, 373, 473, 573 K were considered. The pressure in etching chamber of reactor was equal to p = 0.5 torr. The average electron density was assumed equal to $\bar{n}_e$ = 6 x 10⁹ cm⁻³. The temperature of reactor walls was T_{w1} = 300 K. The O₂ percentage fraction in CF₄/O₂ feed gas mixture varied in the range 10-90%. Two gas flow directions were examined: "inflow", when the gas flow was directed to the centre of reactor, and "outflow" - with the reversed direction. It was of interest to study a distribution of fluorine flows and spontaneous etching rate in dependence on O₂ percentage fraction in feed mixture.

The distribution of fluorine flow in the reactor depends on many factors: operating regime of reactor, gas composition, direction of feed gas flow and etc. Figures 7-10 illustrates the distribution of flow structure, temperature, concentration and full flow of active species for outflow feed gas structure. The direction of fluorine mass transfer in the middle part of reactor height, where $|\mathbf{v}| \gg w_0$, coincides with the direction of gas flow. Near the RF-electrodes, where the velocity of gas flow is small, the fluorine mass transfer is realized by a concentration diffusion. Owing to the intensive forced convection the nonuniform distribution of full flow of active species on the wafer is arisen. Because of that the full flow of active species deviates from the normal to the wafer. Finally, as one can see from Figure 10, the mass transfer of fluorine atoms to the wafer surface depends strongly on the convective transfer that confirms the known representations about process in a radial flow reactor.
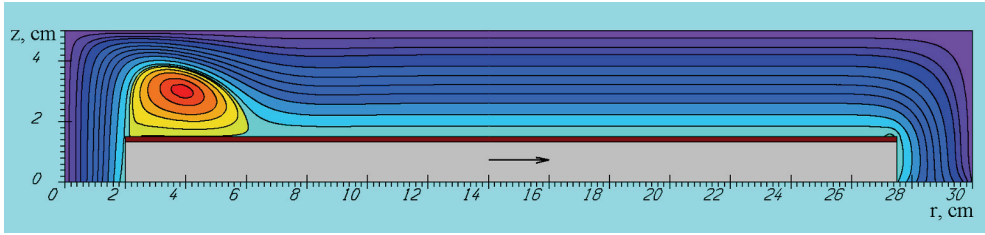
Fig. 7. The flow gas structure.

Processing regime: p = 0.5 torr, Q = 800 cm³/min, T_{w2} = 300 K, 30% fraction O_2 in CF_4/O_2.
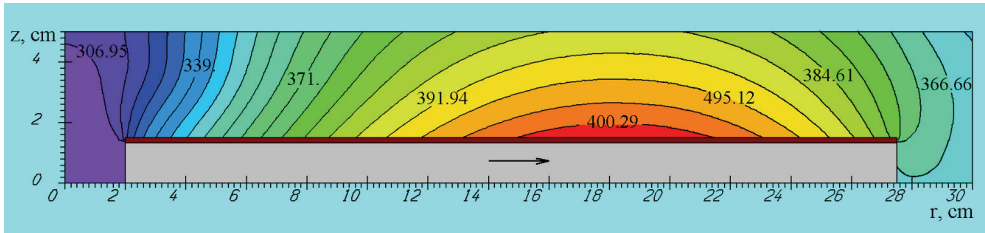
Fig. 8. The temperature distribution.

Processing regime: p = 0.5 torr, Q = 200 cm³/min, T_{w2} = 473 K, 30% fraction O_2 in CF_4/O_2.
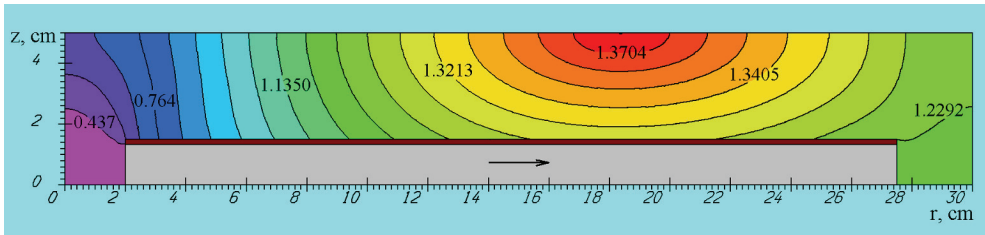
Fig. 9. The fluorine concentration distribution.

Processing regime: p = 0.5 torr, Q = 200 cm³/min, T_{w2} = 300 K, 40% fraction O_2 in CF_4/O_2.
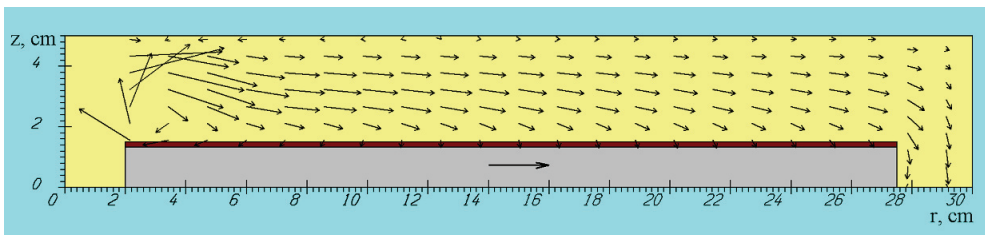
Fig. 10. The distribution of full flow of fluorine.

Processing regime: p = 0.5 torr, Q = 200 cm³/min, T_{w2} = 300 K, 40% fraction O_2 in CF_4/O_2.
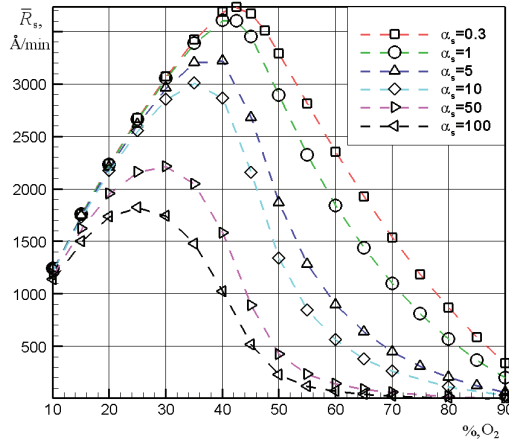
Fig. 11. Average etching rate as a function of percentage fraction of O_2 in CF_4/O_2 input mixture. Processing regime: p = 0.5 torr, Q = 200 cm^3/min, T_{w2} = 300 K. Case of "inflow" direction.

The sticking coefficients of oxygen and fluorine atoms on silicon are defined as in (Kopalidis & Jorine, 1993) S_F = 0.7, S_O = 0.2 which correspond to α_s = 0.3. In (Schoenborn at al., 1989) the chemical model included the similar parameter which was varied in the range 1 - 600. In our opinion such a range is too large, because it is clear the rise of α_s corresponds to increasing of ratio of sticking coefficients (Eq. (36)). Thereat in calculations the following values were examined α_s = 0.3, 1, 5, 10, 50, 100.

It was obtained that the average etching rate dependence on percentage fraction of O_2 in parent gas mixture has the maximum value about 40% of O_2. It takes place in the range of gas flow rate 100-400 cm^3/min (see Fig. 11, curve α_s = 0.3). In this case the maximum value of etching rate exceeds one's value in pure CF_4 more than in three times. With increasing of the gas flow rate up to 800 cm^3/min the maximum of etching rate moves to 30% of O_2. The location of extremum is independent of the feed gas flow direction and the gas temperature. With increasing O_2 percentage in gas mixture the fluorine concentration rises and has the maximum at 40% O_2. The decrease of CF_2, CF_3, CF_4, COF concentrations accompanies by the growth of fluorine concentration. The substances F_2, CO, CO_2, COF_2 are formed as the products of reactions. It corresponds to the one of channels of etching intensification by oxygen addition, which consists in the replacement of fluorine by oxygen in fluorine-containing radicals. With oxygen fraction more than 40% the fluorine concentration falls owing to decrease of the total number of CF_2, CF_3 radicals (depletion of mixture). Thus, the calculations show that the presence of etching rate maximum is connected with chemical reactions of oxygen atoms with CF_2, CF_3 radicals where the additional fluorine atoms are set free.

The etching rate dependence on fluorine concentration with different fractions of O_2 in CF_4/O_2 mixture obtained in calculations has a hysteresis character (see Fig. 12). Due to the passivation of wafer surface by oxygen atoms (with increasing ϑ_O) in the range of O_2 concentration more then 40% the silicon etching rate falls more quickly then fluorine
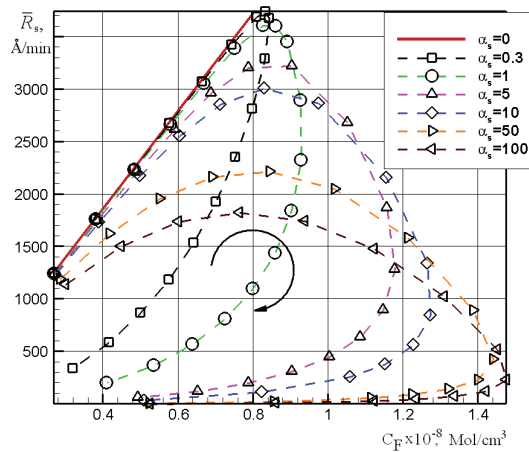
Fig. 12. Average etching rate as a function of average fluorine concentration on a wafer with different fractions of O_2 in CF_4/O_2 mixture. Processing regime: p = 0.5 torr, Q = 200 cm^3/min, T_{w2} = 300 K, "inflow" direction. The markers on the curves indicates O_2 fraction in CF_4/O_2 mixture varied in the direction of the arrow from 10% to 90% with 5% steps.

concentration. At the same time the hysteresis is not observed under the absence of oxygen chemisorption. It is illustrated by the linear dependence in Fig. 12 obtained for α_s = 0 and any fraction of O_2 in mixture. In this case the maxima of etching rate and fluorine concentration were achieved simultaneously at the same concentration of O_2 in CF_4/O_2 mixture that contradicted to experimental data (Mogab at al., 1978).

The parameter α_s essentially influence location and amplitude of maximum etching rate. With increase of parameter from 0.3 up to 100 the maximum of average etching rate moves from 40% fraction of O_2 in CF_4/O_2 to 25% (Fig. 11). At the same time the maximum of average fluorine concentration on Si surface is located near 50% fraction of O_2. The maximum value of etching rate decreases approximately in two time. This fact is connected with the passivation of Si surface by oxygen atoms (Eq. (27)) which prevents the reaction of fluorine with silicon (Eq. (29)). On the contrary because the etching rate decreases the maximum value of fluorine concentration two-fold rises. It is explained by additional yield of fluorine atoms which cannot use in the etching reaction owing to the surface passivation. The deviation between the maxima of normalized etching rate and normalized fluorine concentration at α_s = 0.3 was less than 3% of O_2 in the feed gas. This is inconsistent with experimental data where more large difference was observed. But this difference rises gradually with α_s and it arrives in the range 10 - 15% of O_2 at α_s = 1. This is agree with the real experimental data (Mogab at al., 1978) although the operating regimes of reactor and its dimensions were slightly different from using in the present work. This also means that limiting case for ϑ_O (Eq. (35)) describes an essential contribution in competition process.

Probably a close sticking coefficients of oxygen and fluorine atoms when $\alpha_s \approx 1$ corresponds more realistic scenario of etching process. The displacement of maxima also affects on dependence of etching rate versus fluorine concentration (Fig.12). One can see from Fig.12 that with the rise of α_s the local maximum of etching rate shifts and becomes more lower

with simultaneous increasing maximal fluorine concentration. As a result the hysteresis curves become more circular. At the same time the beginning of curling of curve is gradually shifts to the region with small edition of O_2.

Generally these obtained results using complex kinetic model confirm the known experimental facts and qualitative scenario of silicon etching in CF_4/O_2 gas mixture. Particularly, the hysteresis effect is correctly reproduced. The oxygen addition up to 30 - 40% allows to increase the etching rate in several times and it is an effective factor for controlling the processing regime. The main channel of intensification of silicon etching in CF_4/O_2 gas mixture connects with a replacement of fluorine in CF_x radicals by oxygen atoms which set free the additional fluorine. The fluorine and oxygen chemisorption on silicon leads to hysteresis on the diagram of etching rate with respect to fluorine concentration. The influence of CF_2, CF_3 adsorption layers on the etching rate composes less 1% of nominal value. The variation of sticking coefficients of oxygen and fluorine atoms essential influence etching rate and hysteresis curve. With increase of its ratio α_s the maximum of average etching rate moves to the region with small edition of O_2. Thus, we have verified the very complex model of plasma etching reactor and have shown its opportunities.

3.5 Electron density effect on etching rate in plasma-chemical reactor

Electrical RF-discharges in fluorocarbon mixtures have been strongly studied for their effective employment in the production of semiconductor devices. Owing to different additions to the parent gas a variation of electron and ion concentrations in RF discharge may essentially affects on the main characteristics of etching process. In the case of Si - CF_4/O_2 system widely used in semiconductor production the addition of oxygen increases the spontaneous etching rate in several times and is an effective means for control of etching process as was shown above (see also experiments of Mogab at al., 1978). But at the same time the oxygen addition to tetrafluoromethane has a negative influence. According to the data of spectroscopic measurements (Mogab at al., 1978; Coburn & Chen, 1980; d'Agostino at al., 1980) the silicon etching process in CF_4/O_2 mixture accompanies by the electron concentration decrease with growing oxygen fraction when RF power and gas pressure being constants. This results in essential decrease of active particles generation at large oxygen contents. Further investigations (Coburn & Chen, 1980; d'Agostino at al., 1980) was shown that the influence of oxygen addition to CF_4 on electron density in RF-discharge is independent of construction dimensions of reacting chamber, operating regimes, electron energy (at least in the range 8-20eV) and has universal character.

Thus, the oxygen addition to CF_4 causes the contrary processes: on the one hand - the fluorine active particles production rises with simultaneously depletion of mixture by CF_x radicals. On the other hand - formation of oxygen chemisorption layer on silicon, decrease of electron concentration and dissociation efficiency of parent gas mixture also take place due to O_2 presence. Before our work the effect of electron concentration decrease in CF_4/O_2 mixture was not practically considered on the base of mathematical simulation.

The numerical results obtained illustrate an electron density influence on the main characteristics of silicon etching in dependence on oxygen concentration. In calculations it is compared the plasma-chemical reactors with uniform electron density distributions without oxygen $\bar{n}_e^0$ with $\bar{n} = \bar{n}_e^0 \times \vartheta(x_{O_2})$, where $\vartheta(x_{O_2})$ - the experimental dependence of average electron density on O_2 content (Mogab at al., 1978).
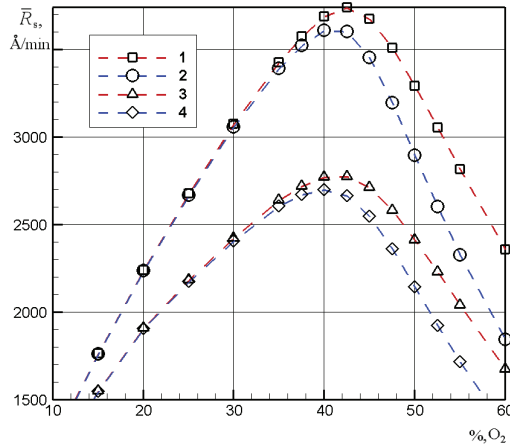
Fig. 13. Average etching rate vs percentage of O_2 in CF_4/O_2 mixture. Processing regime: $p = 0.5$ torr, $Q = 200$ cm³/min, $T_{w2} = 300$ K, "Inflow" direction. Parameters: $1 - \vartheta_e = 1$, $\alpha_s = 0.3$; $2 - \vartheta_e = 1$, $\alpha_s = 1$; $3 - \vartheta_e = \vartheta_e(x_{O_2})$, $\alpha_s = 0.3$; $4 - \vartheta_e = \vartheta_e(x_{O_2})$, $\alpha_s = 1$.

The oxygen addition allows to increase essentially the concentration of active particles and etching rate of silicon wafers. The dependence of average spontaneous etching rate on the percentage of O_2 concentration in CF_4/O_2 illustrates this ability (see Fig. 13, curves 1, 2). At the same time the electron concentration monotonically falls down with the rise oxygen addition to CF_4 under other constant conditions. This decreases electron impact dissociation and ionization rates and leads to the reduction of active particle production. Simultaneously the average etching rate goes down (see Fig. 13, curves 3, 4). On can see from this diagrams that the visible separation of curves begins in the region of maximum etching rates at the oxygen concentrations more then 30%.

For characteristic of decrease of average etching rate and dissociation efficiency of parent gas mixture the relative deviations were calculated:

$$\delta_{\overline{R}_s} = \left| \frac{\overline{R'}_s - \overline{R''}_s}{\overline{R'}_s} \right|, \delta_{\overline{C}_i} = \left| \frac{\overline{C'}_i - \overline{C''}_i}{\overline{C'}_i} \right|, i = F, CF_4, O_2.$$

Here $\overline{R'}_s$, $\overline{C'}_i$ are the average etching rate and fluorine, tetrafluoromethane and oxygen concentrations correspondingly at $\vartheta_e = 1$; $\overline{R''}_s$, $\overline{C''}_i$ are the average values with taking into account the dependence $\vartheta_e(x_{O_2})$.

Due to O_2 addition under other constant conditions an electron density monotonous decreases that causes the fall of fluorine concentration and etching rate. With the rise of percentage of O_2 content in CF_4/O_2 in the limits up to 60% the decrease of average etching rate can reach 30%. The quantitative deviations of average fluorine concentration at the wafer in the limits 0-40% O_2 addition coincide with the deviations of average etching rate. The last shows that in the limits etching rate dependence on fluorine concentration is practically linear and adsorption of CF_2, CF_3 particles on the wafer is weak. With taking into

account the dependence $\vartheta_e(x_{O_2})$ there are no qualitative differences in the average etching rate and fluorine concentration distributions as functions of O_2 addition. The location of the average etching rate maximum is preserved and is defined only by the intensity of oxygen chemisorption on silicon (parameter α_s) (Grigoryev & Gorobchuk, 2007). With increasing α_s the relative deviations of etching rate due to lowering electron density are preserved.

%, O_2	%, $\delta_{\overline{R}_s}$	%, $\delta_{\overline{C}_F}$	%, $\delta_{\overline{C}_{CF_4}}$	%, $\delta_{\overline{C}_{O_2}}$
10	8.95	8.95	2.05	25.98
20	14.79	14.74	8.41	42.07
30	21.22	21.07	22.25	58.75
40	24.88	24.44	44.15	66.49
50	26.66	26.86	68.82	74.99
60	28.90	29.66	81.61	84.58

Table 2. The values of $\delta_{\overline{R}_s}$, $\delta_{\overline{C}_F}$, $\delta_{\overline{C}_{CF_4}}$, $\delta_{\overline{C}_{O_2}}$ in dependence on O_2 addition.

The calculations of silicon etching process in CF_4/O_2 plasma including the experimental data on electron density decrease with oxygen addition allow to formulate the following conclusions. The fall of electron density in a RF-discharge with O_2 addition may decrease a spontaneous etching rate in the range up to 30%. At that the decrease in the range close to optimal etching rates arrived at 25 - 40% O_2 addition exceed 20%. Nevertheless the main positive effect of O_2 addition on the rise of etching rate in CF_4/O_2 mixture as compared with pure tetrafluoromethane preserves.

4. Some further perspectives

The perspectives of applications of created modelling tools in working up of new reactor schemes in the framework of silicon technology are obvious. Traditionally the parametric calculations will help design engineers in choosing the optimal configuration and operating parameters of the reactor.

The new technical challenges will require to model some new processes. For example, in present time the industry moves to production of largest single wafers with diameters up to 12 or even 16 inches (Japanese program). Under processing such large patterns the thermal and gravitational stresses will pose a strong limitation on the maximum wafer temperature and maximum temperature gradients near the wafer surface. This new problem will demand a very precise account of all wafer heating effects such as energetic ion bombardment, plasma radiation, heat generation by the exothermic reactions and glow discharge. It will be necessary to model in detail the real cooling system calculating thoroughly a problem of complex conjugate heat transfer. Furthermore it is need to include in the consideration the problems of an elasticity and thermoelasticity of a semiconductor wafer.

Besides we would like to draw attention to a new direction in mathematical modelling of chip manufacturing. In nearest future the typical fabrication-line for making complex chips

will be characterized by high level of programming automation, high quality operations and high yields. The new generations of micro- and nano-devices will demand a permanent (so called *in situ*) monitoring and control by their manufacturing in a real time (Orlikovsky at al., 2001). Monitoring data are to be operated and introduced in control system to support the optimal characteristics of process.

The relatively simple control system is based on supporting some goal parameter (parameters) on the optimal level using the controller. In this case the controller compares the current values of the goal parameters measured by monitoring system with ones from *apriori* calculated response functions. The calculating of response functions is a new problem of mathematical modelling.

The etching rate R_s can be considered as the goal parameter depending on the variable parameters such as flow rate Q, discharge power P, parent gas composition, etc. The calculated response functions are approximated by interpolation formulas, for example, by 2-dimensional cubic splines:

$$z^{(k)} = \sum_{i,j=1}^{M,N} \sum_{0 \leq m+n \leq 3} a_{ij}^{(m,n)}(x^{(l)} - x_i^{(l)})^m (x^{(r)} - x_j^{(r)})^n, \quad k,l,r = 1,2,...$$

where $\{z^{(k)}\}$ is the set of goal parameters, $\{x^{(l)}\}$ are the sets of variable parameters. These formulas are written in operating storage of controller. The data on the actual values $\{z^{(k)}(t)\}$ from *in situ* sensors are also transmitted there. The results of their comparison in the form of parameter disbalance $\Delta z^{(k)}(t) = z^{(k)}(t) - z^{(k)}$ are converted by the controller into the new set of optimizing variable parameters $\{x^{(l)}\}$. For example, the action of known proportional-integral - differential controller (Orlikovsky at al., 2001) is described by the following equation:

$$x^{(l)} = K\left[\Delta z^{(k)}(t) + \frac{1}{T_i}\int \Delta z^{(k)}(t)dt + T_d \frac{d\Delta z^{(k)}(t)}{dt} \right] \quad l = 1,2,...$$

The executor installs the new set of variable parameters on a reactor.

The most impressive new project in mathematical modelling of glow reactors is a creation of a virtual plasma chemical reactor. It is a comprehensive computational model of reactor processes that in first turn sets aside for the real time control of manufacturing process in flexible automation system. Such a composition is presented on Fig. 14 (Orlikovsky at al., 2001).

It is supposed that the computational model of virtual reactor will be organized in modular principle. It must contain a module 1 for calculations of electromagnetic (E - M) fields. The fields are describe by Maxwell - Lorentz system with taking into account a plasma bulk charge.

In the module 2 according to the given fields the electron density n_e and the energy T_e are calculated. In particular, in this module the Boltzmann kinetic equation for electron energy density distribution function (DDF) can be solved. Let us note that such an approach is a very laborious method which restricts the possibility to work in a real time. But here one can use so called hydrodynamic approach.

In the module 3 the gas flow structure and the temperature field in the reactor are computed. The module 4 carries out the calculations of plasma chemical kinetics and
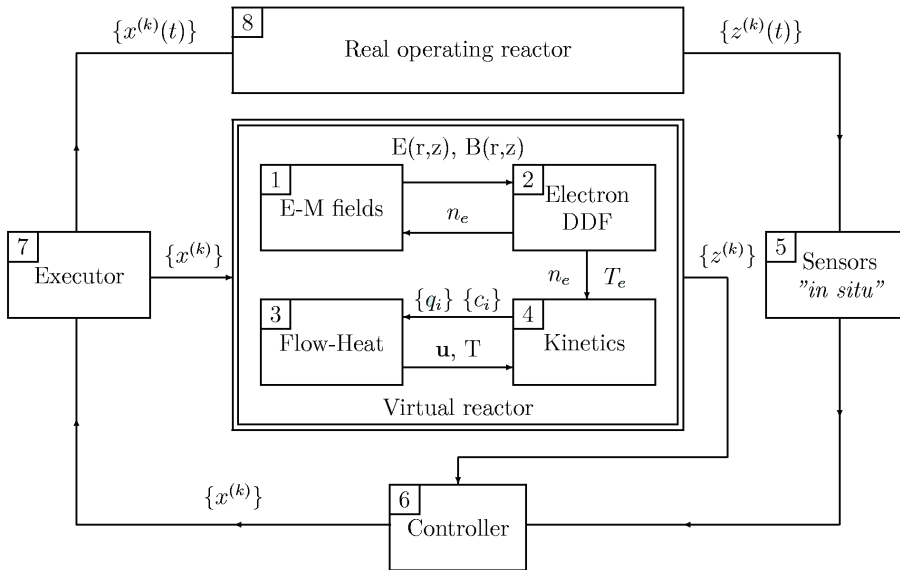
Fig. 14. Scheme of flexible automation system with virtual reactor

computes a mass transfer using of velocity and temperature fields. In turn, the data on gas component concentrations and thermal effects of chemical reactions are taken into account in the module 3. When using the virtual reactor in automation system the response functions $\{z^{(k)}\}$, calculated in real time with the actual process parameters, are transmitted in module 6 (controller). The data of *in situ* sensors $\{z^{(k)}(t)\}$, treated in the module 5, are also transmitted in this module. As it was said above the controller converts the results of their comparison in a set $\{x^{(l)}\}$ of optimizing variable parameters. The latest are simultaneously communicated both in the executor 7 and in the virtual reactor. In the virtual reactor the optimal goal parameters are calculated again. In such a way the feedback is closed.

As a comprehensive computational model of plasma chemical processing the virtual reactor can be used also by design engineers, in the calculations of more precise response functions for using in simple control systems described above, and as a learning automaton for personnel.

5. Acknowledgment

6. References

Aydil, E.S. & Economou, D.J. (1993) Modelling of plasma etching reactors including wafer heating effects. *Journal of the electrochemical society*, Vol.140, No.5. pp.1471-1481.

Coburn, J.W. & Chen, M. (1980) Optical emission spectroscopy of reactive plasmas: A method for correlating emission intensities to reactive particle density. *Journal of applied physics,* Vol.51, pp.3134-3136.

d'Agostino, R., Cramarossa, F., De Benedicts, S. & Ferraro, G. (1980) Spectroscopic diagnostics of CF_4-O_2 plasmas during Si and SiO_2 etching processes. *Journal of applied physics*, Vol.52. No.3. pp.1259−1265.

Dalvie, M., Jensen, K.F. & Graves D.B. (1986) Modelling of reactors for plasma processing. I. Silicon etching by CF_4 in a radial flow reactor. *Chemical engineering science*, Vol.41, No.4. pp.653-660.

Edelson, D. & Flamm, D.L. (1984) Computer simulation of a CF_4 plasma etching of silicon. *Journal of applied physics*, Vol.56. pp.1522-1531.

Edwards, D.K. & Menard, W.A. (1964) Comparison of models for correlation of total band absorption. *Applied optics*, Vol.3, No.5. pp.621-625.

Grigoryev, Yu.N. & Gorobchuk, A.G. (1996) Numerical optimization of planar reactors of individual plasma-chemical etching. *Surface*, No.2, pp.47-63. (in Russian)

Grigoryev, Yn.N. & Gorobchuk, A.G. (1997) Numerical simulation of plasma-chemical etching reactors, *Proceedings of 21st International conference on microelectronics MIEL'97*, pp. 485-488, Nis, Yugoslavia, 14-17 September 1997.

Grigoryev, Yn.N. & Gorobchuk, A.G. (1998) Nonisothermal effects in plasma-chemical etching reactor. *Microelectronics*, Vol.27, No.4. pp.294-303. (in Russian)

Grigoryev, Yu.N. & Gorobchuk, A.G. (2004) Numerical modeling of silicon etching in CF_4/O_2 plasma-chemical system, *Proceedings of 24th International conference on microelectronics MIEL'2004*, pp.475-478, Nis, Serbia and Montenegro, 16-19 May 2004.

Grigoryev, Yn.N. & Gorobchuk, A.G. (2007) Specific features of intensification of silicon etching in CF_4/O_2 plasma. *Russian Microelectronics*, Vol.36, No.5. pp.368-379, ISSN 1063-7397.

Grigoryev, Yu.N., Gorobchuk, A.G. (2008) Effect on electron density in RF-discharge on etching rate in plasma-chemical reactor, *Proceedings of International conferencee 2008 IEEE Region 8 International Conference on "Computational Technologies in Electrical and Electronics Engineering" SIBIRCON 2008*, pp.322-327, Novosibirsk, Russia, July 21-25, 2008.

Kopalidis, P.M. & Jorine J. (1993) Modeling and experimental studies of a reactive ion etcher using SF_6/O_2 chemistry. *Journal of the electrochemical society*, Vol.140, No.10. pp.3037-3045.

Mogab, C.J., Adams, A.C. & Flamm, D.L. (1978) Plasma etching of Si and SiO_2 - The effect of oxygen additions to CF_4 plasmas. *Journal of applied physics*, Vol.49. No.7. pp.3796-3803.

Orlikovsky, A.A., Rudenko, K.V. & Suhanov, Ya.N. (2001) Diagnostics in situ of plasma technological processes of microelectronics: present-day state and nearest perspectives. Part IV. *Microelectronics*, Vol.30, No.6. pp.403-433. (in Russian)

Plumb, I.C., & Ryan, K.R. (1986) A model of the chemical processes occurring in CF_4/O_2 discharges used in plasma etching. *Plasma Chemistry and Plasma Processing*, Vol.6, No.3. pp.205-230.

Sang-Kyu Park & Economou, D.J. (1991) A mathematical model for etching of silicon using CF$_4$ in a radial flow plasma reactor. *Journal of the electrochemical society*, Vol.138, No.5. pp.1499-1508.

Schoenborn, Ph., Patrick, R. & Baltes H.P. (1989) Numerical simulation of a CF$_4$/O$_2$ plasma and correlation with spectroscopic and etch rate data. *Journal of the electrochemical society*, Vol.136, No.1. pp.199-205.

Shokin, Yu.I., Grigoryev, Yu.N. & Gorobchuk, A.G. (1999) Advanced optimization of etching processes in radial flow plasma-chemical reactor, *Proceedings of 8th International symposium on Computational Fluid Dynamics ISCFD'99*, Bremen, Germany, 5-10 September 1999, ZARM University of Bremen, Bremen.

Venkatesan, S.P., Trachtenberg, I. & Edgar, Th.F. (1990) Modelling of silicon etching in CF$_4$/O$_2$ and CF$_4$/H$_2$ plasmas. *Journal of the electrochemical society*, Vol.137, No.7. pp.2280-2290.

Experimental Studies on Doped and Co-Doped ZnO Thin Films Prepared by RF Diode Sputtering

Krasimira Shtereva[1,2], Vladimir Tvarozek[2], Pavel Sutta[3],
Jaroslav Kovac[2] and Ivan Novotny[2]
[1]University of Rousse
[2]Slovak University of Technology in Bratislava
[3]New Technologies Research Center, West Bohemian University
[1]Bulgaria
[2]Slovakia
[3]Czech Republic

1. Introduction

For decades zinc oxide (ZnO) has been in the spotlight due to its unique combination of semiconductor, piezoelectric, optical, and magnetic properties, which open perspectives for wide range of applications from optoelectronic and transparent electronic devices (Ohta & Hosono, 2004), surface and bulk acoustic wave devices and piezoelectric transducers (Wang et al., 2008), spintronics (Ji et al., 2008), to chemical and gas sensors (Carotta et al., 2009), and solar cells (Ganguly et al., 2004). Great industrial advantages of ZnO are its eco-friendly nature, wide abundant sources and low costs of metal Zn.

ZnO is a group II-VI semiconductor with a direct band gap of 3.37 eV at room temperature, which can be modified (~3 eV– 4 eV) via extrinsic doping with either cadmium (Cd) or magnesium (Mg). By its semiconductor properties ZnO is similar to gallium nitride (GaN) (Table 1).

Wide band gap semiconductor	**ZnO**	GaN	ZnSe
Crystal structure	**wurtzite**	wurtzite	zinc-blende
Energy gap at RT, (E_g), (eV)	**3.37**	3.4	2.7
Exciton binding energy (meV)	**60**	21	20
Dielectric constant	**8.75**	9.5	7.1
Melting point (^{o}C)	**1975**	1700	1100
Lattice constant (Å) a-axis (Å) c-axis (Å)	**3.25** **5.21**	3.19 5.19	5.67 -

Table 1. Comparison of some basic properties of ZnO, GaN and ZnSe (Tüzemen & Gür, 2007)

Both materials have wurtzite crystal structure and very closed values of energy gap, lattice constants and thermal expansion coefficients. Therefore, ZnO can provide a high quality substrate for GaN.

Remarkable optical qualities (the width of band gap and large exciton binding energy of 60 meV), enable application for light emitting diodes (LEDs) and UV semiconductor lasers, which are devices with great commercial potential.

Besides, ZnO exhibits impressive stability of its properties when exposed to temperatures up to 700 K and radiation and hence, ZnO devices are suitable for space applications.

Resistance dependence on temperature, pressure and illumination can be utilized in temperature and pressure sensors and fire alarm.

Undisputable advantage of ZnO over GaN and ZnSe is the lowest price of single crystal wafers ZnO (1 side polish, 10x10 mm), fabricate from Wafer World Inc. (Fig. 1).
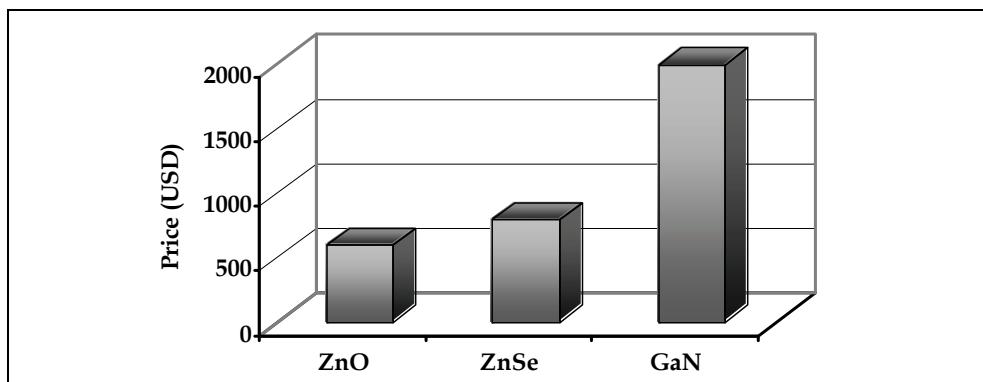
Fig. 1. Comparison between the price of ZnO, GaN and ZnSe wafers
(http://www.waferworld.com/)

Zinc oxide can provide an alternative to other transparent conductive oxides (TCOs), (tin–doped indium oxide (In_2O_3:Sn) (ITO) and antimony–doped tin oxide (SnO_2:Sb)), which are widely used as transparent electrodes for liquid crystal displays (LCDs), organic light-emitting diodes (OLEDs), and in photovoltaic solar cells. ZnO has higher conductivity than tin oxide and is readily etchable than ITO. Furthermore, zinc is an inexpensive and abundant metal, whereas indium is a rare metal, and parallel to the growing market for flat panel displays grow worries over its depletion and stable supplies.

The remarkable properties of zinc oxide along with: (i) presence of high quality single crystals; (ii) ability to grow ZnO thin films; (iii) ability to dope ZnO and this way to modify its physical properties; (iv) ability to metalize for ohmic contacts and interconnections, make ZnO one of the most promising semiconductor materials of the new millennium.

Our work has been motivated by the optoelectronic applications of ZnO and the need of both, high quality *p*-type and *n*-type materials to realize active devices. In this chapter we will present our research on the growing and characterizing of *p*-type ZnO thin films, prepared by radio frequency (RF) diode sputtering, mono-doped with nitrogen (ZnO:N), and co-doped with aluminium and nitrogen (ZnO:Al:N). Important parameters such as crystallite size (D), strain (ε), texture orientation, type and carrier concentration (n/p), carrier mobility (μ) and resistivity (ρ), have to be determined when assessing ZnO quality. In the

following sections the reader will find experimental results and analysis of these micro structural and electrical parameters, and a discussion over their interconnection. We do believe that this research contributes to the better understanding of doping behavior and *p*-type conductivity in ZnO.

2. Doping ZnO

Doping is fundamental to controlling the properties of the semiconductors and to obtaining of new multifunctional materials. There is a wide variety of well developed doping methods, including diffusion, ion implantation, and in situ doping with epitaxial growth. The pure elemental semiconductors, such as silicon (Si) or germanium (Ge), tend to be electrically intrinsic. They can be readily *n*- or *p*-type doped using extrinsic dopants, such as boron (B) or phosphorous (P), therefore, they are denoted symmetrically doped. The doping and re-doping of the same silicon crystal from *n*- to *p*-type and back again is very common during the device fabrication. On the contrary, the wide band gap semiconductors, including ZnO, are denoted asymmetrically doped, i.e., they can be doped easily either *n*-type or *p*-type but not both. This asymmetry has been explained with (i) the low or limited solubility of the desirable dopants; (ii) the high activation energies of these dopants, (iii) the tendency to form spontaneously compensating defects, and (iv) hydrogen acting as an unintentional extrinsic donor.

Zinc oxide exhibits asymmetry in doping and *n*-type conducting ZnO can be readily obtained via either native or extrinsic donor defects whereas *p*-type doping turned out to be more difficult. Density functional theory identifies as donors two native point defects in ZnO structure, the oxygen vacancy (V_O) and the zinc interstitial (Zn_i) (Look et al., 2003). One of them, Zn interstitial, is a shallow donor, whereas O vacancy is a deep donor. The cation vacancies are expected to create acceptor levels in most of group II-VI semiconductors and hence Zn vacancy (V_{Zn}), a part of Zn_i- V_{Zn} Frenkel pair, is an important acceptor in ZnO. Since Zn_i is expected to have high formation energy it puts the question whether (only) this defect is responsible for the donors' concentration in un-doped *n*-type material. Theoretical studies of Van de Walle found hydrogen is a shallow donor in ZnO (Van de Walle, 2000).

2.1 ZnO n-type doping

Conductive ($\rho \sim 10^{-4}$ Ωcm) and transparent ($T > 80$ %) *n*-type ZnO thin films have significant commercial impact due to their use as transparent electrodes for flat panel displays, organic light emitting devices and in solar cells. They can be obtained via doping of ZnO with either group III elements, aluminum (Al), gallium (Ga) or indium (In), as a substitution for Zn, or group VII elements, chlorine (Cl) and iodine (I) as a substitution for O, in the crystal lattice. Almost metal conductivity ($\rho = 4.6 \times 10^{-4}$ Ωcm, sheet resistance $R_S = 32$ Ω/) was reported for aluminum - doped ZnO films (ZnO:Al), deposited by middle-frequency (MF) alternative magnetron sputtering (Fu et al., 2004), and DC magnetron sputtering ($\rho = 9 \times 10^{-4}$ Ωcm) (Guillén & Herrero, 2006). Their low resistivity was combined with high visible transparency (85 – 90 %). By changing the sputter pressure and the film thickness were modified physical properties and obtained Ga-doped films (ZnO:Ga) with resistivity of 2.6×10^{-4} Ωcm, Hall mobility of 18 cm^2/Vs and the carriers concentration of 1.3×10^{21} cm^{-3} (Assunção et al., 2003, Fortunato et al., 2004). It was found that mobility and grain size were linearly dependent, which resulted in high mobility in films with larger crystalline grain size. Optical transmittance (80 – 90 %) depends on both, thickness and sputtering pressure

and its changes are consistent with the changes of the electrical parameters. Other deposition parameter with major influence on the film's properties is the substrate temperature (T_S). Textured ZnO:Ga films with a strong preferential orientation can be deposited only at optimum temperature (Song et al., 2002). ZnO thin films doped with In, Ga and Al were utilized in amorphous silicon solar cells (Nunes et al., 2002). The best parameters showed the cells with ZnO:In films.

Rare-earth elements yttrium (Y) and scandium (Sc) make efficient donors for producing an n-type ZnO material with low resistivity (3.1×10^{-4} Ωcm) (Minami et al., 2000). Conductivity of ZnO:Sc was higher than that of ZnO:Y and its electrical and optical properties, as well as the thermal stability of resistivity were comparable to those of ZnO:Al.

2.2 ZnO p-type doping and co-doping

ZnO can be doped p-type intrinsically, through native defects, and extrinsically through impurity doping.

As discussed above, self-compensation by native donor defects and unintentionally introduced impurities (H), prevent the achievement of p-type conductivity in un-doped ZnO. Intrinsic n-type was changed to moderate p-type ZnO (ρ = 1250 Ωcm, μ = 30 cm²/Vs and $p = 10^9$ cm⁻³), by adjusting the oxygen/argon ratio in the sputtering plasma (Xiong et al., 2002). The authors explained p–type conductivity in un-doped ZnO with the increase in the chemical potential of atomic oxygen that lowers the formation energy of the acceptor defects.

On the other hand, the successful extrinsic p-type doping is restricted by low solubility and high activation energy of the acceptor dopants in addition to above mentioned compensating species.

Group I elements, such as lithium (Li), as substitution elements for Zn in ZnO crystal lattice, are regarded as suitable acceptors. They create shallow acceptor levels but due to smaller atomic radius than this of the replaced atom, they tend to locate on the interstitial sites and hence, to act as donors. Besides, the larger bond than the Zn–O bond induces the lattice strain that will lower the formation energy of the native donor defect, e.g. O vacancy, which will compensate the dopants (Özgür et al., 2005). Investigations of p-type lithium - doped ZnO thin films (ZnO:Li) showed that p-type doping by Li is limited from the formation of Li_{Zn} – Li_i complexes (Zeng et al., 2005). The formation of these complexes can be suppressed by adjusting of the Li content. It was found that low resistivity p-type ZnO:Li films (ρ = 17 Ωcm, μ = 3.47 cm²/Vs and $p = 1.01 \times 10^{17}$ cm-3), could be obtained at a low Li content and optimum temperature. Resistivity of the films, grown at a high Li content, was high and their absorption edge blue shifted.

Group V dopants, such as nitrogen (N), phosphorus (P) and arsenic (As), which substituting for oxygen in the ZnO crystal lattice, are acceptors as well. Quality p-type material was produced by P and As doping, even though their ionic radii are larger than that of O. ZnO n-type by arsenic doping (ZnO:As) ($n \sim 10^{18}$ cm⁻³), turned to p-type ($p \sim 10^{19}$ cm⁻³, μ_H = 4.02 cm²/Vs, ρ = 4.4x10⁻² Ωcm) after heat treatment (Ryu et al., 2000, Moon et al., 2005). The phosphorus - doped p-type ZnO (ZnO:P) films exhibit a hole concentration of 5×10^{18} cm⁻², mobility of 2 cm²/Vs and resistivity of 2 Ωcm (Hwang et al., 2005).

2.2.1 p-type ZnO by nitrogen-doping

Nitrogen ionic radius is almost the same as that of O and therefore, should match well on the O site. Furthermore, its highest solubility between group V elements and shallower

acceptor levels than phosphorus and arsenic, make nitrogen the preferable candidate for *p*-type doping between group V elements. Promises for successful *p*-type doping are also *p*-type conduction in nitrogen doped ZnSe and high concentration of N_O acceptor in N-doped ZnO (ZnO:N) thin films grown by plasma–assisted molecular beam epitaxy (P-MBE) (Sun et al., 2006).

Preparation of *p*-type ZnO thin films by nitrogen doping, using different deposition techniques and nitrogen sources, was reported by a number of groups. Electrical parameters of sputtered *p*-type ZnO:N thin films are compared in Table 2.

Reference	Resistivity (Ωcm)	Mobility (cm^2/Vs)	Concentration (cm^{-3})
(Lu et al., 2003)	4×10^3	2.4	4.6×10^{15}
	31	1.3	7.3×10^{17}
	760	1.9	2.4×10^{16}
(Wang et al., 2006)	339	0.08	2.4×10^{17}
	8.34	0.1	7.5×10^{18}
(Yao et al., 2007)	456	0.1	1.2×10^{17}
(Ye et al., 2004)	1200	84.9	6.10^{13}
(Lu et al., 2005)	3.4×10^4	8.3	2.18×10^{13}
(Tvarozek et al., 2008)	**790**	**25**	**3.2×10^{14}**

Table 2. Electrical properties of sputtered *p*-type ZnO:N

Experimental studies suggest a positive role of the acceptor-hydrogen complex (N_O-H), for suppressing the formation of compensating interstitial defects and for providing *p*-type doping (Lu et al., 2007). Hence, films grown in ammonia (NH_3) – O_2 ambient were *p*-type with a hole concentration of 7.3×10^{17} cm^{-3} (Lu et al., 2003). At higher and lower ammonia concentration, the carrier concentration reduces and mobility goes up due to the oxygen deficiency in the growth ambient, which creates large amount of donor defects, such as Zn interstitials and O vacancies. Therefore, it is difficult to obtain carrier concentrations higher than ~ 10^{-17} cm^{-3} even when the film grows in pure N_2 (Yao et al., 2007). The temperature plays an important role in the activation of the N_O acceptor and annealing can cause an increase in the hole concentrations (Wang et al., 2006). Indeed, mobility of the annealed samples is higher than of as-deposited ones due to the large grain size and the strong c-axis crystalline structure. The XRD patterns of un-doped ZnO thin films show a strong (002) diffraction line, indicating a c-axis preferential orientation. Incorporation of nitrogen caused the randomization of the crystallites orientation, which is manifested by appearance of additional lines ((100), (101) and (110)), besides the main (002) line, in the XRD patterns of ZnO:N films (Zhao et al., 2005). The visible transmittace of *p*-type ZnO:N (~ 90 %) is almost the same as for Al -, Ga -, or In – doped ZnO thin films (Lu et al., 2003). The sharp absorption edge appears around 389 nm and shifts towards shorted wavelengths (blue shift) which corresponds to an increase in the carrier concentration.

2.2.2 p-type ZnO by aluminum-nitrogen co-doping

Theory predicted that co-doping of ZnO simultaneously with acceptors (nitrogen) and donors (Al, Ga or In), enhances incorporation of the N_O acceptors and supports formation of the shallow N_O acceptor levels. Hence, co-doping provides better perspectives for obtaining of a *p*-type ZnO material than nitrogen mono-doping.

The electrical parameters of p-type N - Al co-doped ZnO thin films are compared in Table 3.

Reference	Resistivity (Ωcm)	Mobility (cm^2/Vs)	Concentration (cm^{-3})
(Ye et al., 2005)	157	0.0711	5.59×10^{17}
(Ye et al., 2004)	1.6×10^2	0.3	1.1×10^{17}
(Yuan et al., 2004)	24.5	0.34	7.5×10^{17}
(Lu et al., 2005)	3.4×10^4	8.3	2.18×10^{13}
(Ge et al., 2004)	78	0.1	8×10^{17}
(Zhu et al., 2005)	101	0.05	1.3×10^{18}
	160	0.3	1.1×10^{17}
(Shtereva et al., 2008)	**21**	**0.4**	**7.8×10^{17}**

Table 3. Electrical properties of sputtered *p*-type ZnO:N:Al

Indeed, the obtained hole concentration of *p*-type ZnO:N:Al (1.1×10^{17} cm^{-3}) is more than two orders of magnitude higher than achieved for *p*-type ZnO:N (6.7×10^{14} cm^{-3}) prepared under the same deposition conditions, which gives an evidence for effectiveness of the co-doping concept (Ye et al., 2004). Estimation of nitrogen incorporation by means of SIMS depth profiles shows the same trend of variation of the N and Al content in the film. Deposition parameters that influence p-type doping and the physical properties of ZnO:N:Al are growth ambient and temperature (Lu et al., 2005), and oxygen partial pressure (Ye et al., 2005). In fact, there is optimum temperature or pressure, a kind of temperature/pressure "window" where *p*-type ZnO material can be grown (Yuan et al., 2004). Formation of the compensating donor defects, e.g. nitrogen molecules on the oxygen site, $(N_2)_O$, can be suppressed by adjustment of the growth temperature. The higher hole concentration of *p*-type ZnO:N:Al films prepared using NH$_3$ (1.3×10^{18} cm^{-3}), compared to this of films doped from N$_2$O source (1.1×10^{17} cm^{-3}), is ascribed to hydrogen incorporation into the film together with aluminum and nitrogen that suppresses the formation of the donor defects. AlN defects act as scattering centers and cause low mobility.

The ZnO:Al:N thin films grow with a c-axis preferential orientation and the (002) diffraction line has maximum intensity independently of the substrate temperature or the dopant concentration. They exhibit high visible transparency of about 90%. A decrease in optical band gap at co-doped *p*-type ZnO was observed (Ge et al., 2004).

3. Development and characterization of nitrogen-doped and aluminium-nitrogen co-doped ZnO thin films: Results and discussions

Our previous studies on un-doped ZnO films (chemo-resistive films for gas sensing and piezoelectric high resistive films for SAW sensors and micro-actuators), laid the groundwork for current experiments (Tvarozek et al., 2007). The experiments of Yao et al., mentioned in the previous part (Yao et al., 2007), provided evidence of successful *p*-doping using N$_2$ source. Co-doping simultaneously with nitrogen and group III element should be a more effective strategy to achieve a low resistivity p-type ZnO material than nitrogen mono-doping (Yamamoto, 2002). This concept was verified by many research groups and discussed in the co-doping section of this chapter. In the following subsections we shall describe our experimental results and shall discuss the effects of nitrogen doping and aluminum – nitrogen co-doping on structural and electrical properties of ZnO thin films

prepared by RF diode sputtering. Moreover, structural and electrical features of ZnO films grown on both Corning glass and Si/SiO$_2$ substrate under the same and different growth conditions will be compared.

3.1 Thin films deposition

The nitrogen-doped and the aluminum - nitrogen co-doped (ZnO:Al:N) thin films, discussed in this chapter, were deposited in a planar radio frequency (RF) sputtering diode system Perkin Elmer 2400/8L. Using of RF sputtering in reactive plasma is a new approach to preparation of ZnO thin films of required properties. In RF diode sputtering a flux of charged and neutral particles interacts with the growing film. This energetic particles bombardment increases the substrate/film temperature (T_S) and reduces the formation energy of the nitrogen acceptor providing conditions for effective p-type doping of the ZnO thin films. The total energy density (E_Φ) of the flux significantly affects and modifies the crystalline structure and hence the electrical and the optical properties of the RF sputtered ZnO thin films. The N$_2$ dopant source, chosen among the other nitrogen sources, is an easy getting, economic lucrative and non-toxic. The ZnO:N thin films were deposited on Corning glass substrates or on n-type Si (100) wafers covered by thermal Si oxide of thickness 0.8 µm, from a ZnO ceramic target (purity 99.99%), in Ar/N$_2$ working gases. The diameter of the ZnO target was 203.2 mm. A sintered ceramic target ZnO:Al$_2$O$_3$ (98wt%:2wt%), a mixture of ZnO (purity 99.99%) and Al$_2$O$_3$ (purity 99.99%), in diameter 152.4 mm, was used for the deposition of ZnO:Al:N films. They were deposited on Corning glass substrates in Ar/N$_2$ working gases. In both cases, the vacuum chamber was evacuated to a base pressure of 2x10^{-5} Pa before admission of gases of purity: Ar (99.999%) and N$_2$ (99.999%). The working gas pressure of 1.3 Pa and the sputtering power (500 W for ZnO:N and of 417 W for ZnO:Al:N films), were maintained constant during deposition. Varying deposition parameters were percentage of nitrogen (0 % ÷ 100 %) in the sputtering gas, and the bias voltage applied on the substrate (-25 V, -50 V and -100 V). Depending on the deposition time that varies from 30 to 60 minutes, the thickness of the ZnO films ranged from 430 nm to 870 nm. The film thickness was evaluated by a Talystep instrument. The substrate temperature T_s = room temperature and the ratio E_Φ / $E_{\Phi min} \geq 7$.

3.2 Micro structural parameters of nitrogen-doped and Al:N$_2$ co-doped ZnO thin films

X-ray diffraction provides general purpose qualitative and quantitative information on the composition and structure of the studied material and was used for qualitative structure analysis of nitrogen mono-doped and Al:N$_2$ co-doped ZnO films. Diffraction line contains a lot of information of which four parameters are of special interest: (i) peak position, (ii) FWHM (full width at half maximum), (iii) intensity of its maximum and (iiii) integrated intensity (area below the diffraction line). These parameters were used to identify the contents of the sample, as well as to evaluate important material parameters such as crystallite size (D), crystallinity, stress and strain ($<\varepsilon>$).

X-ray diffraction patterns of the ZnO:N thin films were recorded on an AXS Bruker D8 powder diffractometer (symmetric Θ -2Θ geometry and pseudo-parallel beam), equipped by 2D detector and an Eulerian cradle. CoKα radiation (λ = 0.179 nm) was used. A method proposed by Langford, based on the X-ray diffraction line profile analysis, was used to perform the size-strain analysis (micro-strains and crystallite sizes).

The microstructure of ZnO:Al:N thin films was analyzed on a thin film attachment (asymmetric ω-2Θ geometry and pseudo-parallel beam) and on an X'pert Pro powder diffractometer at a constant incident angle ω of 1˚ with an observation 2Θ range from 20 to 40˚. Cu K_α (λ = 0,154 nm) radiation was used. In both cases, a ceramic alumina from NIST (National Institute of Standards and Technology) was used as an instrumental standard (Fig. 2).
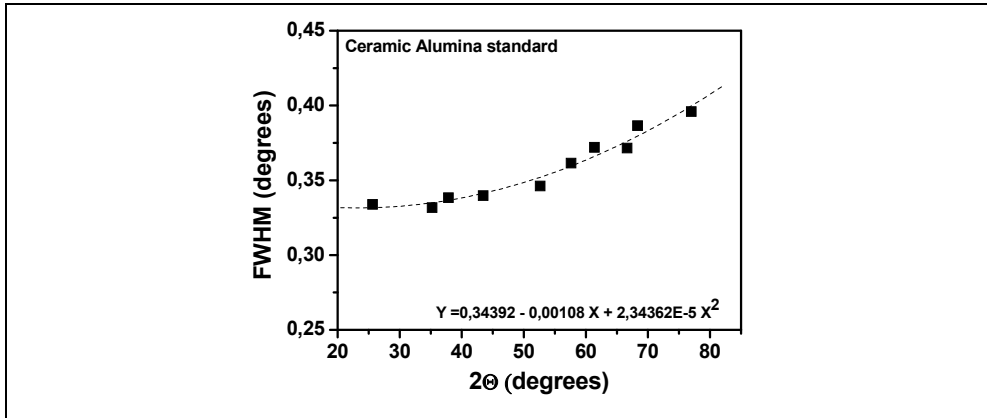
Fig. 2. Dependence of FWHM of Ceramic Alumina standard on 2Θ

The diffraction indices for the crystal planes of pure ZnO, their 2Θ (CuKα) values, d-spacings and relative intensities calculated by APX 63 - Struc software are presented in Table 4.

Plane (hkl)	d-Spacing (nm)	2θ (°)	Relative Intensity (%)
(100)	0.281450	37.088	48
(002)	0.260330	40.221	36
(101)	0.247591	42.388	100
(110)	0.162495	66.851	28
(103)	0.147725	74.591	30
(004)	0.130165	86.892	2

Table 4. XRD reference data for hexagonal ZnO calculated by using APX 63 - Struc software (Kraus, I., 1993)

3.2.1 ZnO:N thin films

Structural parameters of ZnO:N films, deposited on Corning glass and Si/SiO$_2$ substrates, were examined as a function of two deposition variables: the N$_2$ content in the working gas and the negative bias voltage applied on the substrate (Table 5). It was found out that both variables influence the film crystalline quality.

The XRD patterns presented in Fig. 3, from the observed range (30 °< 2Θ <80 °), show the diffraction lines of ZnO:N thin films deposited on Corning glass 7059 substrates. The plots show the results for a different content of N$_2$ in the sputtering gas under 0 V bias.

N_2 content (%)	Mode	Substrate	D (nm)	$<\varepsilon>$ (-)
0	sputter	Corning 7059	37	1.0×10^{-2}
10	sputter	Corning 7059	140	1.6×10^{-2}
25	sputter	Corning 7059	20	1.6×10^{-2}
50	sputter	Corning 7059	12	1.3×10^{-2}
75	sputter	Corning 7059	8	1.2×10^{-2}
100	sputter	Corning 7059	24	1.2×10^{-2}
75	bias -25 V	Corning 7059	11	1.26×10^{-2}
75	bias -50 V	Corning 7059	11	0.75×10^{-2}
25	sputter	Si/SiO$_2$	12	9.57×10^{-3}
50	sputter	Si/SiO$_2$	-	-
75	sputter	Si/SiO$_2$	37	13.2×10^{-3}

Table 5. Crystallite size and strain as a function of nitrogen percentage in the working gas and the negative bias voltage

Three dominant lines (100), (002), and (101) appear in the XRD patterns of all ZnO films. The forth diffraction line (110), is available only in the patterns of the films deposited at 25, 50 and 75 % N_2 in the sputtering gas. The XRD patterns clear show the change in the growth direction with increasing nitrogen content. A strong (002) line at $2\Theta \sim 40°$, and weak (100) and (101) features at $2\Theta \sim 36.9°$, 42°, for un-doped (0 % N_2) films and those grown with 10 and 100 % N_2, provide evidence for their polycrystalline hexagonal structure with a preferential c-axis orientation, perpendicular to the substrate. As nitrogen content increases (25, 50 and 75 %), the (002) peak intensity decreases while its full width at half maximum (FWHM) increases. The diffraction lines corresponding to (100), (101) and (110) crystal planes rise, showing that the ZnO:N thin films become more randomly orientated.

Fig. 3. X–ray diffraction patterns of ZnO:N deposited on Corning glass 7059 substrates as a function of nitrogen percentage in the working gas

The 2Θ diffraction angle values are slightly smaller than the standard values given in Table 4 for all three dominant lines (100), (002), and (101) respectively. Depending on N_2 content, their position shifts 0.2° to 0.6° lower for nitrogen - doped ZnO films compared to the standard values (Table 4) for un-doped ZnO. This shift can be explained by compressive

lattice strains (stresses) created during sputtering process and can be quantitatively evaluated from the equation for biaxial lattice stress (Šutta & Jackuliak, 1998)

$$\sigma_1 + \sigma_2 = -\frac{E}{\mu} \cdot \frac{d - d_0}{d_0} \qquad (1)$$

where E is Young's modulus, μ is Poisson's ratio, d_0 is the reference strain-free interplanar spacing and d the interplanar spacing obtained from the experiment. According the Bragg law the decrease of the diffraction angle with increasing N_2 content will increase the interplanar spacing d thus introducing stress into the film. Therefore, the shift of the position of (002) line toward the lower angles is an indicator of tensile stress in the ZnO thin films.

The estimated crystallite size ranges from 8 to 140 nm with varying N_2 content in working gas. Crystallite size is a measure of the size of a coherently diffracting domain and can be estimated using a Scherrer's formula:

$$<D> = \frac{K\lambda}{\beta_C^f \cos\Theta} \qquad (2)$$

where $K = 2\sqrt{\ln 2/\pi} \approx 0.94$ is the Scherrer's constant, λ is the wavelength of the X-rays used, β_C^f is the pure (physical) Cauchy component of integral breadth of the line taken in radians and Θ is the Bragg's angle (Delhez et al., 1982). Average micro-strain is about 1×10^{-2} and was determined using the equation:

$$<\varepsilon> = \frac{\beta_G^f}{4 \mathrm{tg}\Theta} \qquad (3)$$

where β_G^f is the pure (physical) Gaussian component of integral breadth of the line taken in radians (Delhez et al., 1982). Some researchers have interpreted the strain in nitrogen-doped thin films and structural deformations in terms of the increase of complex defect density in the material (Park et al., 2008). It was observed that ZnO:N films grown at low temperatures (< 500°C), suffer from high residual tensile stress. One reason for this stress can be the effect of the film thickness and the other is the incorporation of nitrogen. Since Zn-N bond length (2.04 Å) is somehow longer than the length of Zn-O bond (1.93 Å), it is expected that nitrogen incorporation will cause lattice expansion.

It is commonly accepted that substrate temperature has a significant effect on physical properties of the film. Since a negative bias voltage applied on the substrate causes an intensive positive ions bombardment on the substrate, it will increase the substrate temperature. The XRD patterns for ZnO:N thin films deposited at three different negative bias voltages, 0, - 25 and - 50 V, applied on the substrate, and at a constant N_2 content of 75%, are displayed on a Fig. 4. The crystalline structure of the ZnO:N thin films deposited without bias and at bias voltage of – 25 V is random orientated with a dominant (100) feature at $2\Theta \sim 36.7°$, while that of the films deposited at – 50 V bias gives well defined (002) diffraction line at $2\Theta \sim 39.7°$ and two weaker (100) and (101) diffraction lines, indicating preferential orientation growth. The narrowing of the (002) and the widening of (100) and (101) diffraction lines are clear observable. The 2Θ diffraction angle positions remain uninfluenced from the negative bias voltage. The crystallite size and average microstrain are slightly influenced from the increase of the bias voltage.
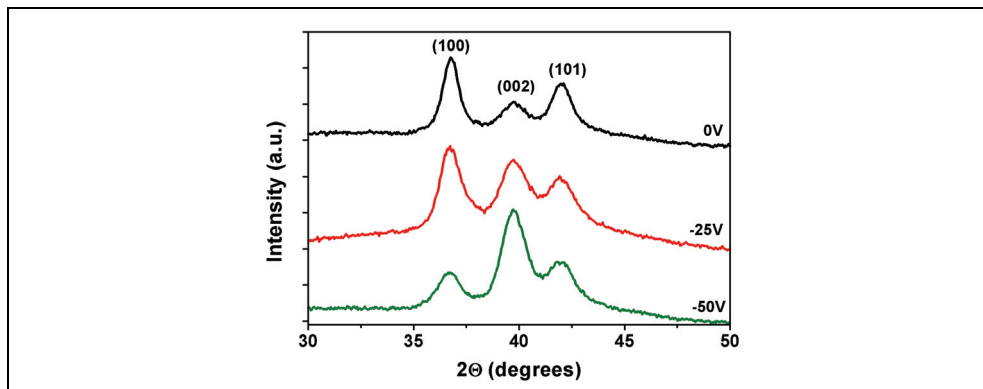
Fig. 4. XRD patterns of ZnO:N thin films deposited on Corning glass 7059 substrates at 75 % N_2 content as a function of the bias voltages

The XRD patterns presented in Fig. 5, recorded for 30° < 2Θ < 80 °, show the diffraction lines for ZnO:N thin films deposited on Si/SiO$_2$ substrates at different contents of N_2 in the sputtering gas. The XRD patterns reveal four dominant lines at 2Θ diffraction angles of approximately 36.2 °, 39.4 °, 41.46 °, and 65.6 °, which correspond to (100), (002), (101), and (110) planes of the hexagonal ZnO structure respectively.
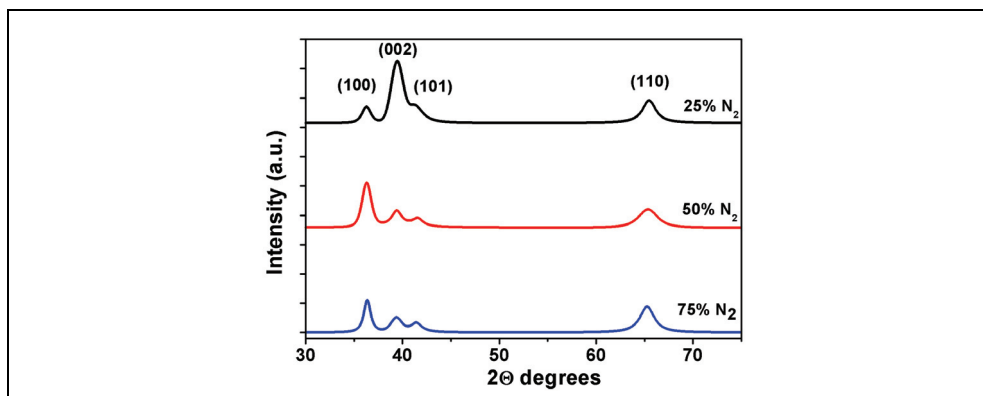
Fig. 5. X–ray diffraction patterns of ZnO:N deposited on Si/SiO$_2$ substrates as a function of nitrogen percentage in the working gas

Diffraction lines from (103) and (004) crystal planes listed in Table 4 are missing, and the intensities from detected lines do not match the listed relative intensities. In fact, the variation tendencies in the diffraction angle positions and the intensities with N_2 content are the same as for the films deposited on a glass substrate.

The dominant (002) diffraction line of ZnO thin films deposited at 25 % N_2 in the sputtering gas indicates a good crystalline quality with a preferential c-axis orientation, perpendicular to the substrate. As nitrogen percentage in the working gas increases (50 %, 75 % N_2), the intensity of (002) line decreases and its width increases. In parallel, the rise of (100) and (101) diffraction lines is observed indicating that the crystallites become randomly oriented.

The 2Θ diffraction angles are smaller than the standard values given in a table 4 for all four dominant lines (100), (002), (101), and (110) respectively. They shift from 0.8° to 1° lower in ZnO:N compared to the standard values (Table 4) for un-doped ZnO.

The grain size (12 to 37 nm) increases and average microstrains decrease (0.97×10^{-2} to 1.32×10^{-2}) with increasing N_2 content in the sputtering gas.

XRD patterns of ZnO:N deposited at 75 % N_2 on Si/SiO$_2$ and Corning glass 7059 substrates are compared in Fig. 6.
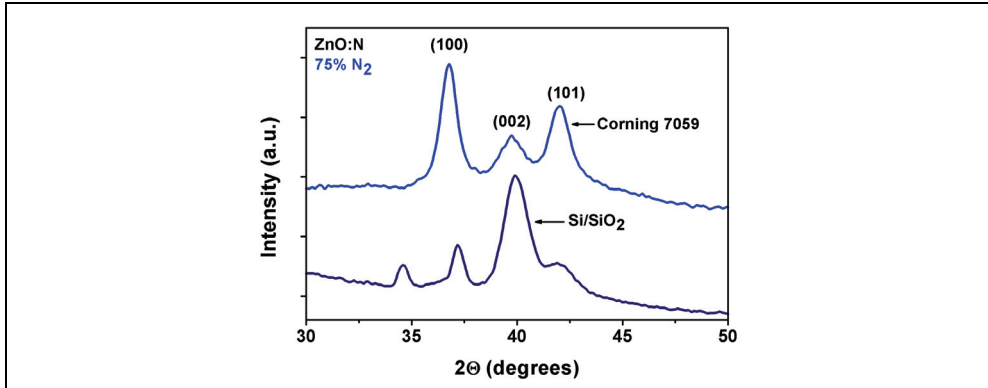
Fig. 6. X–ray diffraction patterns of ZnO:N deposited on Corning glass 7059 and Si/SiO$_2$ substrates at 75 % N_2 content in the working gas

3.2.2 ZnO:Al:N thin films

Structural properties of aluminum-nitrogen co-doped ZnO films, sputtered on Corning glass substrate were investigated as a function of nitrogen percentage in the sputtering Ar/N_2 gas mixture. X-ray diffraction show that co-doping improves the crystalline structure and all ZnO:Al:N thin films show a c-axis texture in direction declined about 16 deg from the surface normal. The XRD patterns of ZnO:Al:N thin films deposited at four nitrogen contents (0, 25, 50 and 75 % N_2), are recorded for 30° < 2Θ < 50 °(Fig. 7). They reveal more or
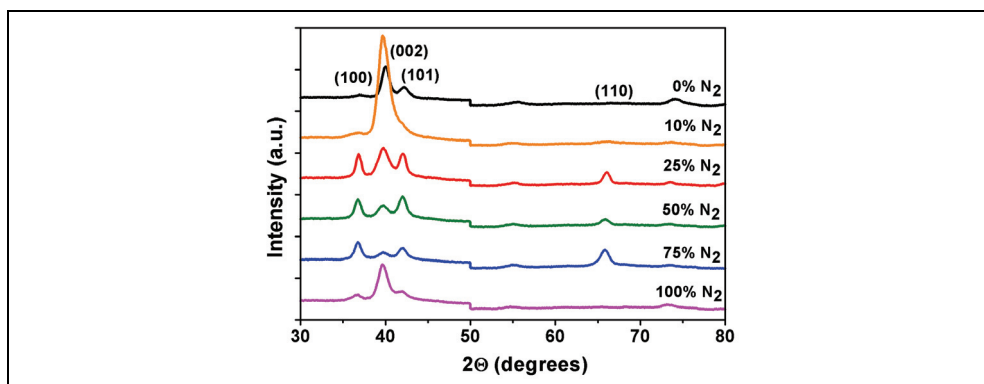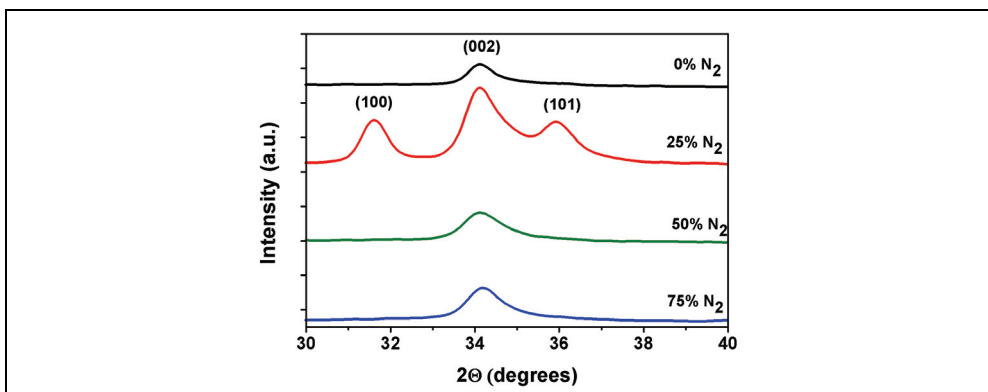

Fig. 7. X–ray diffraction patterns of ZnO:Al:N deposited on Corning glass 7059 substrates as a function of nitrogen percentage in the working gas

less stronger expressed c-axis preferential orientation of the films depending on the N_2 content in the sputtering gas. (100), (002) and (101) diffraction lines appear in the XRD patterns of the films deposited at 25 % N_2, indicating that their preferential orientation is more random. A strong preferential orientation of the crystallites is observed for high dopant contents (50 % and 75 % N_2) and the intensities of the (002) diffraction lines are almost the same. The diffraction lines are asymmetric more likely due to the incorporation of the Al and the formation of different phases (AlO) at the interface, which causes broadening of the line toward the higher diffraction angles. The estimated grain size changes from 21 to 33 nm and microstrains vary from 4.5×10^{-3} to 9.6×10^{-3} with the N_2 content.

3.3 Electrical characterization of doped and co-doped ZnO thin films

Hall-effect measurements are widely used technique to determine the electrical properties, and to evaluate the quality of the semiconductor materials. This technique allows to measure directly the carrier type and concentration. Other attractive properties, which make this characterization technique so popular, are its cost, simplicity and ease of use, although a special sample geometry is required and measurement is sensitive to the contact. Contemporary semiconductor physics acknowledges carrier concentration (n/p) and carrier mobility (μ_n/μ_p), fundamental electrical parameters, while resistivity is related to those two parameters. The Hall coefficient (R_H) and resistivity (ρ), are determined experimentally. The hole concentration is a function of the Hall constant and is given by

$$p = \frac{1}{qR_H} \tag{4}$$

Similarly for an n-type semiconductor

$$n = -\frac{1}{qR_H} \tag{5}$$

For known resistivity ρ, the carrier drift mobility is evaluated using the formula:

$$\mu = \frac{|R_H|}{\rho} \tag{6}$$

The Hall-effect measurements of ZnO:N and ZnO:Al:N thin films were carried out at room temperature (RT). The measured samples have geometry of a 1x1 cm square. The ohmic contacts were made by small indium dots at the four corners of square samples, so that, their average diameter and also the film thickness were significantly smaller than the distance between the contacts. In order to exclude photoconductive and photovoltaic effects the samples were measured in the dark. The Hall system was assembled from a Tesla Multimeter model BM518, a voltage source Statron model 3205 covering the voltage range from 0 to 30 V and a high input impedance voltmeter. The magnetic field of 0.385 T was created from a permanent magnet. The obtained measurement results were acquired and processed by a personal computer (PC) and home developed software. The data reported from PC were a carrier type and concentration p/n, resistivity ρ, Hall coefficient R_H and carrier (or Hall) mobility μ_H. When evaluating the Hall measurement results should be

considered some underlying limits associated with them like are Hall scattering factor (r_H) (generally assumed to be 1), and the small magnitude of the Hall voltage. For thin films with low Hall mobility like ZnO:N and ZnO:Al:N the difference between the Hall voltage with and without a magnetic field will be small, making accurate measurement of the Hall voltage difficult. This required repeating and careful measurements to make sure the measured changes in the Hall voltage were accurate and consistent for the applied magnetic field and current.

3.3.1 ZnO:N thin films

The room temperature Hall measurements were carried at specified currents, depending on the film's resistivity. Four types of variations are followed in the data depending on the N_2 content in the sputtering gas, negative bias voltage on the substrate, type of the substrate and heat treatment.

3.3.1.1 *Effect of the N_2 content in the sputtering gas*

The electrical parameters of ZnO:N deposited on Corning Glass substrates as a function of the N_2 content into the sputtering gas are specified in Table 6.

N_2 content (%)	Mobility (cm^2/Vs)	Concentration (cm^{-3})	Resistivity (Ωcm)
0	NM	NM	NM
10	NM	NM	5.4×10^4
25	3	$p \sim 1.2 \times 10^{15}$	1.5×10^3
50	23	$n \sim 4 \times 10^{14}$	7.0×10^2
75	25	$p \sim 3.2 \times 10^{14}$	7.9×10^2
100	23	$p \sim 1.3 \times 10^{14}$	2.1×10^3

Table 6. Electrical parameters of ZnO:N deposited on Corning glass 7059 substrates as a function of the N_2 content

The term NM refers to "not measurable", when the sample's resistance is too high (higher than $5.4 \times 10^4 \, \Omega cm$) and the current source is unable to drive the necessary current. This occurred for ZnO:N thin films deposited without N_2 and with 10 % N_2 in the sputtering gas. The effect of the N_2 content on the resistivity, mobility and carrier concentration of *p*-type ZnO:N thin films grown on Corning glass substrates can be seen in Fig. 8. The resistivity of the un-doped ZnO thin films (0 % N_2 in the sputtering gas) is too high to be measured. In fact, the films doped at 10 % N_2 have the highest resistivity of $5.4 \times 10^4 \, \Omega cm$.

Owing to nitrogen incorporation and creation of nitrogen acceptors, the films prepared at 25 %, 75 % and 100 % N_2 in the sputtering gas show *p*–type features. Nitrogen incorporation was confirmed by secondary ion mass spectroscopy (SIMS) (Shtereva et al., 2006).

Conductivity type of the samples grown at 50% nitrogen is rather controversial and statistically was determined to be *n*-type. Four subsequent measurements were conducted with a current of 200 nA driven through this sample. With the measured Hall voltages, film's thickness and current, type of conductivity was determined three times as n–type, and once as *p*–type. Measurement was repeated and three subsequent measurements were carried out with a current of 500 nA. The calculations with newly obtained Hall voltages determined the sample two times as a *p*–type and once as an *n*–type. Since denoted *n*-type, the electrical parameters of this sample are not included in Fig. 8 and the following discussion.

Fig. 8. Effect of the N_2 content on (a) the carrier concentration and mobility, and (b) resistivity for ZnO:N deposited on Corning glass 7059 substrates

Resistivity variations of two orders of magnitude, mobility variations of a factor of ~ 8, and concentration variations of an order of magnitude occur with N_2 content changing from 25 to 100%. More than 50% reduction of the resistivity, from 1.5×10^3 Ωcm to 7.9×10^2 Ωcm, follows the increase of nitrogen percentage from 25 to 50 %. The lowest resistivity is obtained for 75 % N_2 and is a result of the highest mobility. The further supply of nitrogen to the Ar/N_2 gas to amount of 100 %, leads to 2.7 times increase of the resistivity to the value of 2.1×10^3 Ωcm.

Although the SIMS depth profile shows the incorporation of nitrogen in the films, the carrier concentration is reduced from 1.2×10^{15} cm^{-3} to 1.3×10^{14} cm^{-3} when nitrogen percentage in working gas increasing from 25 to 100 %. The reason for this decrease in the carrier concentration is more likely due to compensation of the N_O acceptor by native and nitrogen related donor defects as well as by unintentionally introduced donor defects. First-principals calculations denote O vacancies and N-acceptor – Zn-antiside (N_O-Zn_O) complexes major compensating donors for a normal N_2 source, whereas nitrogen molecules $(N_2)_O$ and N_O - $(N_2)_O$ complexes are the major compensating species for a plasma N_2 source (Lee et al, 2001). Hence, achievement of high hole concentrations at low nitrogen doping levels is restricted by the O vacancies, whereas at high doping levels the N_O acceptor is compensated by donor complexes that nitrogen forms with the native defects. More recent theoretical

results point to the role of the complexes that nitrogen acceptor forms with unintentionally introduced donor defects, such as hydrogen (N_O - H complex with a low binding energy of about 1 eV), and the carbon impurities (NC_O and $(N_2)_O$) (Limpijumnong et al., 2006). They both are donors and compensate the N_O acceptor. N and C were registered by SIMS in all our ZnO:N films.

Chemical potential for nitrogen (μ_N) increases and the activation energy of the N_O acceptor decreases as a result of the increase in N concentration. Chemical potential for nitrogen is determined by the formation energy of the nitrogen molecule ($\mu_N^{max} = \dfrac{1}{2}\mu_{N_2}$), hence, the formation energy of $(N_2)_O$ molecule decreases double, which means that in thermal equilibrium conditions the concentration of $(N_2)_O$ donors will increase faster than the concentration of acceptors (Zhang et al., 2001). This explains why the increase in the N concentration often does not result in high hole concentrations p-type ZnO.

The carrier transport is influenced from various scattering mechanisms that determine carrier mobility in semiconductors. The major scattering mechanisms in ZnO thin films are presumed to be: (i) ionized impurity scattering; (ii) neutral impurity scattering, and (iii) grain boundary scattering. In the case of ionized impurity scattering, mobility shall decrease with increasing carrier concentration when the free carriers' density is equal to the concentration of the ionized donors/acceptors. The dependence of mobility on carrier concentration is plotted on Fig. 9. Mobility increases with increasing N_2 content from 25 to 75 % in consequence of the decreasing hole concentration (Fig. 8a). The highest Hall mobility of 25 cm^2/Vs along with a hole concentration of 3.2×10^{14} cm^{-3} are found in ZnO:N prepared at 75 % N_2 and result in the lowest resistivity of these films.
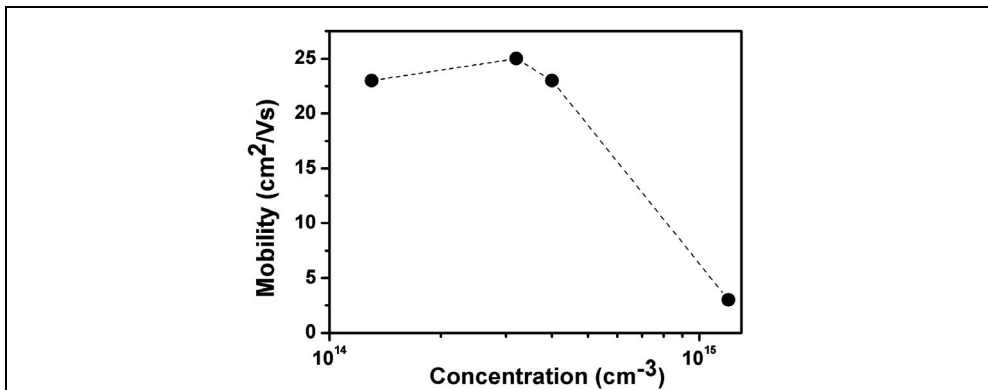
Fig. 9. Mobility of ZnO:N films as a function of the carrier concentration

3.3.1.2 *Effect of the negative bias voltage applied on the substrate*

The electrical parameters of ZnO:N thin films deposited on Corning Glass and Si/SiO$_2$ substrates as a function of the negative bias voltage applied on the substrate are given in Table 7. All films were prepared at 75 % N_2 in the sputtering gas.

Fig. 10 illustrates how the hole carrier concentration, resistivity and mobility vary with the negative bias voltage for ZnO:N films deposited on glass substrates. Irrespectively of the used substrate, the carrier concentration is lower, and mobility and resistivity are higher, for samples prepared at negative bias compared to those grown without a bias voltage. Indeed,

Substrate	Bias voltage (V)	Mobility (cm^2/Vs)	Concentration (cm^{-3})	Resistivity (Ωcm)
Si/SiO$_2$	0	469	$p \sim 2.0 \times 10^{12}$	6.7×10^3
Si/SiO$_2$	-25	371	$p \sim 2.6 \times 10^{12}$	6.4×10^3
Si/SiO$_2$	-50	480	$n \sim 1.7 \times 10^{12}$	7.7×10^3
Si/SiO$_2$	-100	28	$p \sim 4.0 \times 10^{13}$	5.7×10^3
Glass	0	25	$p \sim 3.2 \times 10^{14}$	7.9×10^2
Glass	-25	121	$p \sim 1.2 \times 10^{13}$	4.4×10^3
Glass	-50	51	$p \sim 2.6 \times 10^{13}$	4.7×10^3
Glass	-100	357	$n \sim 5.0 \times 10^{12}$	3.5×10^3

Table 7. Electrical parameters of ZnO:N deposited on Corning glass 7059 and Si/SiO$_2$ substrates as a function of the negative bias voltage

the intensive bombardment by positive ions on the substrate, induced from the negative bias voltage, will increase the substrate temperature and hence, improvement of the crystalline structure of the film discussed in the previous subsection. On the other hand, this ion bombardment can induce defects in the film, which influence its electrical parameters.
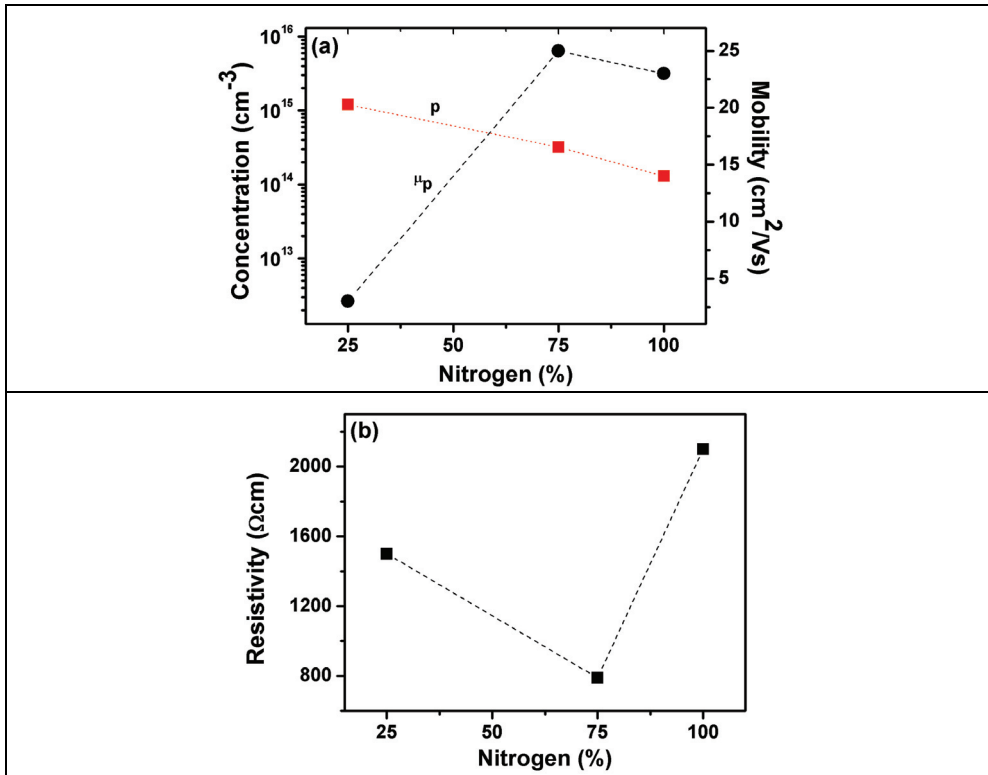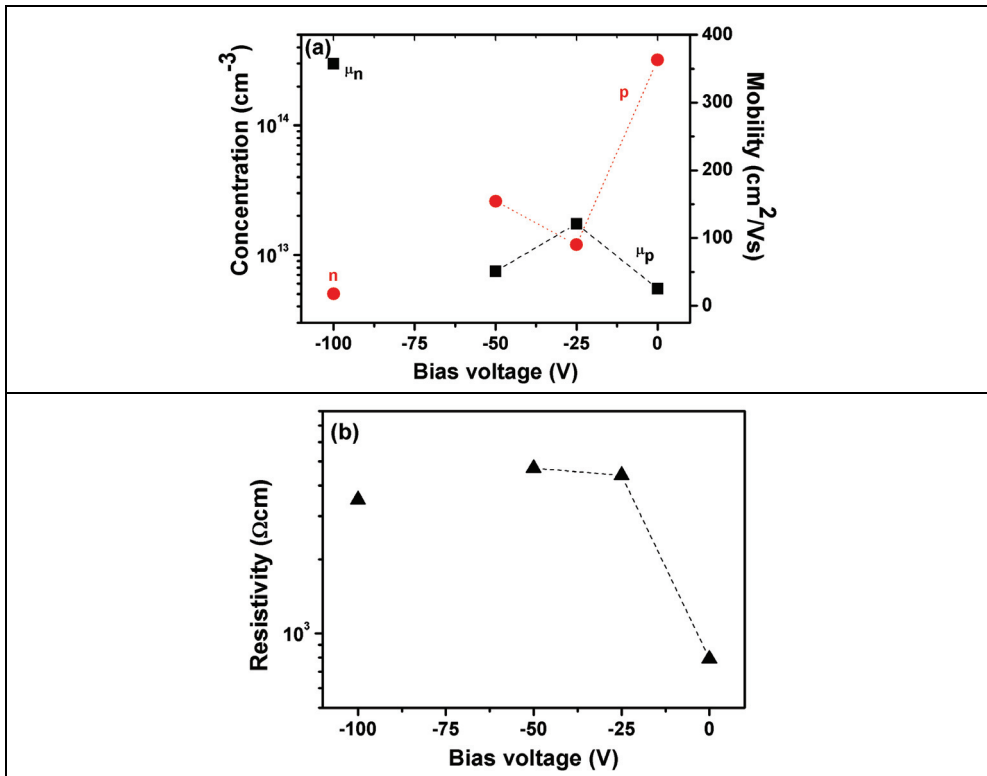

Fig. 10. Effect of the bias voltage on (a) the carrier concentration and mobility and (b) resistivity, for ZnO:N deposited on Corning glass 7059 substrates

3.3.1.3 *Effect of annealing*

Post-deposition annealing is known to improve crystalline structure and to decrease the background defects in ZnO. The ZnO:N thin films deposited with 75 % and 100 % N_2 in the sputtering gas (Samples No 12 and 13) were annealed in N_2 atmosphere at a temperature of 600°C for 10 minutes. The electrical parameters of the as deposited and annealed samples are presented in a Table 8.

Sample Noe	N_2 content (%)	Mobility (cm²/Vs)	Concentration (cm⁻³)	Resistivity (Ωcm)
12	75	25	p~ 3.2×10^{14}	7.9×10^2
13	100	23	p~ 1.3×10^{14}	2.1×10^3
12A	75	5	n~ 5.4×10^{18}	2.4
13A	100	3	n~ 1.1×10^{18}	2

Table 8. Electrical parameters of ZnO:N thin films as a function of annealing in N_2 atmosphere at temperature of 600°C.

The annealed samples are marked as 12A and 13A respectively.

After annealing, type of conductivity changes from *p*–type to *n*–type for both samples more likely due to the out-diffusion of nitrogen from the film during heat treatment and the decrease in concentration of the nitrogen acceptors. Mobility and resistivity are reduced as a result of the increase carrier concentration.

Raman spectra of ZnO:N films reveal a decrease in intensity of the peak at 271.5 cm⁻¹ that is assigned to the local vibrational modes (LVM) of nitrogen (Tvarozek et al. 2008). According to some works, intensity of the dopant's LVM has a linear dependence on the dopant concentration (Kaschner et al. 1999). On the other hand, intensity of the peak at 576.7 cm⁻¹, which is associated with the presence of defects, such as O vacancies and Zn interstitials, increases after annealing.

3.3.2 ZnO:Al:N thin films

The aim of these experiments was to investigate the influence of Al – N_2 co–doping on the electrical properties of sputtered ZnO thin films and thereby to verify the effectiveness of the co-doping concept. The importance of this task is determined by the following requirements: (i) to decrease their resistivity and increase their transmittance and (ii) to understand better the role of aluminum, as nitrogen co–dopant, on their physical properties. The ZnO:Al:N thin films were deposited at different N_2 contents in the sputtering gas mixture: 0 %, 25 %, 50 % and 75 %. Their electrical parameters were determined by means of Hall-effect measurements, carried out at room temperature. In contrary to ZnO:N, the Al – N_2 co – doped thin films have lower resistivity which allowed the Hall voltages to be taken at higher currents, depending on the film resistivity. Hall data, taken twice with two weeks period between the first and the second measurement, are listed in Table 9.

As expected, the thin films doped only with aluminum, without nitrogen in the sputtering gas (ZnO:Al), are *n*–type and high conductive. The high carrier concentrations of ZnO:Al compared to nitrogen-doped ZnO results from the contribution of the Al^{3+} substituents in addition to the native donor defects.

The conductivity type of the samples prepared with 25 % N_2 in the sputtering gas was unstable, being *p*–type during the first measurement and turning to *n*–type during the

N$_2$ content (%)	Mobility (cm^2/Vs)	Concentration (cm^{-3})	Resistivity (Ωcm)
0	1	n ~ 5.1x10^{19}	8.9x10^{-2}
25	4	n ~ 1.1x10^{18}	1.5
50	2	n ~ 6.5x10^{17}	6.1
75	0.4	p ~ 7.8x10^{17}	21

Table 9. Electrical parameters of ZnO:Al:N deposited on Corning glass 7059 substrates as a function of the N$_2$ content

second measurement, which is an indication of a compensate semiconductor. The films prepared at 50% N$_2$ exhibit persistent *n*–type conductivity. These results are consistent with the results obtained for nitrogen mono-doped ZnO and discussed in the previous subsection.

The repeated measurements show a stable *p*–type features of the ZnO:Al:N thin films prepared at 75 % N$_2$ in the sputtering gas. It suggests existence of the "nitrogen content window" where p-type doping in ZnO can be realized. The resistivity of *p*–type ZnO:Al:N thin films measured in this experiment is 21 Ωcm, which is roughly a factor of 40 lower than the lowest ZnO:N resistivity (790 Ωcm), reported in 3.3.1.1. Accordingly, the hole carrier concentration of the former (7.8x10^{17} cm^{-3}) is more than three orders of magnitude higher than the ZnO:N hole concentration (3.2x10^{14} cm^{-3}). This electrical parameters improvement is due to Al co–doping.

The carrier concentration, resistivity and Hall mobility are plotted as a function of the N$_2$ content in Fig. 11.

4. Conclusions

Our research on the growing and characterizing of *p*-type ZnO thin films, prepared by radio frequency (RF) diode sputtering, mono-doped with nitrogen, and co-doped with aluminium and nitrogen, is a response of the need from *p*-type *ZnO* thin films for device applications. The dopants determine the conductivity type of the film and its physical properties. We obtained *p*-type *ZnO* thin films by RF diode sputtering and using a nitrogen dopant source. The novelty in our approach is in the use of a plasma assisted deposition method, to increase nitrogen solubility and the concentration of the N$_O$ acceptors in the film.

The structural parameters, such as preferential orientation, crystallite size and strain in the film, depend on percentage of nitrogen in the working gas and the negative bias voltage. The XRD diffraction patterns reveal more or less stronger expressed c-axis preferential orientation of the films depending on the N$_2$ content in the sputtering gas. The diffraction lines of ZnO (100), (002), (101), and (110) are observed in the XRD diffraction patterns of all ZnO:N thin film. Aluminium- nitrogen co-doping improves the crystalline structure and all ZnO:Al:N thin films show a c-axis texture in direction declined about 16 deg from the surface normal. Three diffraction lines ((100), (002) and (101)), appear only in the XRD patterns of the films deposited at 25 % N$_2$, indicating that their preferential orientation is more random. A strong preferential orientation of the crystallites is observed for high dopant contents (50 % and 75 % N$_2$) and intensities of the (002) diffraction lines are almost the same.
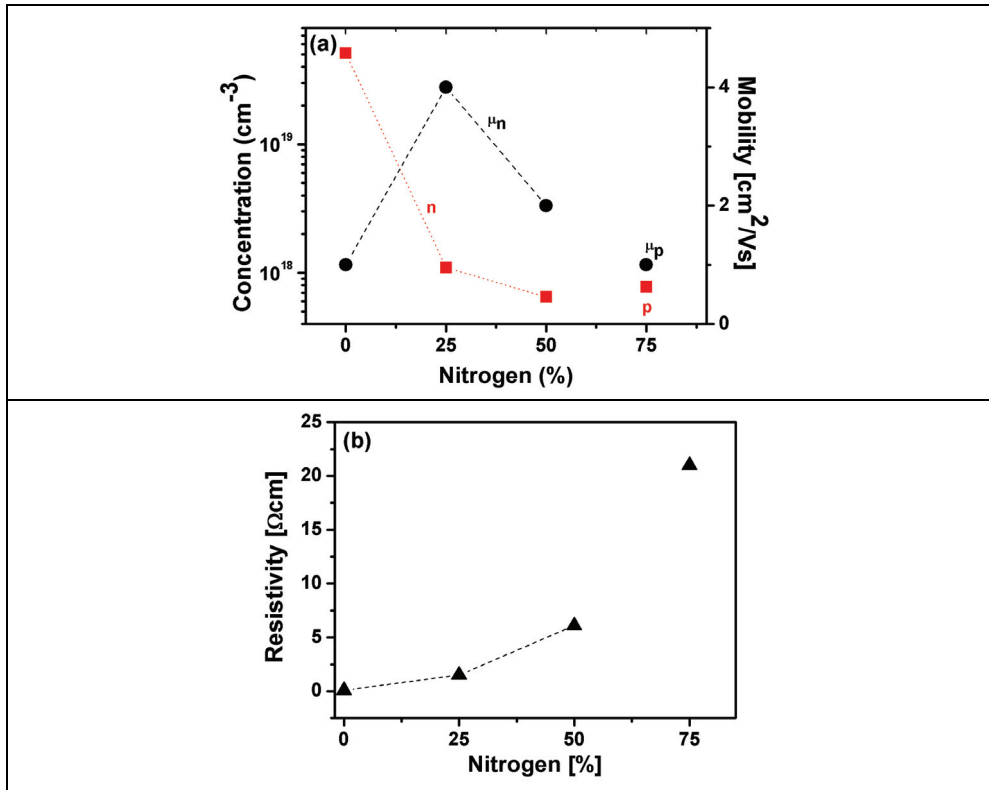
Fig. 11. Effect of the N_2 content on (a) the carrier concentration and mobility and (b) resistivity of ZnO:Al:N films deposited on Corning glass 7059 substrates

The undoped ZnO thin films (0 % N_2 in the sputtering gas) have high resistivity. The minimum resistivity of 8.9×10^{-2} Ωcm and the highest carrier concentration of 5.1×10^{19} cm^{-3} obtained for ZnO:Al films result from the contribution of the Al^{3+} substituents (Al^{3+} that substitute for Zn^{2+} in the ZnO lattice), in addition to the native donor defects, such as Zn interstitials and O vacancies.

p–type conductivity in the ZnO:N thin films and ZnO:N:Al is a result of the nitrogen which substitutes for oxygen in the crystal lattice, creating N_O acceptors. The incorporation of N_2 was confirmed by means of SIMS depth profiling and Raman – scattering measurements. Besides native donor defects, p-type doping in ZnO is limited from the formation of molecules $(N_2)_O$ and complexes $(N_O - (N_2)_O, N_O – Zn_O)$ which are the surplus donors to the native ZnO donors.

5. Acknowledgements

6. References

Assunção, V., Fortunato, E., Marques, A., Águas, H., Ferreira, I., Costa, M. E. V. & Martins, R. (2003). Influence of the deposition pressure on the properties of transparent and conductive ZnO:Ga thin-film produced by r.f. sputtering at room temperature, *Thin Solid Films*, 427, (2003), 401-405

Carotta, M. C.; Cervi, A., Natale, V., Gherardi, S., Giberti, A., Guidi, V., Puzzovio, D., Vendemiati, B., Martinelli, G., Sacerdoti, M., Calestani, D., Zappettini, A., Zha, M. & Zanotti L. (2009). ZnO gas sensors: A comparison between nanoparticles and nanotetrapods-based thick films, *Sensors and Actuators B*, 137 (2009), page numbers (164–169)

Delhez, R., De Keijser, Th. H., Mittemeijer, E.J. & Fresenius Z. (1982). *Anal. Chem.*, 312, (1982), 1-16

Fortunato, E., Assunção, V., Gonçalves, A., Marques, A., Águas, H., Pereira, L., Ferreira, I., Vilarinho, P., Martins, R. (2004). High quality conductive gallium-doped zinc oxide films deposited at room temperature, *Thin Solid Films*, 451 –452, (2004), 443–447

Fu, E. G., Zhuang, D. M., Zhang, G., Ming, Z., Yang, W. F., Liu, J. J. (2004). Properties of transparent conductive ZnO:Al thin films prepared by magnetron sputtering, *Microelectronics Journal*, 35, (2004), 383–387

Ganguly, G.; Carlson, D. E., Hegedus, S. S., Ryan, D., Gordon, R. G., Pang, D., Reedy, R. C. (2004). Improved fill factors in amorphous silicon solar cells on zinc oxide by insertion of a germanium layer to block impurity incorporation, *Applied Physics Letters*, Vol. 85, No 3, (July 2004), page numbers (479-481)

Ge, F. Z., Ye, Z. Z., Zhu, L. P., Lü, J. G., Zhao, B. H., Huang, J. Y., Zhang, Z. H., Wang, L. & Ji, Z. G. (2004). Electrical and optical properties of Al–N co-doped p-type zinc oxide films, *Journal of Crystal Growth*, 268, (2004), 163–168

Guillén, C., Herrero J. (2006). High conductivity and transparent ZnO:Al films prepared at low temperature by DC and MF magnetron sputtering, *Thin Solid Films*, 515, (2006) 640 – 643

Hwang, D. K., Kim, H. S., Lim, J. H., Oh, J. Y., Yang, J. H., Park, S. J., Kim, K. K., Look, D. C. & Park, Y. S. (2005). Study of the photoluminescence of phosphorus-doped p-type ZnO thin films grown by radio-frequency magnetron sputtering, *Applied Physics Letters*, 86, (2005), 151917

Ji, G.; Zhang, Z., Chen, Y., Yan, Sh., Liu, Y. & Mei, L. (2009). Current spin polarization and spin injection efficiency in ZnO-based ferromagnetic semiconductor junctions. *Acta Metallurgica Sinica* (*English Letters*), Vol. 22, No. 2, (April 2009), page numbers (153-160)

Joseph, M., Tabata, H., Saeki, H., Ueda, K. & Kawai, T. (2001). Fabrication of the low-resistive p-type ZnO by codoping method, Physica B 302–303, (2001), 140–148

Kaschner, A., Siegle, H., Hoffmann, A. & Thomsen C. (1999). Influence of doping on the lattice dynamics of gallium nitride, *MRS Internet J. Nitride Semicond. Res.*, 4S1, G3.57, (1999)

Kraus, I. (1993). Struktura a vlastnosti krystalů, ACADEMIA Praha, str. 221

Lee, E. Ch., Kim, Y. S., Jin, Y. G. & Chang, K. J. (2001). First-principles study of the compensation mechanism in N-doped ZnO, *Phisica B*, 912, (2001), 308-310.

Limpijumnong, S., Li, X., Wei, S.-H. & Zhang, S. B. (2006). Probing deactivations in Nitrogen doped ZnO by vibrational signatures: A first principles study, *Physica B*, 376–377, (2006), 686–689

Look, D. C., Coşkun, C., Claflin, B. & Farlow, G. C. (2003). Electrical and optical properties of defects and impurities in ZnO, *Physica B*, 340–342, (2003), 32–38.

Lu, J., Zhang, Y., Ye, Z., Wang, L., Zhao, B. & Huang, J. (2003). p-type ZnO films deposited by DC reactive magnetron sputtering at different ammonia concentrations, *Materials Letters*, 57, (2003), 3311 –3314

Lu, J. G., Fujita, S., Kawaharamura, T. & Nishinaka, H. (2007). Roles of hydrogen and nitrogen in p-type doping of ZnO, *Chemical Physics Letters*, 441, (2007), 68–71

Lu, J. G., Zhu, L. P., Ye, Z. Z., Zhuge, F., Zeng, Y. J., Zhao, B. H. & Ma, D. W. (2005). Dependence of properties of N–Al codoped p-type ZnO thin films on growth temperature, *Applied Surface Science*, 245, (May 2005), 109-113

Minami, T., Yamamoto, T. & Miyata, T. (2000). Highly transparent and conductive rare earth-doped ZnO thin flms prepared by magnetron sputtering, *Thin Solid Films*, 366, (2000), 63-68

Moon, T. H., Jeong, M. C., Lee, W. & Myoung J. M. (2005). The fabrication and characterization of ZnO UV detector, *Applied Surface Science* 240, (2005), 280–285

Nunes, P., Fortunato, E., Tonello, P., Fernandes, F. B., Vilarinho, P. & Martins, R. (2002), Effect of different dopant elements on the properties of ZnO thin films, *Vacuum*, 64, (2002), 281–285

Ohta, H.; Hosono, H. (2004).Transparent oxide optoelectronics, *Materialstoday*, (June 2004), page numbers (42-51), ISSN:1369 7021 © Elsevier Ltd 2004

Özgür, Ü., Alivov, Y. I., Liu, C., Teke, A., Reshchikov, M. A., Doğan, S., Avrutin, V., Cho, S.-J. & Morkoç, H. (2005). A comprehensive review of ZnO materials and devices, Journal of Applied Physics 98, (2005), 041301

Park, S. H., Chang, J. H., Ko, H. J., Minegishi, T., Park, J. S., Im, I. H., Ito, M., Oh, D. C., Cho, M. W. & Yao, T. (2008). Lattice deformation of ZnO films with high nitrogen concentration, *Applied Surface Science*, 254, (2008), 7972–7975

Ryu, Y. R., Zhu, S., Look, D. C., Wrobel, J. M., Jeong, H. M. & White, H. W. (2000). Synthesis of p-type ZnO films, *Journal of Crystal Growth*, 216, (2000), 330-334

Shtereva, K.; Flickyngerova, S.; Kovac, J.; Tvarozek, V.; Novotny, I.; Skriniarova, J.; Srnanek, R. & Rehakova, A.; (2008). Preparation of p-/n- type ZnO:Al thin films by RF diode sputtering for solar and optoelectronic applications, *Proceedings of 26th International Conference on Microelectronics (MIEL 2008)*, pp. 247 – 250, Nis Serbia, 11-14 May 2008

Shtereva, K., Tvarozek, V., Novotny, I., Kovac, J., Sutta P. & Vincze A. (2006). p-Type Conduction in Sputtered ZnO Thin Films Doped by Nitrogen, *Proceedings of 25th International Conference on Microelectronics (MIEL 2006)*, pp. 357–360, Belgrade, Serbia & Montenegro, 14-17 May 2006

Song, P. K., Watanabe, M., Kon, M., Mitsui, A. & Shigesato, Y. (2002). Electrical and optical properties of gallium-doped zinc oxide films deposited by dc magnetron sputtering, *Thin Solid Films*, 411, (2002), 82–86

Sze, S. M. (2002). *Semiconductor devices, Physics and Technology*, John Wiley&Sons, New York, 2002, ISBN 0-471-33372-7

Sun, J. W., Lu, Y. M., Liu, Y. C., Shen, D. Z., Zhang, Z. Z., Li, B. H., Zhang, J. Y., Yao, B., Zhao, D. X. & Fan, X. W. (2006). The activation energy of the nitrogen acceptor in p-

type ZnO film grown by plasma-assisted molecular beam epitaxy, *Solid State Communications*, 140, (2006), 345-348

Šutta, P. & Jackuliak Q. (1998). Macro-stress formation in thin films and its investigation by X-ray diffraction, *Proceedings of 2nd International Conference on Advanced Semiconductor Devices and Microsystems (ASDAM'98)*, pp. 227-230, Smolenice Castle, Slovakia, 5-7 October 1998

Tvarozek, V., Novotny, I., Sutta, P., Flickyngerova, S., Schtereva, K. & Vavrinsky, E. (2007). Influence of sputtering parameters on crystalline structure of ZnO thin films, *Thin Solid Films*, 515, (2007), 8756-8760

Tvarozek, V., Shtereva, K., Novotny, I., Kovac, J., Sutta, P., Srnanek, R. & Vincze, A. (2008). RF diode reactive sputtering of n-and p-type zinc oxide thin films, *Vacuum*, 82, (2008), 166-169

Tüzemen, S. & Gür, E. (2007). Principal issues in producing new ultraviolet light emitters based on transparent semiconductor zinc oxide, *Optical Materials*, vol. 30, (2007), 292–310, 2007.

Van de Walle, Ch. G. (2000). Hydrogen as a Cause of doping in Zinc Oxide, *Physical Review Letters*, 31, vol. 85, No5, (31 July 2000), 1012-1015

Wang, C., Ji, Z., Xi, J., Du, J. & Ye, Z. (2006). Fabrication and characteristics of the low-resistive p-type ZnO thin films by DC reactive magnetron sputtering, *Materials Letters*, 60, (2006), 912–914

Wang, Q. J.; Pflügl, Ch., Andress, W. F., Ham, D. & Capasso F., Yamanishi M. (2008). Gigahertz surface acoustic wave generation on ZnO thin films deposited by radio frequency magnetron sputtering on III-V semiconductor substrates. *Journal of Vacuum Science & Technology B*, Vol. 26, No 6, (Nov/Dec 2008), page numbers (1848 -1851)

Xiong, G., Wilkinson, J., Mischuck, B., Tüzemen, S., Ucer, K. B. & Williams, R. T. (2002) Control of p- and n-type conductivity in sputter deposition of undoped ZnO, *Applied Phisics Letters*, 80, (February 2002), 7-18

Yamamoto, T. (2002). Codoping for the fabrication of *p*-type ZnO, *Thin Solid Films*, 420-421, (2002), 100-106

Yao, B., Guan, L. X., Xing, G. Z., Zhang, Z. Z., Li, B. H., Wei, Z. P., Wang, X. H., Cong, C. X., Xie, Y. P., Lu, Y. M. & Shen, D. Z. (2007). P-type conductivity and stability of nitrogen-doped zinc oxide prepared by magnetron sputtering, *Journal of Luminescence*, 122–123, (2007), 191–194

Ye, Z. Z., Ge, F. Z., Lu, J.-G., Zhang, Z. H., Zhu, L. P., Zhao, B. H. & Huang, J. Y. (2004). Preparation of p-type ZnO films by Al+N-codoping method, *Journal of Crystal Growth*, 265, (2004), 127–132

Ye, Z. Z., Qian, Q., Yuan, G. D., Zhao, B. H. & Ma, D. W. (2005). Effect of oxygen partial pressure ratios on the properties of Al–N co-doped ZnO thin films, *Journal of Crystal Growth*, 274, (2005), 178–182

Yuan, G., Ye, Z., Zhu, L., Zeng, Y., Huang, J., Qian, Q. & Lu, J. (2004). p-Type conduction in Al–N co-doped ZnO films, Materials Letters 58, (2004), 3741–3744

Zeng, Y. J., Ye, Z. Z., Xu, W. Z., Chen, L. L., Li, D. Y., Zhu, L. P., Zhao, B. H. & Hu, Y. L. (2005). Realization of p-type ZnO films via monodoping of Li acceptor, *Journal of Crystal Growth*, 283, (September 2005), 180-184

Zhang, S. B., Wei, S. H. & Yan, Y. (2001). The thermodynamics of codoping: how does it work?, *Physica B*, 302–303, (2001), pp.135–139.

Zhao, J.-L., Li, X.-M., Bian, J.-M., Yu, W.-D. & Zhang, C.-Y. (2005). Growth of nitrogen-doped p-type ZnO films by spray pyrolysis, *Journal of Crystal Growth*, 280, (2005), 495–501

Zhu, L., Ye, Z., Zhuge, F., Yuan, G. & Lu, J. (2005). Al–N codoping and p-type conductivity in ZnO using different nitrogen sources, *Surface & Coatings Technology*, 198, (2005), 354–356

Self-Aligned π-Shaped Source/Drain Ultrathin SOI MOSFETs

Yi-Chuen Eng and Jyi-Tsong Lin

National Sun Yat-Sen University, Department of Electrical Engineering
70, Lienhai Road, Kaohsiung 80424,
Taiwan, R.O.C.

1. Introduction

In this chapter, we shall study the short-channel characteristics of **self-aligned π-shaped source/drain ultrathin silicon-on-insulator metal-oxide semiconductor field-effect transistor** (SA-πFET). The only difference between conventional and proposed **ultrathin silicon-on-insulator** (UTSOI) transistors is that a path from the **source/drain** (S/D) to the **silicon** (Si) substrate is created and called the **S/D tie** (SDT). Thus, UTSOI **metal-oxide semiconductor field-effect transistor** (MOSFET) thermal performance can be enhanced drastically by opening up the SDT rather than increasing Si body thickness.

Although the path between S/D and Si substrate has degraded the device properties slightly, the short-channel characteristics of SA-πFET are within acceptable limits due to the existence of **UT body** (UTB). After changing the S/D structure in the proposed SOI transistor, the **n-channel enhancement-type MOSFET** (NMOS) current drive gets improved accordingly. Furthermore, the effects of self-heating on SA-πFET performance can be reduced greatly. This is ascribed to the fact that the forms of additional leakage paths truly help dissipate the heat generated by the thermal vibrations of the crystalline lattice phonons. For these reasons, quasi-SOI devices are strong contenders for future **complementary MOS** (CMOS) technology.

The objectives of this chapter are to describe the physical structure of the **π-shaped S/D** (π-S/D) transistor and its process, to give an understanding of why SDT design must use, and to discuss the short-channel characteristics compared with those of a conventional UTSOI. By the end of this chapter, the reader should be able to know the importance of the design of SDT.

2. Structure and process of the SA-πFET

Figure 1 shows the physical structure of the SA-πFET. Observe that the SDT has a length L_{SDT} and a location which is determined by the length of **Si nitride** (SiN) L_{SP}, two important parameters of the SA-πFET.

A simplified description of the fabrication of a SA-πFET is as follows (see Fig. 2(a)−(f)). The SOI wafer structure is used to make π-S/D transistors, which has a Si layer located on top of the **buried oxide** (BOX) and a **bulk Si** (bulkSi) substrate layer located below the oxide insulating layer. The final Si layer thickness is obtained by thermal oxidation and etching down to 5 nm. A channel implantation process is first performed with **boron difluoride** (BF_2), 2.3 KeV, $1.15×10^{12}$ cm^{-2}. Following this, device isolation is achieved using a traditional

<u>s</u>hallow <u>t</u>rench <u>i</u>solation (STI) approach. A gate insulator of **Si dioxide** (SiO_2) is thermally growth and a **<u>poly</u>crystalline-<u>Si</u>** (poly-Si) layer as a gate electrode deposited by using the **<u>c</u>hemical <u>v</u>apor <u>d</u>eposition** (CVD) process is then formed. In order to form a Π-S/D scheme, the layer of SiN as hard mask is deposited by CVD. After the patterning of the gate stack (see Fig. 2(a)), a SiN layer for forming the spacer is deposited and etched back, as shown in Fig. 2(b). The sidewall spacer hard mask is used for etching Si and BOX, respectively (see Fig. 2(c)). A layer of poly-Si is deposited as SDT shown in Fig. 2(d). After the deposition and planarization of the SiO_2 layer, the etching process is performed in order to form a BOX layer under the source and drain regions (see Fig. 2(e)). The poly-Si layer is deposited, patterned, and etched to create the active region of the S/D, as shown in Fig. 2(f). Next, the S/D implantation process is carried out by **<u>ar</u>senic** (As), 10 KeV, 2.1×10^{14} cm^{-2}. **<u>R</u>apid <u>t</u>hermal <u>a</u>nnealing** (RTA) process is followed to activate the dopants and repair the lattice damage that is caused by the implantation process. Finally, a conventional SOI fabrication flow can be used for **<u>back-end-of-line</u>** (BEOL) processing.

The simulation parameters are T_{BOX} = 40 nm, T_{BOI} = 50 nm, $T_{S/D}$ = T_{Si} = 5 nm, and T_{GOX} = 1.4 nm for the Π-S/D. Various gate lengts L_G (L_{CH} – 9 nm) of 10 nm ~ 70 nm were investiaged. Notice that all the parameters of the UTSOI NMOS are equivalent to those of the Π-S/D NMOS, expect that the T_{BOX} is equal to the T_{BOI} (= 50 nm).
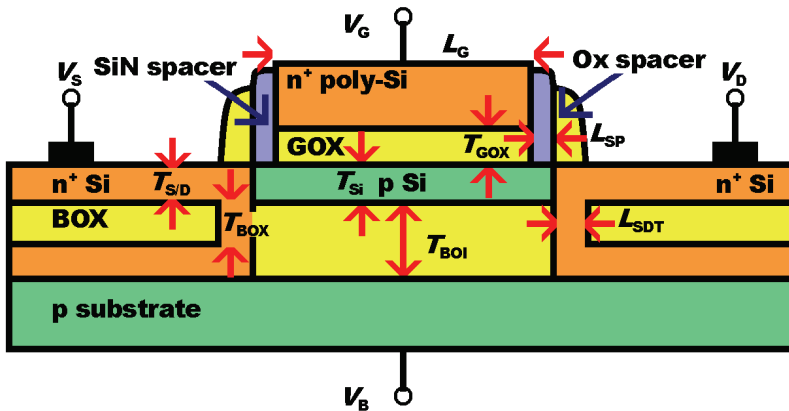
Fig. 1. Schematic cross-sectional view of an n-channel SA-ΠFET. Note that the L_{SDT} and L_{SP} are two important parameters of the SA-ΠFET.

(a)

(b)

(c)

(d)

(e)

(f)
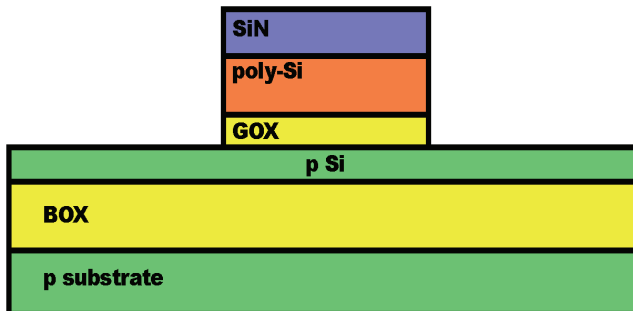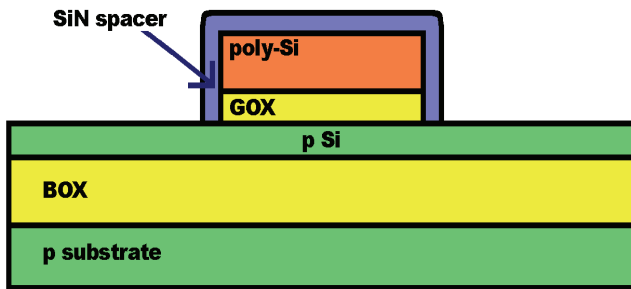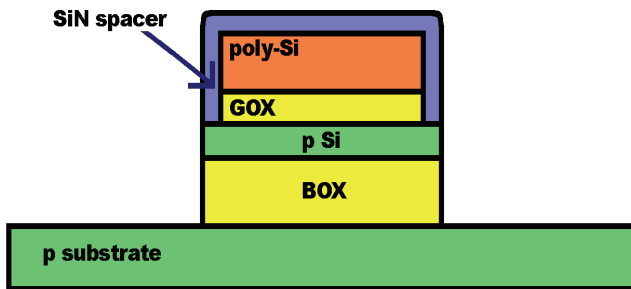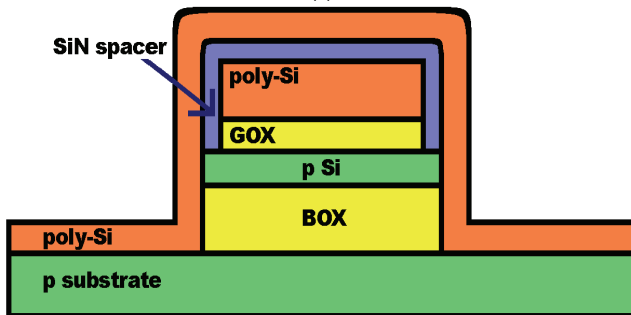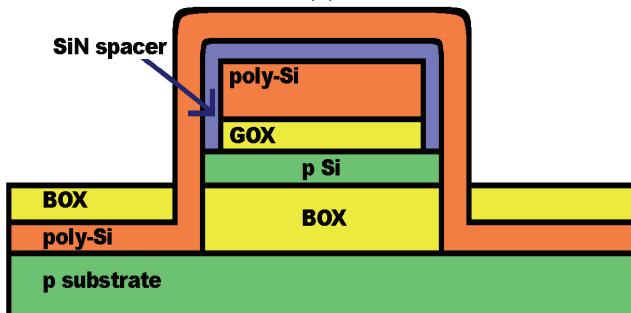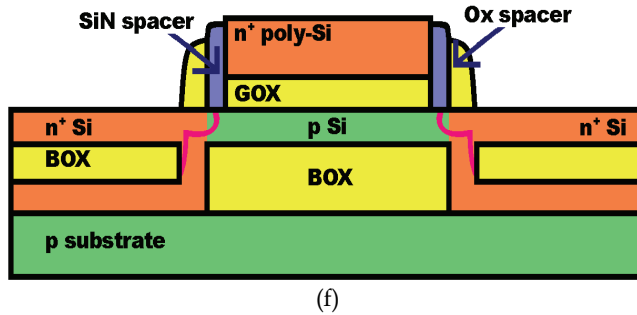
Fig. 2. The SA-πFET fabrication process [1], [2]; (a) gate patterning, (b) SiN spacer formation, (c) Si/BOX etch with a SiN mask, (d) poly-Si deposition, (e) formation of BOX, and (f) S/D poly-Si deposition and formation by CMP and wet etching.

3. Electrical characteristics of the SA-πFET

In this section, we study the physical and electrical characteristics of the SA-πFET. It should be clear that the design of SDT is important for scaled π-S/D transistors. In order to control the **short-channel effects** (SCEs), a conventional UTSOI MOSFET is considered as a strong contender for replacing the position of the bulkSi in near future [3]. However, note that because the SOI family of devices has a BOX layer (which is underneath the active region), the self-heating is undesirable for the performance due to lattice scattering. As device dimensions decrease, the self-heating is more pronounced. Hence the importance of SDT in the conventional UTSOI MOSFET is growing owing to self-heating.

I_{DS}-V_{GS} characteristics of the SA-πFET

Figure 3 shows the drain current I_{DS} versus gate voltage V_{GS} characteristic curves of the SA-πFET compared with a conventional UTSOI FET. Obviously, the leakage current of a conventional UTSOI FET is lower than that of the SA-πFET. This means that an (SDT) additional path appended becomes an encumbrance to the UTSOI NMOS, resulting in increased p-n junction leakage current. However, the short-channel such as **drain-induced barrier lowering** (DIBL) and **subthreshold swing** (S.S.) characteristics of SA-πFET are within acceptable limits because of the presence of UTB. Replacing the conventional UTSOI MOSFET S/D structure with an SDT results in the slight performance degradation in the proposed SOI transistor, but the results can get accepted.

Short-channel effects in π-S/D transistors

As shown in Fig. 4, the dependence of S.S. and **threshold voltage** (V_{TH}) on **gate length** (L_G) for both π-S/D and UTSOI devices is compared. It is obvious in Fig. 4a that the conventional UTSOI MOSFET having a L_{SDT} = 0 nm exhibits low S.S. values. On the contrary, our proposed UTSOI structure having a L_{SDT} = 10 nm shows how the S.S. is slightly degraded. The reason is that for the π-S/D NMOS, the SDT provides additional paths for high electric field from the drain side to influence the channel, which leads to S.S. degradation. Fortunately, the results are within acceptable limits (< 100 mV/dec at L_G = 10 nm). Fig. 4b
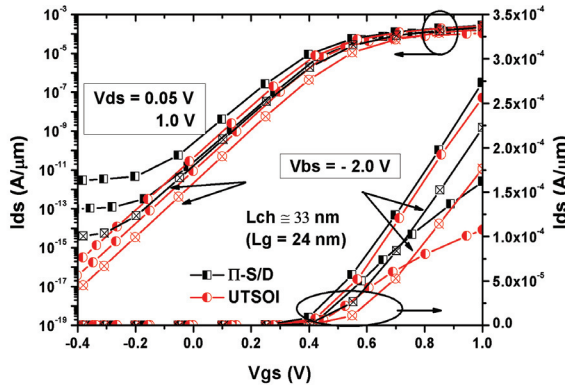
Fig. 3. Comparison of I_{DS}-V_{GS} curves between two transistors Π-S/D and UTSOI.

(a)

(b)
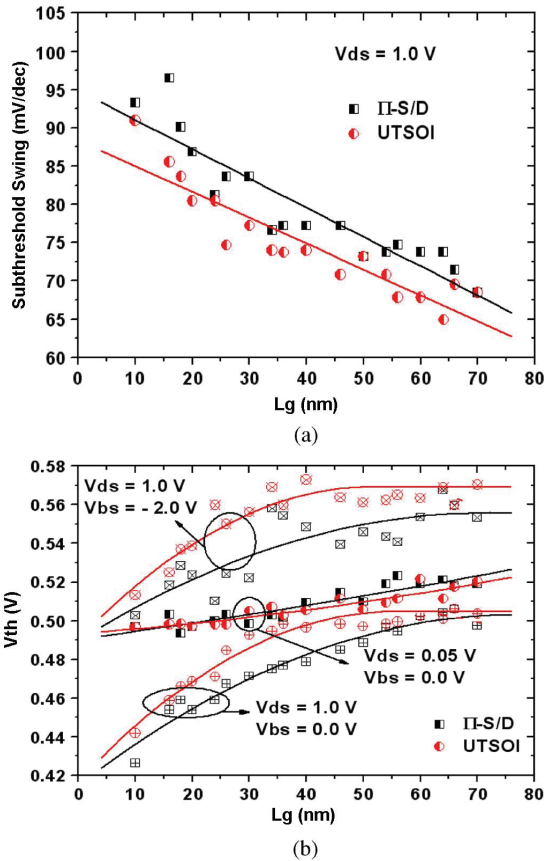
Fig. 4. Effects of the L_G on the (a) S.S. and (b) V_{TH} characteristics of the UTSOI MOSFET w/ and w/o additional SDT.

shows the impact of L_G on V_{TH}. (The intersection of the maximum and minimum slope lines in the $\log(I_{DS})$-V_{GS} characteristic curves was used to extract the V_{TH}.) The V_{TH} is found to decrease with decrease in L_G mainly due to increased charge sharing [4]. It can be seen in figure that the saturation V_{TH} roll-off is worse than linear V_{TH}. When a **substrate bias** (V_{BS}) of – 2.0 V is applied, the V_{TH} is increased, in comparison with V_{BS} equal to 0 V. Apparently, the V_{TH} roll-off behavior of the Π-S/D transistor is quite similar to UTSOI transistor, since both devices have the same Si channel thickness.

Figure 5 shows the impact of L_G on body factor γ and DIBL. The γ was extracted using the method described in [5] ($\gamma = |\Delta V_{TH,SAT}/\Delta V_{BS}|$), which is the shift of the $V_{TH,SAT}$ caused by the change in the V_{BS}, whereas DIBL is the difference between $V_{TH,LIN}$ and $V_{TH,SAT}$. The decrease of the L_G leads to large values of γ and DIBL. This is because the short L_G has less control over the channel region, thereby increasing the **S/D subthreshold off-state leakage current** ($I_{sd,leak}$). For short channel, the V_{BS} has a large effect on $V_{TH,SAT}$. Consequently, the application of V_{BS} in $V_{TH,SAT}$ results in a large γ for short-channel devices. The DIBL of the Π-S/D NMOS is slightly larger than the UTSOI NMOS, as in the case of the S.S. shown in Fig. 4a. The DIBL is similarly impacted by the increased penetration of the fringing fields from the drain region (see Fig. 6).
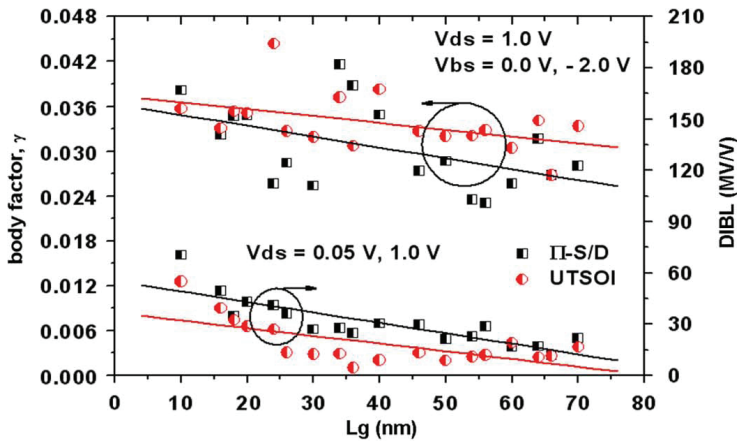
Fig. 5. Effects of the L_G on the γ and DIBL characteristics of the UTSOI MOSFET w/ and w/o additional SDT.

Moreover, we find that the **effective parasitic series S/D resistance** ($R_{S/D}$) for both devices, as shown in Fig. 7 does not decrease dramatically by reducing the L_G. Nevertheless, the Π-S/D transistor has a smaller $R_{S/D}$ compared to a conventional UTSOI transistor, which implies that the SDT added to the UTSOI structure helps increase the drain current. Thus, the additional SDT mainly contributes the $R_{S/D}$ value. Apparently, small $R_{S/D}$ leads to high drain current.

(a)

(b)
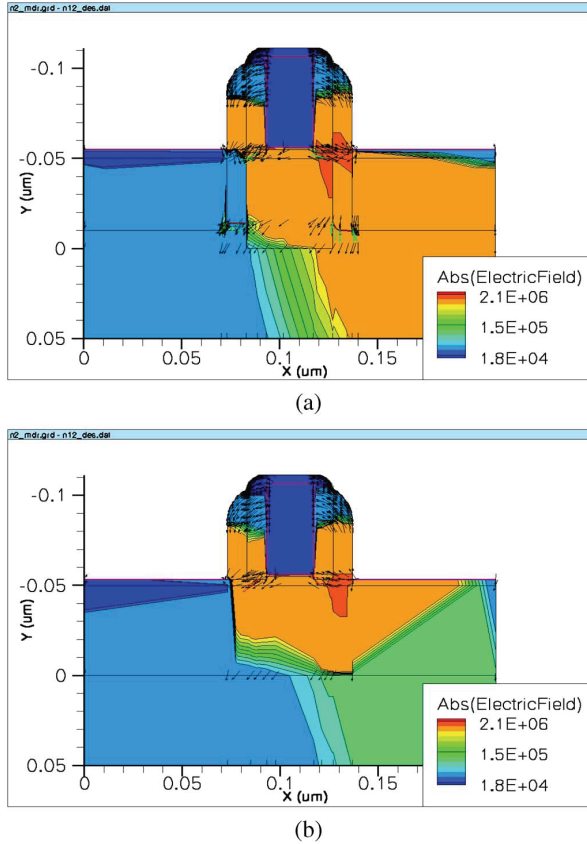
Fig. 6. Contour plot of simulated electric field (in volts per centimeter) at V_{DS} = 4.0 V and V_{GT} (V_{GS}–V_{TH}) = 1.0 V for n-channel UTSOI MOSFET w/ and w/o additional SDT. L_G = 24 nm.
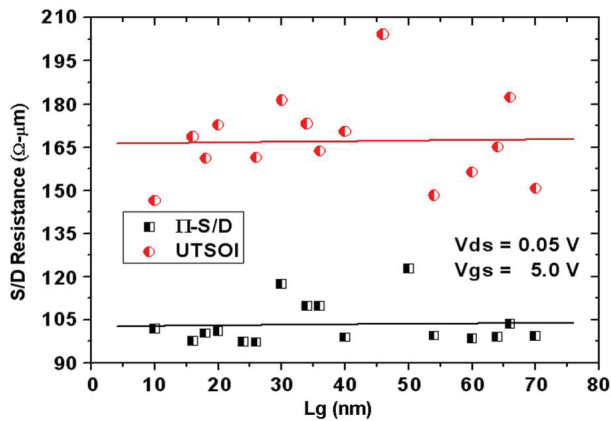
Fig. 7. Effects of the L_G on the $R_{S/D}$ characteristics of the UTSOI MOSFET w/ and w/o additional SDT.

Self-heating effects in π-S/D transistors

In order to investigate the thermal behavior of the Π-S/D and UTSOI devices, the curves in Fig. 8 compare the drain current I_{DS} versus drain-to-source voltage V_{DS} for various values of gate overdrive voltage V_{GT}. When the V_{GT} increases from 0.2 V to 1.2 V, the drain-source saturation current I_{DS} also increases in both types of transistors. In addition, a self-heating induced negative differential output conductance is observed for only the UTSOI-FET. The electron mobility decreases when the local lattice temperature increases due to effects of self-heating. The SDT is shown to overcome self-heating issues. The Π-S/D structure not only obtains high I_{DS} but also reduces the **self-heating effects** (SHEs). An interesting observation is that the reliability of the UTSOI MOSFET can be improved by the addition of an SDT.

(a)

(b)
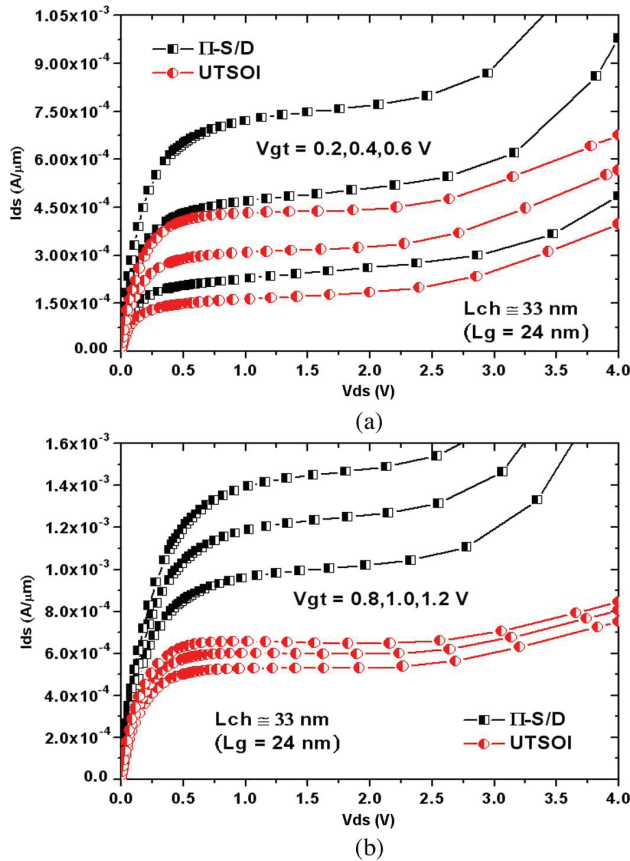
Fig. 8. Comparison of I_{DS}-V_{DS} curves between two transistors Π-S/D and UTSOI.

To probe the physical mechanisms involved for improved thermal performance of the Π-S/D structure, electron velocity and lattice temperature profiles for the Π-S/D and UTSOI are shown in Fig. 9. It should be noticed that the generated electron-hole pair will flow through the SDT, leading to a symmetric lattice temperature near the edges of both the source and drain regions. Moreover, this is due to the fact that the Si channel is thin enough,

the generated hole carriers can flow into the ground terminal only through its source region, resulting in symmetric lattice temperature near the edges of both the source and drain regions. Since the SDT exists only in the π-S/D NMOS, the SDT is to construct additional pathways to link Si substrate, which helps diminish the SHEs caused by the thermal vibrations. The two additional pathways can quickly disperse the heat generation in Si body, resulting in a higher electron velocity and better G_M-V_{GS} characteristics, as shown in Fig. 10. For a UTSOI MOSFET, the mobility decreases as the lattice temperature increases; this implies that the reduced electron velocity and decreased transconductance are inevitable due to self-heating.
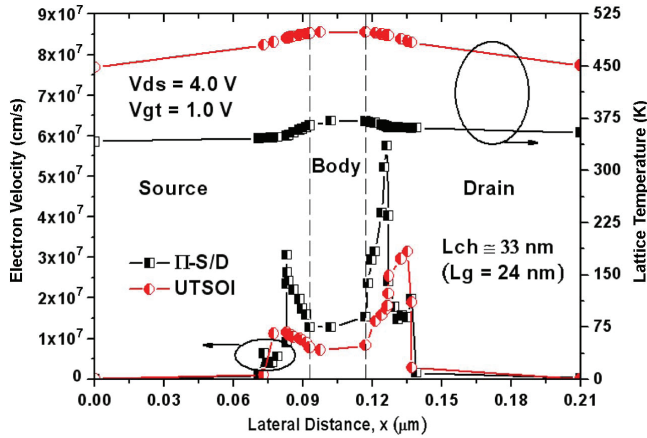
Fig. 9. Comparison of electron velocity and lattice temperature profiles between two transistors π-S/D and UTSOI.
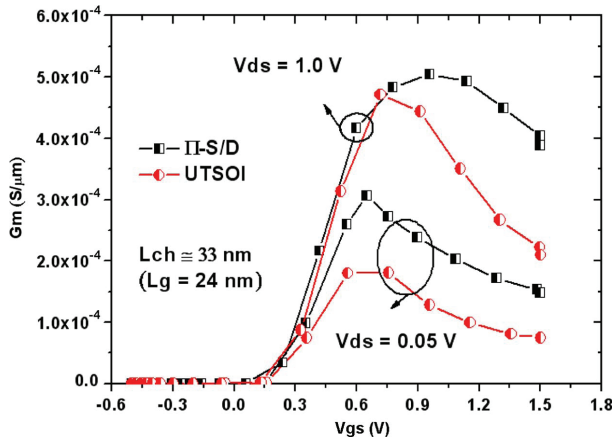
Fig. 10. Comparison of G_M-V_{GS} curves between two transistors π-S/D and UTSOI.

In this chapter, we demonstrate a new self-aligned π-shaped S/D UTSOI MOSFET that reduces device self-heating but without losing the desirable electrical characteristics. According to simulation results, we find that although the π-S/D structure appears to be less advantageous in terms of the charge sharing between the gate and the S/D diffusion

regions, the source-drain current is enhanced. Additionaly, the thermal stability of the Π-S/D NMOS are improved because the additional SDT increases the heat conductin area.

4. Summary

A novel UTSOI with SDT MOSFET (Π-S/D transistor) is proposed, in order to reduce self-heating errors. A path from the S/D to the Si substrate is created and called SDT that which does not significantly degrade the UTSOI MOSFET characteristics due to UTB usage.

Self-heating can be reduced greatly due to the presence of the SDT. The heat generated by thermal vibration of the atoms can be quickly dissipated via SDT. Furthermore, the short-channel characteristics of **fully depleted** (FD) SOI MOSFET with SDT, such as DIBL and S.S., are not significantly degraded or impacted because the BOX layer is directly under the UTB channel region.

5. References

[1] Yi-Chuen Eng, Jyi-Tsong Lin, Po-Hsieh Lin, Hau-Yuan Huang, Shiang-Shi Kang, Kung-Kai Kao, Jeng-Da Lin, Yi-Ming Tseng, and Ying-Chieh Tsai, "Self-aligned Π-shaped Source/Drain Ultra-thin SOI MOSFETs," in *Proc. 26th Int. Conf. Microelectronics*, p.499, 2008.

[2] Jyi-Tsong Lin, Yi-Chuen Eng, Hau-Yuan Huang, Shiang-Shi Kang, Po-Hsieh Lin, Kung-Kai Kao, Jeng-Da Lin, Yi-Ming Tseng, Ying-Chieh Tsai, and Hung-Jen Tseng, "Short-Channel Characteristics of Self-Aligned Π-Shaped Source/Drain Ultrathin SOI MOSFETs," *IEEE Trans. Electron Dev.*, ED-55, 1480 (2008).

[3] International Technology Roadmap for Semiconductors. (2007). [Online].   Available: www.itrs.net.

[4] Donald A. Neamen, *An Introduction to Semiconductor Devices*, 1st ed.   New York: McGraw-Hill, 2005.

[5] Toshiro Hiramoto and Toshiharu Nagumo, "Multi-Gate MOSFETs with Back-Gate Control," in *Proc. IEEE Int. Conf. Integer. Circuit Des. Technol.*, p.80, 2006.

[6] User's Manual, 2004. ISE-TCAD.

Accurate LDMOS Model Extraction using DC, CV and Small Signal S Parameters Measurements for Reliability Issues

Mouna Chetibi-Riah[1], Mohamed Masmoudi[1], Hichame Maanane[2], Jérôme Marcon[1], Karine Mourgues[1], Mohamed Ketata[1] and Philippe Eudeline[2]

[1]*University of Rouen*
[2]*THALES AIR SYSTEMS*
France

1. Introduction

It is well recognized that excellent power performance and linearity can be achieved at low cost using Laterally Diffused Metal-Oxide-Semiconductor (LDMOS) field effect transistor. In fact, it is the preferred technology for base station applications as well as many other RF and microwave applications (Wood et al., 2006). But system's availability and reliability are important parameters in terms of ownership cost for the customer. That's the reason why a precise knowledge of the device's degradation mechanisms and lifetime is of paramount importance. Thus modelling and reliability study of the LDMOS technologies are being increasingly used by the power amplifier design community.

Numerous applications for power integrated circuits are emerging, which needs to operate at high temperatures. In the radar field, a crucial issue to tackle with is the reliability of RF LDMOS subjected to RF pulses with high drain-source DC bias for maximum output power under wide temperature range (Maanane et al., 2006). Those requirements raise the stress (thermal, electrical and RF power) applied to the transistors and have a direct impact on their lifetime. A deep understanding of this impact is necessary for better device and radar module reliability assessment. Moreover, a study has been engaged to elaborate new methods for RF power device reliability investigations under pulsed RADAR conditions (Maanane et al., 2004).

In this chapter, we present an innovative reliability bench designed and implemented in our laboratory by (Maanane et al., 2004). It is specifically dedicated to high RF power device lifetime tests under pulse conditions for radar application and able to keep track of RF powers, voltages and temperatures whose values correspond to stress operating conditions. It clearly appears the need to track electrical parameters that lead to modifications of both the device RF performances and the critical electrical parameters with time (Poole & Walshak, 1974, Sirenza Microdevices, 2002). In this work, we will go step by step through the individual characterization issues and develop the model parameters extraction strategies which will provide the base for accurate device modelling. A commercial RF LDMOS transistor has been chosen for RF life-tests and a complete device electric characterization, I-V, C-V and S parameters, before and after test ageing has been

conducted. RF LDMOS modelling and parameters extraction are performed before and after life test in order to have a better insight on impact of the stress tests. With this intention, a refinement to the electro-thermal MET LDMOS model is presented, including the inductances in series as parasitic components in the extrinsic circuit. A methodology for the accurate extraction of model parameters is developed. The LDMOS model is compared to DC and small signal measurements. It is shown that the DC and RF performances of the device model closely match the measured data. The extracted non-linear model will be used as a reliability tool in order to correlate RF LDMOS electrical parameter drifts with any kind of degradation phenomenon, after different life-tests (DC and/or RF life-tests). Thus, a whole review of critical electrical parameters of the device is exposed and analysed. All the electrical parameters (P_{OUT}, I_{DSS}, C_{RSS}, R_{DSON}, f_t, f_{max}, G_{max}, etc.) drifts after accelerated ageing tests have been shown and discussed. RF-figures of merit such as f_t, f_{max} and power gain of RF LDMOS show significant vulnerability to the self heating and hot-carrier effects. This study clearly explains the physical degradation mechanism occurred during RF life-tests, by means of 2D ATLAS-SILVACO simulations. According to the measurement and simulation results analysis, an observation can be made. At low temperature, the drift of significant parameters is more important. Finally, we have clearly demonstrate that N-LDMOS degradation mechanism, therefore, is self-heating effect and hot carrier generated interface states (traps) and trapped electrons, which results in a build-up of negative charge at Si/SiO_2 interface. Last but not least, more interface states (i.e. important build-up of negative charge) are created at low temperature due to a located maximum impact ionisation rate at the gate edge (drain side). This is the reason why the electrical degradations are so high at low temperature.

2. Innovative reliability bench

An innovative reliability bench designed to apply both electrical and thermal stress (Fig.1) has been implemented. The bench capacity has been limited to eight devices to be tested simultaneously for the simplicity of use. The bench is divided in three modules (Maanane et al., 2004):
- a microwave module,
- a control/command module driven by PC,
- one thermal module for each device (Peltier or high temperature module).

Peltier and high temperature modules give us the capability to cool or heat each device independently for optimal temperature regulation and maximum flexibility.

3. Life-test conditions

Life-tests are run under RF conditions using different temperature settings and a high drain-source voltage in order to get more power from the device for radar applications. The applied drain voltage is 10% higher than the one applied on dedicated radar bipolar devices having same breakdown voltage (75 Vmin). In fact, the real RF LDMOS device breakdown voltage has been measured to be about 87 V. This justifies the applied drain-source voltage value (44 V).

The discrete RF device is a commercial 10 W telecom dedicated transistor with a gate length equal to 0.8 µm and that operates in class-B at saturation. The parameters set for the tests are as following:
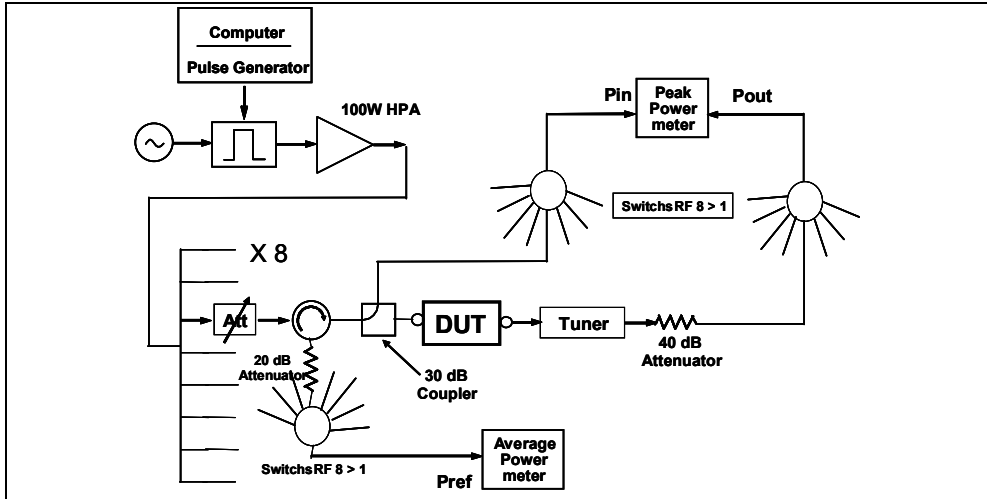
Fig. 1. Synoptic of RF power reliability bench in pulsed mode

Frequency =2.9 GHz,
Pulse width/ Duty cycle= 500 µs/50%,
Device base plate Temperature= 10°C, 80°C, 110°C and 150°C.

The RF transistor (16 samples) has been subjected to a 1500 hours ageing test on the reliability bench. Table 1 resumes the other operating conditions of the device.

Device base plate temperature (°C)	P_{IN} (dBm)	P_{OUT} (dBm)	P_{REF} (dBm)	I_{DQ}@ device base plate temperature (mA)	I_{DSS} during RF pulses (mA)
10	30.5	43.8	23.9	1.76	557.89
80	30.5	43.57	22.2	5.023	550
110	30.5	42.6	17.3	7.5647	537.68
150	30.5	40.2	22.2	13.348	500

Table 1. Summary of the device operating conditions

4. Electrical characterizations

It is crucial to characterize all devices in static and dynamic mode in order to extract the critical electrical parameters before and after ageing. This work should allow us to correlate RF ageing to any significant parameter drift and understand better the degradation phenomena, particularly those linked to hot carrier injection (Burger & Gola, 2002). I-V, C-V and S-parameters measurements were performed. Thanks to the commercial software package IC-CAP (Sischka, 2001). In addition, the cross-section of the RF LDMOS (see Fig. 2a) device used in this study was implemented and simulated with ATLAS-SILVACO in order to explain qualitatively electrical parameter shifts (Cortés et al., 2005, Brisbin et al., 2005, Silvaco, 1998). The structure is a modified 2D RF power N- channel LDMOS structure, previously developed by (Raman et al., 2003), with a Gaussian doping profile along LDD and channel surface, see also Fig. 2b.
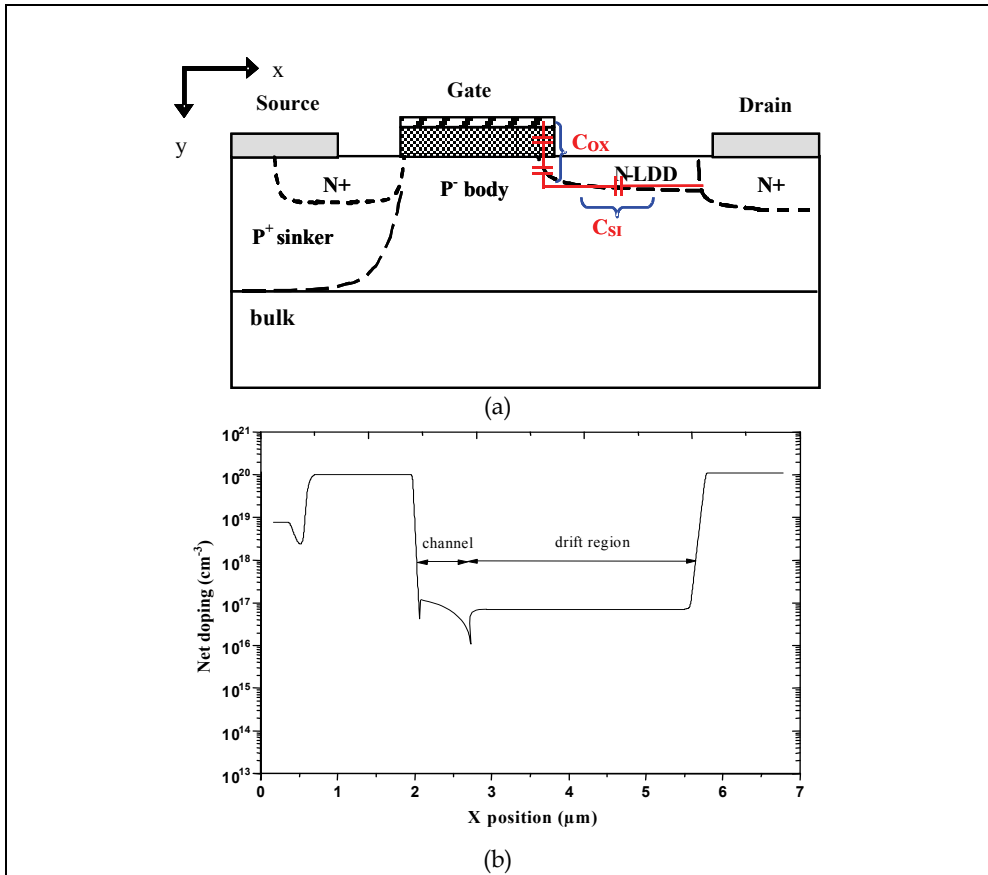
Fig. 2. (a) RF LDMOS device cross section implemented in ATLAS-SILVACO, with its intrinsic capacitances. (b) Net doping profile along silicon surface

4.1 I-V characterization

This kind of measurements gives us an insight into the device behaviour in its various operating mode (linear and saturated), and allows to quantify some important electrical parameters before and after device ageing. Static measurements were performed with the help of a DC analyser AGILENT E5270 with 20W power supply (see Fig. 3 and Fig. 4). The device was mounted on a Peltier module in order to stabilize the self-heating during DC measurements. The use of data management under IC-CAP allows us to check the consistency of measurements (Sischka, 2002). From these measured data and by model simulation, a set of critical electrical parameters can be deduced, strongly correlated to a specific area of RF LDMOS structure: the on-state resistance (R_{DSON}), the drain source current in saturation mode (I_{DSAT}) and the threshold voltage (V_{TH}) (Moens et al., 2004).

These electrical parameters are correlated with the device performances and will be an indicator of the electrical device degradation state at any moment of the ageing (Brisbin et al., 2005, Moens et al., 2004, Versari & Pierraci, 1999, Nigam et al., 2004).
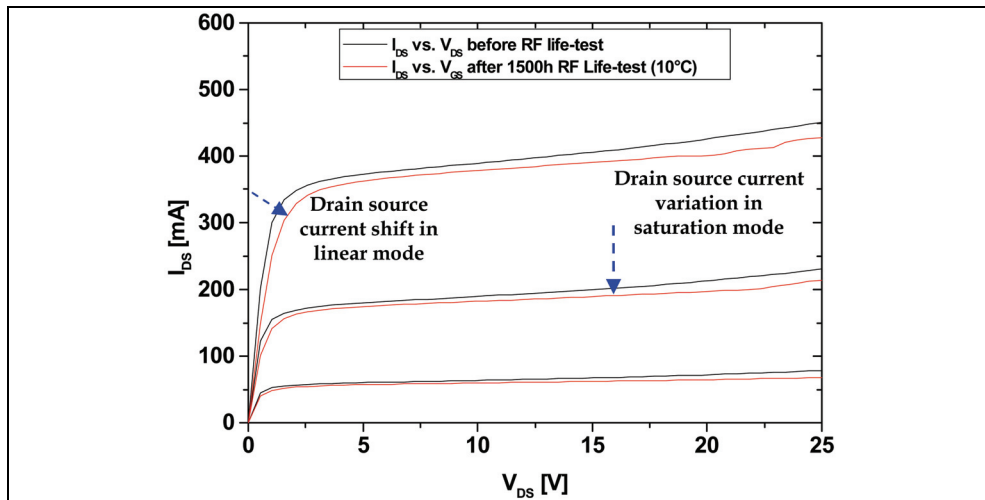
Fig. 3. Isothermal measurements of RF LDMOS output characteristics before and after 1500 h RF Life-test (10°C)
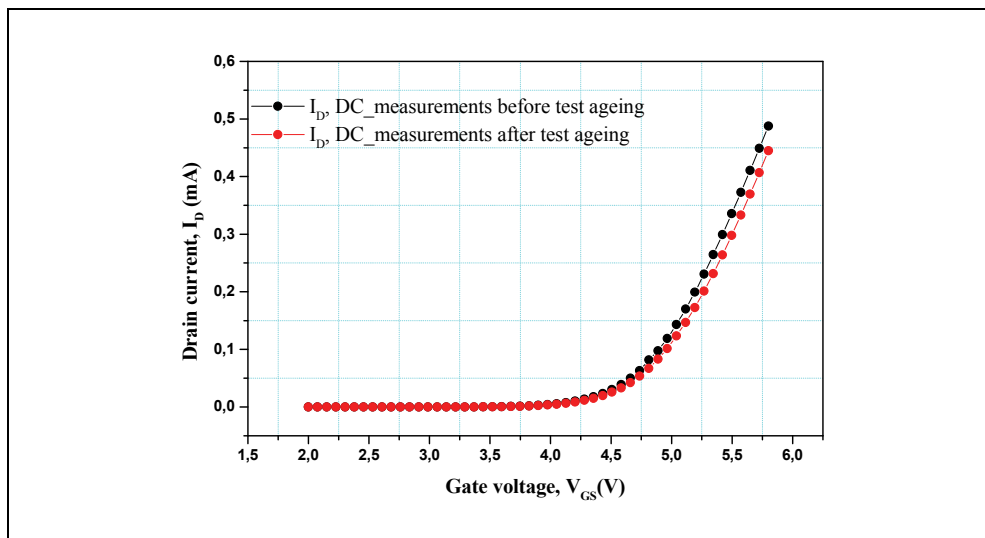
Fig. 4. The device isothermal I_{DS}-V_{GS} characteristics from DC measurements, before and after 1500 h RF Life-test (150°C), with V_{DS} = 20 V

4.2 C-V characterization at 1 MHz

After such a DC characterization, so-called C-V (Capacitance versus Voltage) measurements were performed, by using an HP 4194A impedance analyser, in order to characterize the device capacitances at a standard frequency of 1 MHz. This frequency is high enough to allow a resolution down to a few fempto-Amperes, yet still low enough to neglect second order parasitic like series resistances with the capacitances, or like inductances. Thus we

have take into account this effect in modelling. We have used a 2-pin method and the third pin of the transistor being unconnected (Sischka, 2002), which yields the total capacitance between the measurement pins: C_{RSS}, C_{ISS} and C_{OSS}, and Guarded measurements for inter electrode capacitances: C_{GS}, C_{GD} and C_{DS}.

4.3 Small signal S-parameters characterization

S-parameters were measured at room temperature from 490 MHz up to 5 GHz using an Agilent E8362B Network Analyser. During the measurements of S-parameters we carry out the DC characterization (the output and transfer characteristics) of the component in order to solicit the self-heating effects. Fig. 5 presents the transfer characteristics $I_{DS}\_V_{GS}$ measured with the small RF signal, at room temperature, before and after the ageing test.
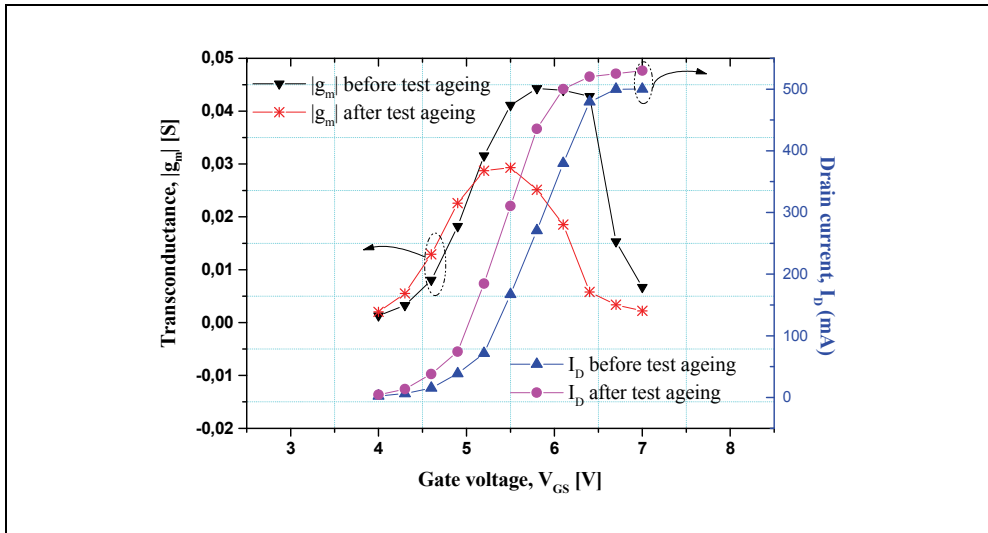
Fig. 5. Device transconductance and drain current from S-parameter measurements as a function of the gate voltage before and after life test at 150 °C with $V_{DS} = 20$ V

5. Model extraction

The new large signal equivalent circuit of RF LDMOS transistor used in this paper is shown in Fig. 6. The model includes new parasitic elements, series inductances, for modelling parasitic effects of the transistor topology (Chetibi-Riah et al., 2008). Self-heating effect is treated with a special circuit as shown in Fig. 6. The model has four ports (Motorola, 1999), with the extra port used for measuring the rise in temperature. Efficient and systematic extraction procedures have been developed and implemented in Agilent technologies IC-CAP software using a Symbolic Defined Device (SDD) into IC-CAP's ADS circuit page (Sischka, 2001). Currents parameters and the model capacitances are extracted from I-V and C-V measurements data respectively, using an optimisation program implemented in ICCAP to superimpose the simulated and the measured curves. The external inductances are used in addition to better fit measured S-parameter data, after the extraction of the intrinsic model elements and bias-dependent capacitance functions. This is done by using

the developed ICCAP routine to manipulate S-parameter data taken at many bias conditions. The current source in the thermal circuit is equal to the instantaneous power dissipated in the FET. The voltage between the external thermal circuit port and the source node in Fig. 6 is equal to the junction temperature rise (Yang et al., 2001) and the resistance R_{TH} is numerically equal to the thermal resistance. The RC product of the thermal circuit is the thermal time constant. A new DC technique for the extraction of thermal resistance of LDMOS transistors was applied (Menozzi et al., 2005). It is based on the DC measurements of the I–V output curves at different ambient temperatures. The thermal time constant (τ) measurement is based on the decrease in the transistor output current, for a sufficient long pulse biasing. The thermal capacitance (C_{TH}) is determined by:

$$\tau = R_{TH}.C_{TH} \tag{1}$$
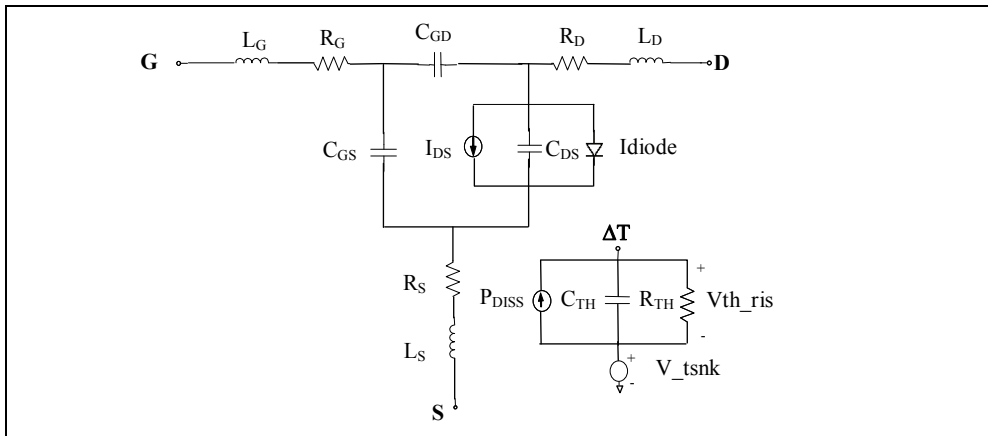
Fig. 6. RF LDMOSFET large-signal equivalent circuit taking account of thermal effect

5.1 Modelling results
5.1.1 Extraction of the R_{TH} of 0.8 µm 10 W LDMOSFET

The first step toward the modelling of self-heating effects is the determination of the device thermal resistance R_{TH}, which links the channel temperature T_C to the DC power dissipated by the device (P_D) through the simple equation:

$$T_C = T_A + R_{TH}.P_D \tag{2}$$

T_A is the ambient temperature. We consider a bias point in the saturation region, defined by an ambient temperature T_{A0}, a gate–source voltage V_{GS0}, a drain–source voltage V_{DS0}, a drain current I_{D0}, and a corresponding channel temperature T_{C0}. If we increase the ambient temperature T_A above T_{A0}, and assume that there is a linear dependence of the drain current on T_A, we can write (Menozzi et al., 2005):

$$I_D(V_{DS0}, T_A) = I_{D0}.(1 + h.(T_A - T_{A0})) \tag{3}$$

and

$$\frac{dI_D(V_{DS0},T_A)}{dT_A} = I_{D0}.h \tag{4}$$

The parameter h can thus be extracted by plotting $I_D(V_{DS0}, T_A)$ as a function of T_A, and taking the slope of the linear regression line.

We also assume that in the same T_A range, the drain current is a linear function of the channel temperature T_C and we write (Menozzi et al., 2005):

$$I_D(V_{DS0},T_C) = I_{D0}.\left(1 + h^{'}.(T_C - T_{C0})\right) \tag{5}$$

$$\frac{1}{h} = \frac{1}{h^{'}} - R_{TH}.P_D(V_{DS},T_{A0}) \tag{6}$$

Thus, a plot of $1/h$ versus $P_D(V_{DS}, T_{A0})$ should yield a straight line, from the slope of which we deduce R_{TH}.
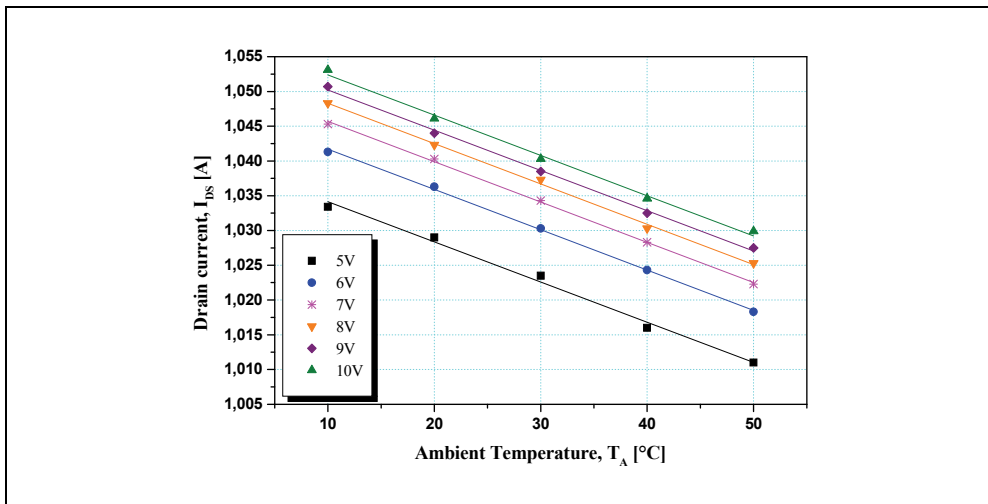
Fig. 7. Drain current of the 0.8 µm, 10 W LDMOS transistor measured at VGS = 7 V as a function of the ambient temperature T_A for different values of V_{DS} (legend). Best fit interpolations (solid lines)

Fig. 7 shows the measured drain current (I_D) as a function of the ambient temperature T_A for different values of drain–source voltage (V_{DS}) at a gate–source voltage V_{GS} = 7 V. The V_{DS} range used for R_{TH} extraction varies from 5 to 10 V and the ambient temperature T_A ranging from T_{A0} = 10 to 50°C. Since, this method assumes the thermal resistance to be constant across the range of T_A and V_{DS} values used in the extraction. Best fit interpolations (solid lines) show that the linearity assumption of (2) holds good in the chosen temperature and voltage ranges.

In Fig. 8, the reciprocals of the values of h show a linear dependence on $P_D(V_{DS}, T_{A0})$ = V_{DS}. $I_D(V_{DS}, T_{A0})$, T_{A0} = 10 °C, as predicted by (5). The absolute value of the slope of the linear regression line is therefore the device thermal resistance R_{TH} = 5.34 °C/W.
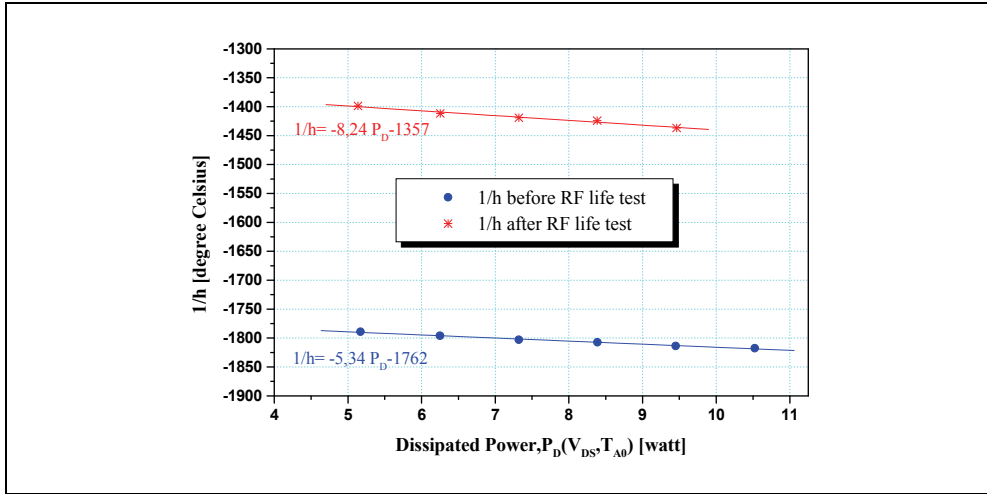
Fig. 8. Extraction of R_{TH} of the 0.8 µm and 10 W LDMOS transistor before and after RF life test (the solid line is the linear best fit, the slope of which yields R_{TH})

5.1.2 Extraction of current and capacitances parameters of the model

The extraction of current and capacitance parameters of the model is performed from I-V and C-V measurements data respectively, using an optimization program implemented in ICCAP to fit simulations with measurements. Small-signal simulations are performed to validate the model for small signal operation (Chetibi-Riah et al., 2008). Fig. 9 shows good agreement between small-signal S-parameters simulations obtained from the new model (with/without inductances) and the small-signal measured data. The extracted parasitic inductances L_G, L_D and L_S with the three dependent parasitic resistances R_G, R_D and R_S are presented in the Table 2.
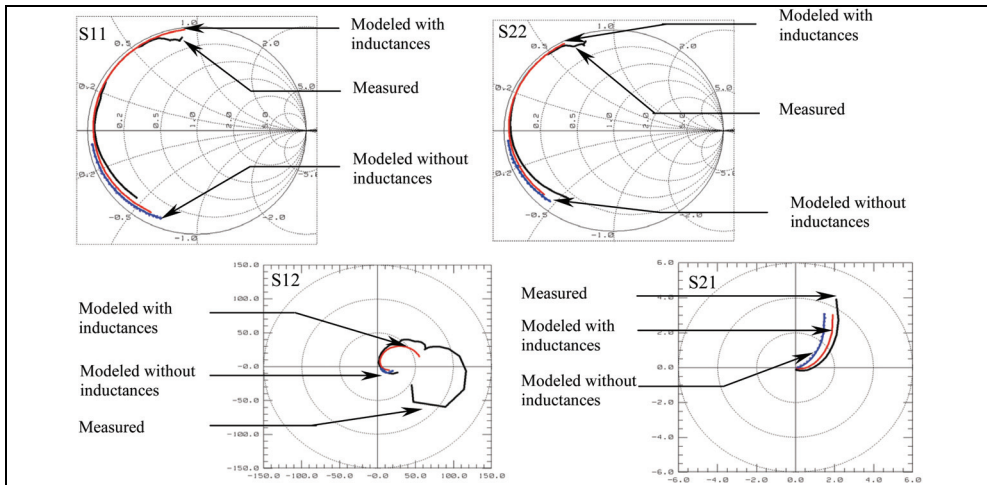
Fig. 9. Measured and modelled S-parameters before and after refinements with frequency = [0.5, 5] GHz and operating at class B

Parameter	Value	Unit
L_G	1.459	nH
L_D	970	pH
L_S	75	pH
R_G	795	mΩ
R_D	700	mΩ
R_S	100	mΩ

Table 2. Extracted parasitic elements

5.2 The new main significant RF LDMOS figures of merit

The threshold voltage, the breakdown voltage, the saturation current… are the classical parameters used as indicator in a reliability study. Thanks to the modelling approach, we can also extract other relevant parameters such as the power gain, the cut-off frequency (*ft*), the maximum oscillation frequency (*f_{Max}*), … which can be used to study the device reliability.

5.2.1 Current gain and cut-off frequency

The current gain is defined from the S-parameters as:

$$\left|H_{21}\right| = \left|\frac{-2.S_{21}}{(1 - S_{11})(1 + S_{22}) + S_{12}.S_{21}}\right| \tag{7}$$

The cut-off frequency represents the frequency for which the current gain is equal to 1 (Fig. 10).
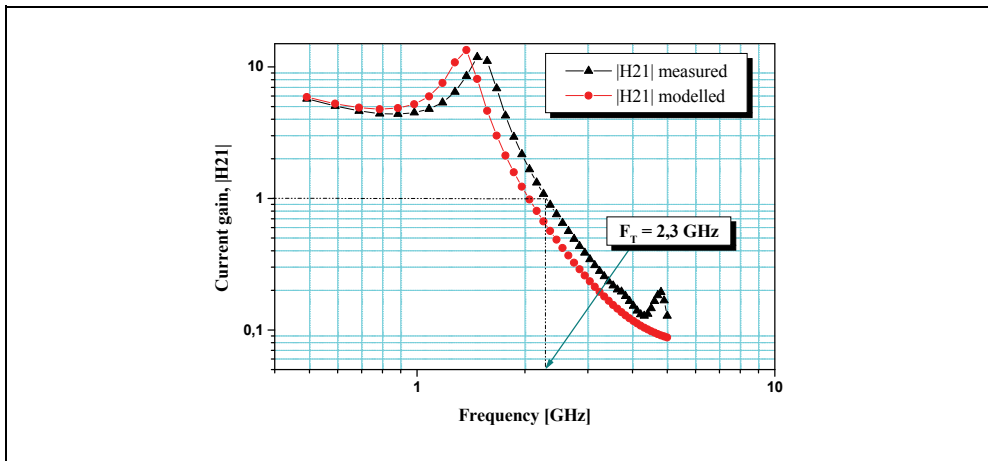
Fig. 10. Current gain and cut-off frequency for operation in class B

5.2.2 Maximum available gain and maximum oscillation frequency

The expression of the maximum available gain from the transistor S-parameters is given by:

$$G_{Max} = \frac{S_{21}}{S_{12}}\left(K - \sqrt{K^2 - 1} \right)$$

(8)

where K is the stability factor of Rollet.

The maximum oscillation frequency represents the frequency for which the maximum available gain is equal to 0 dB (Fig. 11).
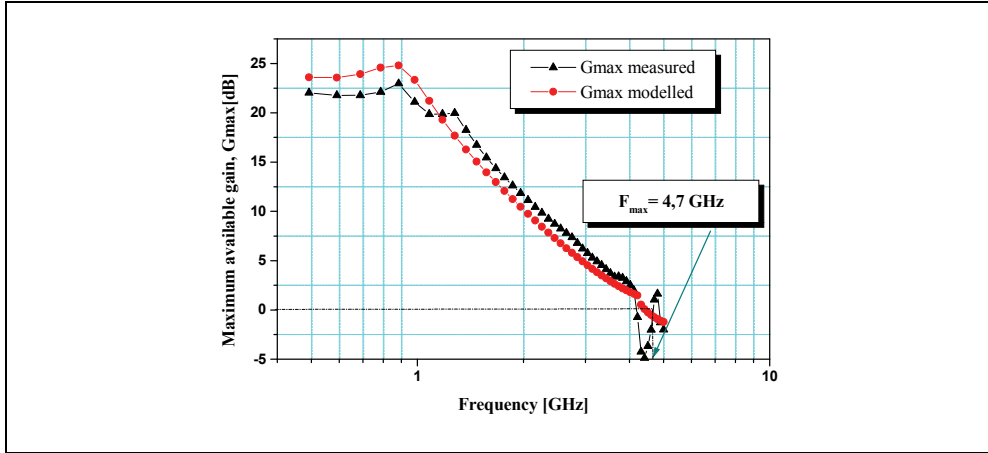
Fig. 11. Maximum available gain and maximum oscillation frequency for operation in class B

6. Experimental results

After RF life-tests, the degraded device under test was characterized at ambient temperature. A set of parameters is extracted and detailed as follows.

6.1 RF output power degradation

Two critical parameters are monitored during ageing tests. For high power devices working at saturation, the significant performance parameters concern output power and drain-source current (measured during RF pulse) (Brown et al., 2004).

Measurements were plotted under these two figures of merit and for four different temperature conditions, see Fig. 12 and Fig. 13.

The means for each parameter from the 24, 48, 168, 500, 1000 and 1500 h test down-points have been empirically fitted to log curves, as an overall trend seems to be log linear. Projecting the curves forward from 1500 hours 20 years, we find that I_{DSS} are expected to change less than 10% from 24 hours to 20 years, whatever the thermal conditions are. Similarly, P_{SAT} degradation is less than 1dB. On the other hand, in Fig. 12 and Fig. 13, it can be observed that the more the temperature decreases (10°C) the more the two parameter drifts increase. This is the invert situation under high temperature (150°C).

6.2 Evolution of I_{DS} after ageing (output characteristics)

Fig. 3 shows static measurements of the output characteristics ($V_{DS}=[0, 26V]$ and $V_{GS}=[2, 5.8V]$) after 1500 hours ageing at 10°C. At high V_{GS}, where the current is dominated by the
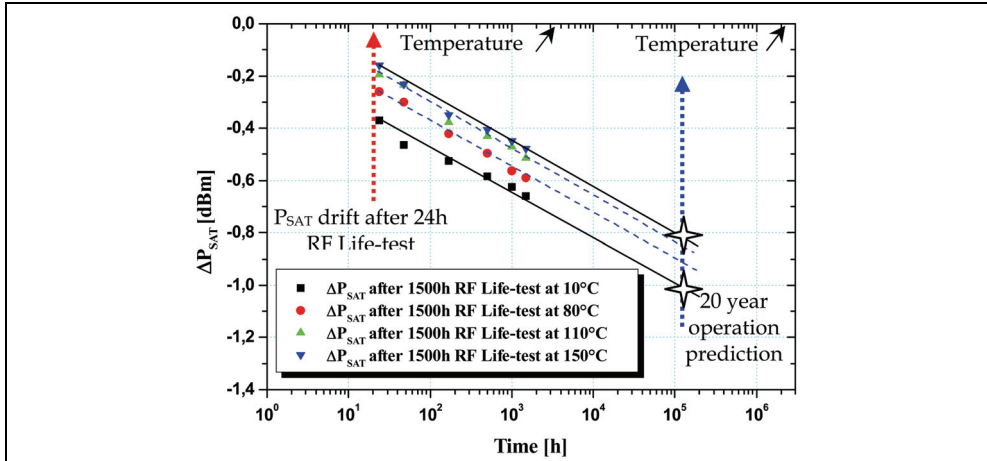
Fig. 12. RF saturated output power evolution over ageing time (1500 h) for various temperature conditions
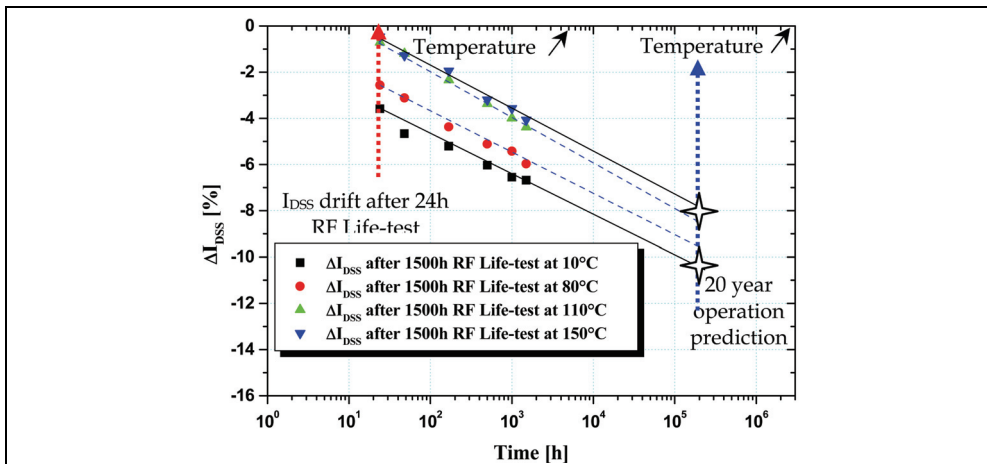
Fig. 13. Drain source current evolution over ageing time (1500 h) for various temperature conditions

drift region of the RF N-LDMOS transistor, a drop in I_{DSAT} can be observed, particularly at 10°C, but in a lesser extent at 150°C (see Table 3). An explanation could be that degradation has occurred in the drift area.

6.3 Evolution of C_{ISS}, C_{OSS} and C_{RSS} after ageing

Fig. 14 shows the input and output capacitance evolution after ageing at the lowest device base plate temperature. C_{ISS} did not drift during 1500 hours ageing. The output capacitance characteristic C_{OSS} experienced a slight shift (see Table 4), such a small change should not impact much the device behaviour. The feedback capacitance study can now be tackled. An interpretation is proposed to explain the discernable change observed on the feedback

	V_{TH} (V)	G_M (S)	I_{DSAT} (mA)	R_{DSON} (Ω)
Before RF Life-test	4.19	0.53	240	1.42
After 1500h RF Life-test at 10°C	4.14	0.52	200	1.74
Variation (%)	**-1.2**	**-1.9**	**-16.6**	**+22.5**
Before RF Life-test	4.21	0.53	238	1.36
After 1500h RF Life-test at 80°C	414	0.52	206	1.62
Variation (%)	**-1.6**	**-1.9**	**-13.5**	**+19**
Before RF Life-test	4.24	0.55	244	1.42
After 1500h RF Life-test at 110°C°	4.21	0.54	222	1.5
Variation (%)	**-0.7**	**-1.8**	**-9**	**+8.6**
Before RF Life-test	4.17	0.54	231	1.4
After 1500h RF Life-test at 150°C°	4.14	0.53	221	1.52
Variation (%)	**-0.72**	**-1.8**	**-4**	**+7**

Table 3. Summary of DC electrical parameter shifts after 1500h RF Life-tests

capacitance, once again more noticeable at 10°C (see Table 4). For LDMOS devices, the zero-volt feedback capacitance value is mainly due to the oxide capacitance (Pritiskutch & Hanson, 2000). Thus, C_{RSS} is defined by two capacitances in series, oxide capacitance and drift region capacitance (C_{SI}), the mathematic relation is the following (Xu et al., 1999):

$$C_{RSS} = \frac{C_{OX} \cdot C_{SI}}{C_{OX} + C_{SI}} \qquad (9)$$
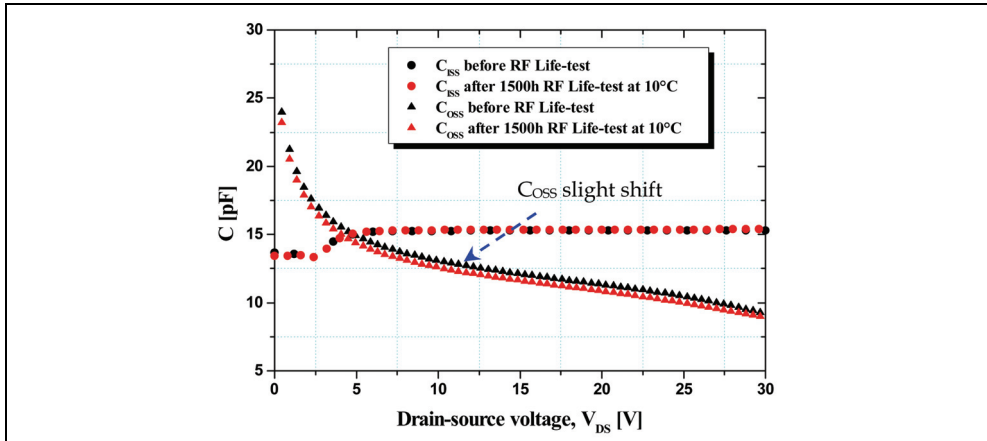
Fig. 14. C_{ISS} and C_{OSS} profiles before and after 1500h RF Life-test (10°C), with Freq= 1 MHz

Fig. 15 informs us about C_{RSS} behaviour after ageing. It can be reminded that C_{OX} is decided by the gate/N-LDD overlap area and the oxide thickness (see Fig. 2a) (Xu et al.,1999, Luo et al., 2003). After 1500 hours ageing, C_{OX} shows an important decrease after ageing (see Table 4). It is clear that a degradation mechanism is activated in the gate/N-LDD region. This capacitance value reduction is explained by the fact that the carriers (mainly electrons) flow in the presence of high field intensity peaks at the gate edge.

	C_{ISS} (pF) at 26 V @ 25°C	C_{OSS} (pF) at 26 V @ 25°C	C_{RSS} (pF) at 0 V @ 25°C	C_{RSS} (pF) at 26 V @ 25°C
Before RF Life-test	15,3	10,25	2,58	0,43
After 1500h RF Life-test at 10°C	15,36	9,77	1,82	0,34
Variation (%)	**+0,4**	**-4,7**	**-29,5**	**-21**
Before RF Life-test	15,3	10,1	2,52	0,46
After 1500h RF Life-test at 80°C	15,3	9,79	1,8	0,35
Variation (%)	**+0**	**-3**	**-28,6**	**-23,9**
Before RF Life-test	14,9	10,2	2,12	0,34
After 1500h RF Life-test at 110°C	14,9	9,9	1,68	0,27
Variation (%)	**+0**	**-3**	**-20,75**	**-20,58**
Before RF Life-test	15,3	10,1	2,52	0,44
After 1500h RF Life-test at 150°C	15,4	9,9	2,1	0,37
Variation (%)	**+0,65**	**-3**	**-16,6**	**-15,9**

Table 4. Summary of CV electrical parameter shifts after 1500h RF Life-tests
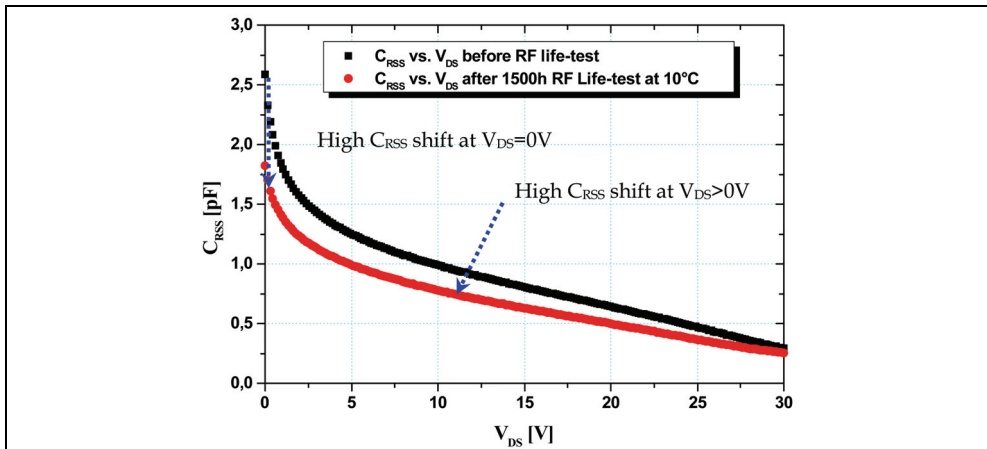
Fig. 15. C_{RSS} profiles before and after 1500h RF Life-test (10°C), with Freq=1 MHz

A detail of the lateral electric field distribution for RF LDMOS along the surface of the active silicon layer in channel and drift regions is shown in Fig. 16.

Hence, electrons are concentrated in silicon surface (see Fig. 17), in such a way that provides a rise of the surface current density near the gate edge. Consequently, many electrons are accelerated to high velocities by this high electric field peaks. They become highly energized and should be accelerated away from their normal directional flow. These highly energized electrons may create interface states by breaking Silicon bonds (Acovic et al., 1996) or be injected into generated surface traps (hot electron injection) at interface between gate oxide and N-LDD overlap area beneath SiO_2 layer. The trapped electrons reduce the electric charge density and therefore the total charge in the area affected by the trapped carriers.
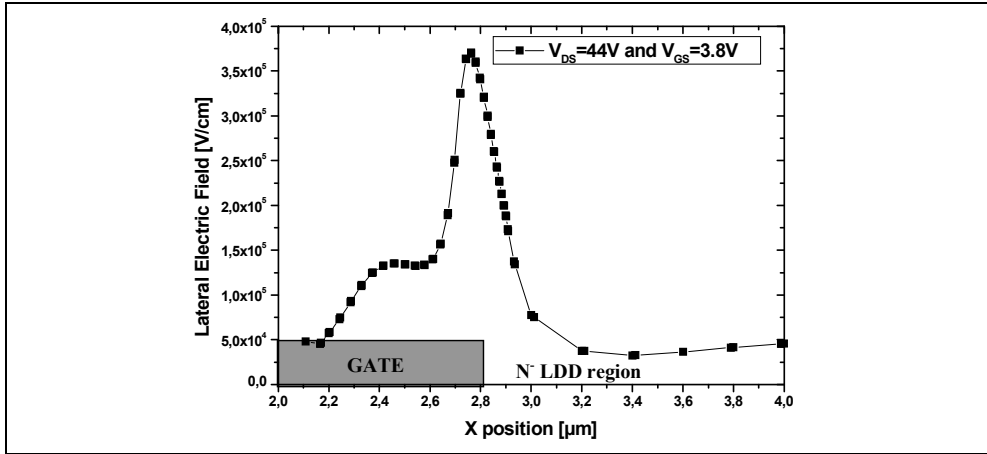
Fig. 16. Lateral electric field distribution in RF LDMOS structure
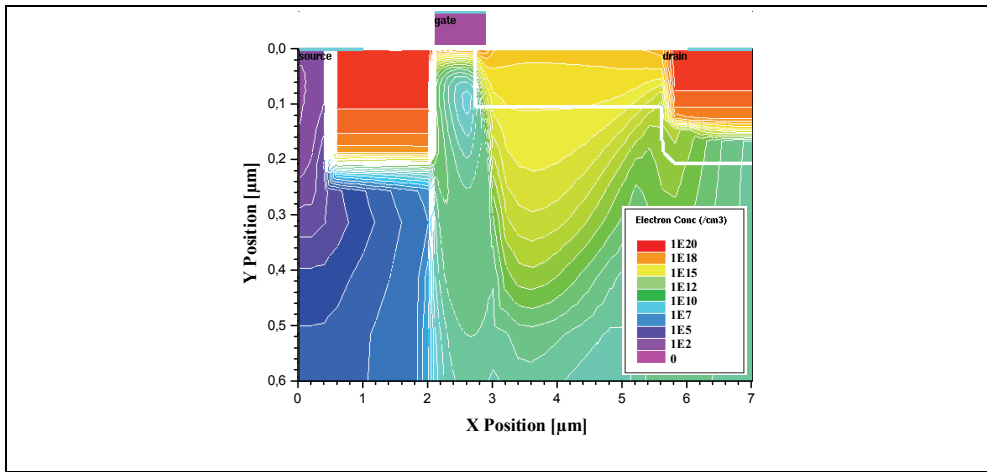
Fig. 17. The electron concentration along the silicon surface at Gate/N- LDD junction area for uniformly doped drift structure, biased at V_{DS}=44 V and V_{GS}=3.8 V

The latter probably changed C_{SI} value (see Fig. 2a and Eq.9) and as final consequence, the whole feedback capacitance characteristic (see Fig. 15) is shifted by the trapped charges. Hence, this shift is more remarkable at 10°C, due to the fact that the maximum impact ionisation rate is located near the gate edge (see Fig. 18a). So the interface trap density (at Si/SiO$_2$ region) is raised, thus increasing the probability of electrons being trapped.

The opposite situation can be observed at 150°C, where the maximum impact ionization rate is in the depth of the silicon material (see Fig. 18b). In this stage study, it would have been judicious to practice the charge pumping technique, which allows to monitor the amount of damage generated, but also the nature and location of the damage (Nigam et al., 2004). However it was not possible in our case, because of the substrate access absence (commercial packaged device). Finally, **C_{RSS}** can be considered as a very sensitive parameter to the electrons injected in the already existing SiO$_2$/N- LDD interface traps.

(a)

(b)
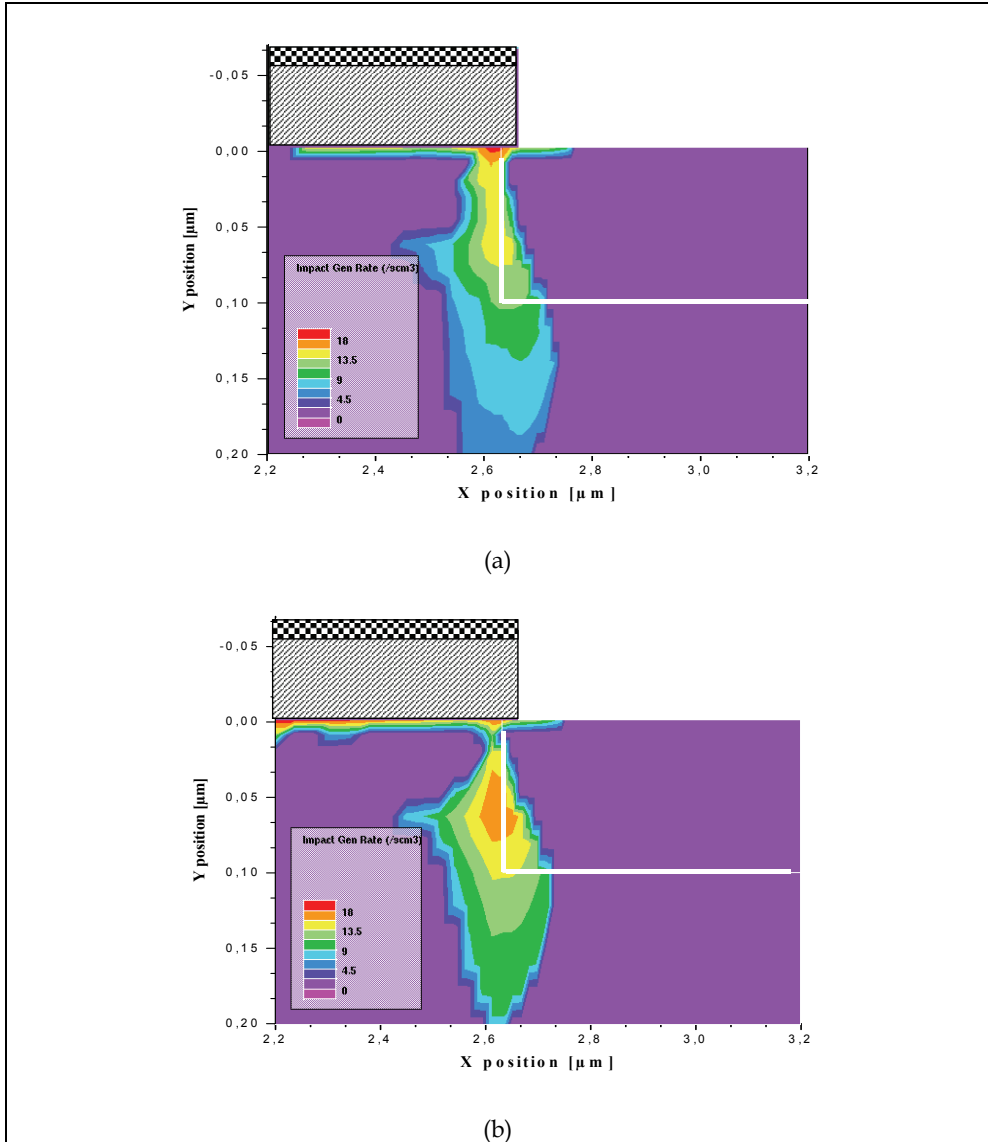
Fig. 18. Simulated impact ionisation rate contours in the gate/ N-LDD junction region for 10°C, at bias conditions (V_{DS}=44 V and V_{GS}=3.8 V)

6.4 Evolution of R_{DSON} after ageing

Fig. 19 shows the drain-source on-state resistance obtained by an output characteristic extrapolation (V_{DS}=[0, 2V] and V_{GS}=[4, 10V]). The measured R_{DSON} on-state resistance value at V_{GS}=7 V raised after 1500 hours ageing (see Table 3). This tendency is even more significant with a 10°C temperature stress.
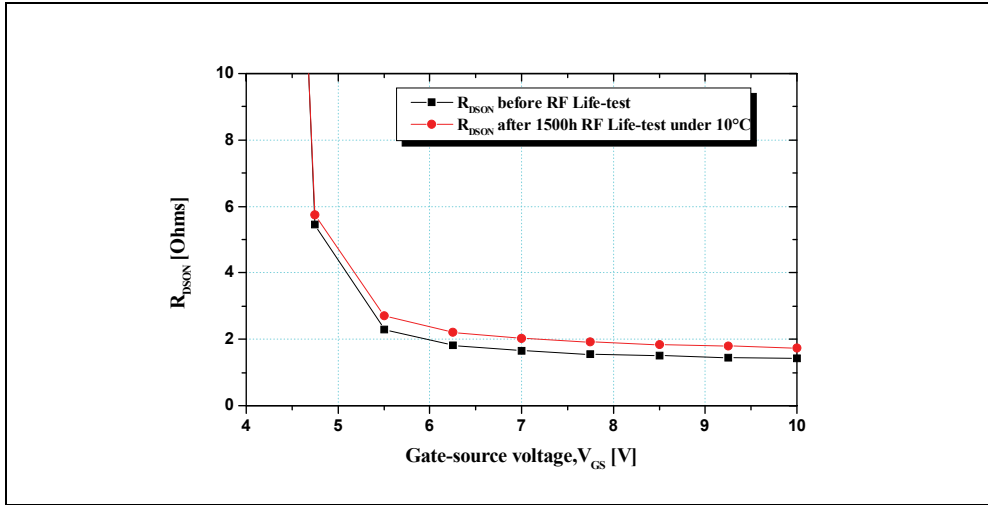
Fig. 19. Comparison between R_{DSON} before and after 1500h RF Life-test (10°C)

When carriers are trapped at the interface above N- LDD, it effectively changes charge concentration in N- LDD region, such that the overall Drain resistance (**R_{DSON}**) is increased. In other words, the hot carriers produce interface states (traps) and trapped electron charge, which results in a build-up of negative charge at Si/SiO$_2$ interface (Brisbin et al., 2005). The location of this negative charge is likely in the vicinity of the intersection of the impact ionisation (at 10°C) with the Si/SiO$_2$ interface as seen in Fig. 18a, which is not the case at 150°C (see Fig. 18b). This negative charge attracts holes depleting the negative charge in the LDMOS N- drift region increasing the R_{DSON} device resistance. Finally R_{DSON} reduces the Peak Current capability and therefore the RF Peak Power capability. All these aspects explain clearly why I_{DSS} and P_{SAT} drifts over the time are more significant at 10°C than 150°C (Fig. 12 & Fig. 13).

6.5 Evolution of R_{TH} after ageing
In Fig. 8, the reciprocals of the values of h after the RF life test show a good linear dependence on $P_D(V_{DS}, T_{A0}) = V_{DS} . I_D(V_{DS}, T_{A0})$, T_{A0} = 20 °C, as predicted by (5). The absolute value of the slope of the linear regression line is therefore the device thermal resistance after the test ageing *R_{TH} = 8.24 °C/W*. It is observed that R_{TH} increases after life test. We can explain this by the degradation of the device thermal performance. It justifies the increase of the current in Fig. 5.

6.6 Evolution of V_{TH} after ageing (transfer characteristics)
The transfer characteristics (V_{GS}=[2, 5.8 V] and V_{DS}=20 V) from the static measurements (without RF signal) after 1500h ageing at 150°C is shown in Fig. 4. A decrease of current is observed, but no significant drift of the threshold voltage parameter has been noted, whatever the device base plate temperatures are (see Table 3). It is known that the threshold voltage is strongly correlated to the drain quiescent current (Rice, 2002), and this last one do not show any drift during ageing. By consequent, the small V_{TH} shift (see Table 3) indicates

that hot carrier injection (hole or electron) into the gate oxide traps does not play an important role in the N-LDMOS degradation mechanism. Fig. 5 shows the transfer characteristics measured with RF small signal I_{DS}−V_{GS} (V_{GS}=[4, 7 V] and V_{DS}=20 V) before and after 1500 hours ageing at 150°C. The drain–current shifts upward due to the self heating of the transistor.

6.7 Evolution of g_m after ageing

The transconductance extracted from small signal S-parameters measurements is shown in Fig. 5. The RF power LDMOSFET suffers from a sharp falloff of the transconductance at high gate bias (De souza et al., 2007). It is observed that g_m increases after life test at low gate bias. We can explain this by the increase of the saturation velocity v_{max} of the MOS channel, due to the self heating (Taghi Ahmadi et al., 2007), which at low gate voltage dominates the transconductance given by :

$$g_{m,\max} = W \quad C_{OX} v_{\max} \tag{10}$$

where W is the width and C_{OX} is the oxide capacitance. Unlike that in the low-voltage, g_m decreases at high gate bias after the life test. We can explain this by the decrease in the oxide capacitance C_{OX} which dominates the transconductance at high gate bias (De souza et al., 2007).

6.8 Evolution of f_t, f_{max} and power gain after ageing

From the simplified small signal model of LDMOS transistor shown in Fig. 20, the cut-off frequency as a function of device parameters is given by (Yu et al., 2006):

$$f_t = \frac{g_m}{2\pi(C_{GD} + C_{GS})} \tag{11}$$

$$C_{GD} + C_{GS} \approx \frac{\mathrm{Im}(Y_{11})}{\omega} \tag{12}$$

where g_m is the transconductance. C_{GS} and C_{GD} are the gate–source and gate–drain capacitances. Y_{11} is an Y matrix element which can be obtained from S-parameter, and ω (ω =$2\pi f$) is the angular frequency (f is the operation frequency).

From (11) and (12) we have the following expression of f_t:

$$f_t = \frac{g_m \cdot f}{\mathrm{Im}(Y_{11})} \tag{13}$$
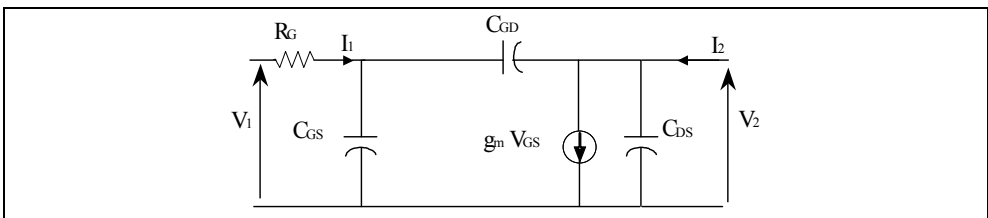
Fig. 20. Simplified small-signal model of LDMOS transistor

The imaginary part of Y11 as a function of the frequency before and after the test ageing is shown in Fig. 21. It is observed that electron trapping in the channel causes a slight increase in the $Im(Y11)$, therefore in the input capacitance C_{ISS} given by $C_{ISS} = C_{GS} + C_{GD}$. The main reason of a slight C_{ISS} rise is due to a reduction of the gate overlap capacitance as a result of enhanced depletion in the underlying drift region caused by trapped electrons (De souza et al., 2007).
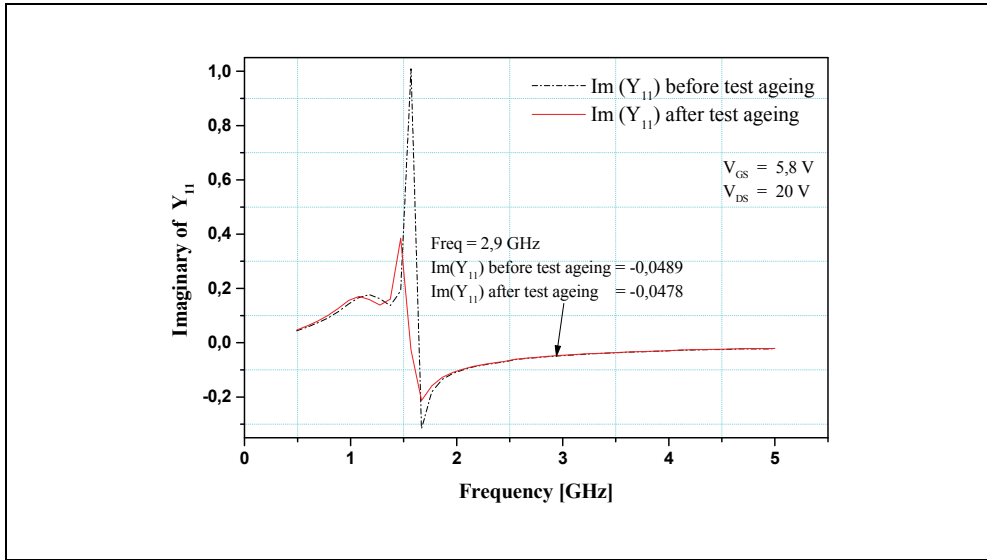
Fig. 21. Y parameter before and after life test at 150°C with V_{GS} = 5,8V and V_{DS} = 20V

The device cut-off frequency extracted from the RF measurement evolved as a function of the gate voltage is shown in Fig. 22 before and after life test. Establishing the correlation between f_t, g_m and $Im(Y11)$ given by equation (11) allows us to explain that the variation of f_t is dominated by the variation of g_m at low gate voltage. Fig. 22 compares f_{max} as a function of gate voltage for the device before and after test ageing at 150°C. It can be seen that the reduction in f_{max} is greater than the reduction in f_t. This can be understood by considering the approximation for f_{max}:

$$f_{max} = \sqrt{\frac{f_t}{8\pi R_G C_{GD}}} \qquad (14)$$

The decrease in f_t and the increase in C_{GD} and $R_{G,}$ shown by the aged model, all act to reduce f_{max}.

Fig. 23 compares the device power gain as a function of frequency (490MHz - 5GHz) before and after the test ageing at 150°C. It can be seen that the power gain performance degraded after the life test.
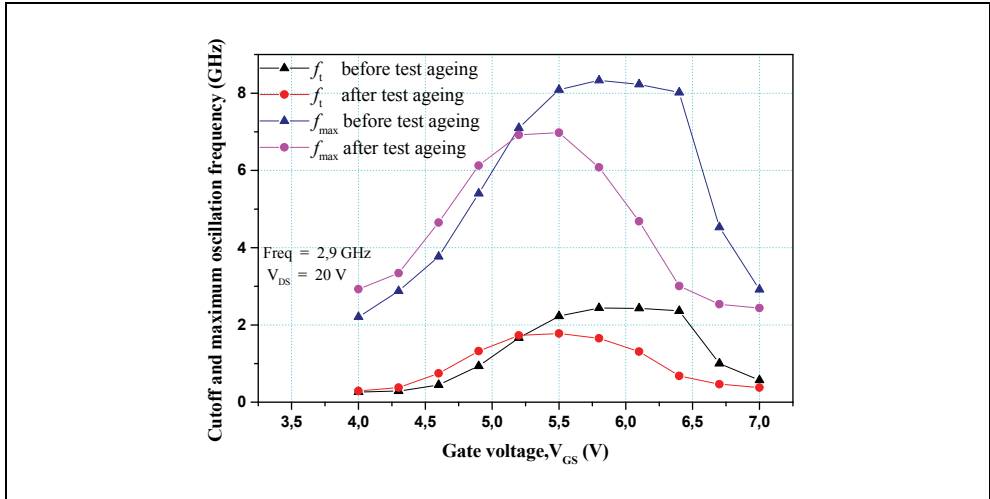
Fig. 22. The device cut-off frequency and maximum oscillation frequency as a function of the gate voltage before and after life test at 150°C with V_{DS} = 20 V
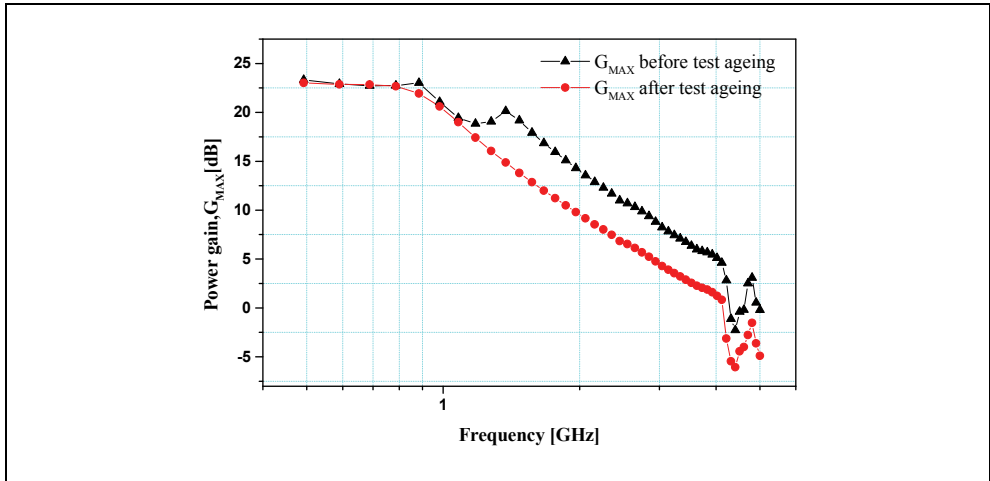
Fig. 23. The device Power Gain before and after life test at 150°C for V_{GS}=5.8 V and V_{DS}=20 V

7. Conclusion

In this work, we show the importance of electro-thermal modelling to study the reliability and the effect of self-heating and channel hot carrier on the DC and RF performances of silicon RF LDMOSFETs. An innovative RF test bench has been presented. The reliability was reviewed under microwave operating conditions. Then the critical parameters were put forward by linking them to the RF degradations (P_{SAT} and I_{DSS} in RF amplification) ones using 2D ATLAS simulations. This study essentially clarified the problems related with self heating, hot carriers and impact ionization under operating conditions met by the RF

LDMOS. Experimental data and simulations indicate that device degradation is mainly caused by the thermal resistance degradation. After RF accelerated temperature life test the thermal resistance increases, because of what, the self-heating effect is more important. To conclude, a focus was made on the electrons injected in Gate/SiO$_2$ interface traps, which have a strong influence on the feedback capacitance (**C$_{RSS}$**) and on state drain source resistance (*R$_{DSON}$*).

8. References

Acovic, A.; La Rosa, G. & Sun, Y.C. (1996), A review of hot-carrier degradation mechanism in MOSFETs, *Microelectronics Reliability*, Vol.36, (1996), (845-869)

Amerasekra, E.A. & Najm F.N. (1997), *Failure mechanisms in semiconductor devices*, John Wiley & Sons, (Ed.2), British library, 0 471 95482 9, New York

Brisbin, D.; Strachan, A. & Chaparala, P. (2005), Optimizing the hot carrier reliability of N-LDMOS transistor arrays, *Microelectronics Reliability*, Vol. 45, (2005), (1021-1032)

Brown, J.D.; Nagy, W.; Singhal, S.; Peters, S.; Chaudhari, A.; Li, T.; Nichols, R.; Borges, R.; Rajagopal, P.; Johnson, J.W.; Therrien, R.J.; Hanson, A.W. & Vescan, A. (2004), Performance of AlGaN/GaN HFETs fabricated on 100mm silicon substrates for wireless base station applications, *IEEE MTT-S Int. Microwave Symposium*, Dig., (2004), (833-836)

Burger, W. & Gola, P. (2002), Semiconductor technologies for RF power, *ESNDF* 2002

Chetibi-Riah, M.; Gares, M.; Masmoudi, M.; Maanane, H.; Marcon, J.; Mourgues, K. & Eudeline, Ph. (2008), RF Power LDMOSFET Characterzation and Modeling for Reliability Issues : DC and RF Performances, *Proceedings of International Conference on Microelectronics*, pp. 171-174, ISBN 978-1-4244-1881-7, Niš, May 2008, IEEE Serbia and Montenegro Section, Serbia

Cortés, I.; Roig, J.; Flores, D.; Urresti, J.; Hidalgo, S. & Rebollo, J. (2005), Analysis of hot-carrier degradation in a SOI LDMOS transistor with a steep retrograde drift doping profile, *Microelectronics Reliability*, Vol. 45, (2005), (493-498)

De Souza, M.M.; Fioravanti, P.; Cao, G. & Hinchley, D. (2007), Design for reliability: The RF Power LDMOSFET, *IEEE Transactions Device and Materials Reliability*, Vol.7, No.1, (March 2007), (162-174)

Luo, J.; Gao, G.; Ekkanath Madathil, S.N. & De Souza, M.M. (2003), A high performance RF LDMOSFET in thin film SOI technology with step drift profile, *Solid-State Electronics*, Vol.47, (2003), (1937-1941)

Maanane, H.; Bertram, P.; Marcon, J.; Masmoudi, M.; Belaid, M.A.; Mourgues, K.; Eudeline, P. & Ketata, K. (2004). Reliability study of power RF LDMOS for radar application, *Microelectronics Reliability*, Vol.44, (2004), (1449-1454)

Maanane, H.; Masmoudi, M.; Marcon, J.; Belaid, M. A.; Mourgues, K.; Tolant, C.; Ketata, K. & Eudeline, Ph. (2006). Study of RF N- LDMOS critical electrical parameter drifts after a thermal and electrical ageing in pulsed RF, *Microelectronics Reliability*, Vol.46, (2006), (994-1000)

Menozzi, R. & Kingswood, A. C. (2005), A New Technique to Measure the Thermal Resistance of LDMOS Transistors, *IEEE Transactions on Device and Materials Reliability*, Vol.5, No.3, (Sep 2005), (515-521)

Motorola (1999), *Motorola's Electro Thermal (MET) LDMOS Model*, on line: http://www.freescale.com/files/abstract/ldmos_models/MET_MODEL_DOCU MENT_0704.pdf

Moens, P.; Van den bosch, G.; De Keukeleire, C.; Degraeve, R.; Tack, M. & Groeseneken, G. (2004), Hot Hole Degradation Effects in Lateral nDMOS Transistors, *IEEE Trans Electron Devices*, Vol.51, No.10, (Oct. 2004), (1704-1710)

Nigam, T.; Shibib, A.; Xu, S.; Safar, H. & Steinberg, L. (2004) Nature and location of interface traps in RF LDMOS due to hot carriers, *Microelectronic Engineering*, Vol.72, (2004), (71-75)

Poole, W.E. & Walshak, L.G. (1974). Median-time-failure (MTF) of an L-band power transistor under RF conditions. *Reliability physics, Proceedings of the Twelfth International Symposium*, pp. 109-117, A75-13601 03-33, Las Vegas, April 1974, New York

Pritiskutch, J. & Hanson, B. (2000), Relate LDMOS device parameters to RF performance, *St Microelectronics*, Application note: AN 1228

Raman, A.; Walker, D. G. & Fisher, T. S. (2003), Simulation of nonequilibrium thermal effects in power LDMOS transistors, *Solid-State Electronics*, Vol.*47*, (2003), (1265-1273)

Rice, J. (2002), LDMOS linearity and reliability. (Technical Feature).(laterally diffused metal-oxide-semiconductors), *Microwave Journal*, (Jun 2002)

Silvaco International, (1998), *Atlas User's Manual-Device Simulation Software*, Santa Clara, California

Sirenza Microdevices, (2002). Bias drift in LDMOS power FETs-A primer, *Application Note*: AN049, 2002

Sischka, F. (2001), IC-CAP user's Guide 2001. *Agilent technologies,*

Sischka, F. (2002), Characterization and modelling handbook, Munich

Taghi Ahmadi, M.; Saad, I. ; L.P.Tan, M. ; Razali, I. & Vijay, K.A. (2007), The Ultimate Drift Velocity in Degenerately-Doped Silicon, *RSM 2007* Proc. 2007, Penang, Malaysia

Versari, R. & Pierraci, A. (1999), Experimental study of hot-carrier effects in LDMOS transistors, *IEEE Transactions on Electron Devices*, Vol. 46, No.6, (Jun 1999), (1228-1233)

Wood, A.; Brakensiek, W.; Dragon, C.; & Bruger, W. (1998). 120 watt, 2 GHz Si LDMOS RF power transistor for PCS base station applications, *IEEE MTT-S Int. Microwave Symp. Dig.*, (1998), (707–710).

Xu, S.; Foo, P.; Wen, J.; Liu, Y.; Lin, F. & Ren, C. (1999) RF LDMOS with extreme low parasitic feedback capacitance and highhot-carrier immunity, *Electron Devices Meeting*, 1999 IEDM Technical Digest. International, (1999), (201-204)

Yang, Y.; Yi, J.; & Kim, B. (2001), Accurate RF Large-Signal Model of LDMOSFETs Including Self-Heating Effect, *IEEE Transactions Microwave Theory and Techniques*., Vol.49, (Feb 2001), (387-390)

Yu, C.; Yuan, J. S. & Suehle, J. (2006), Channel Hot-Electron Degradation on 60-nmHfO_2 – Gated NMOSFET DC and RF Performances, *IEEE Transactions on Electron devices*, Vol. 53, No. 5, (May 2006), (1065-1072)

Comparative Analysis of High Frequency Characteristics of DDR and DAR IMPATT Diodes

Alexander Zemliak
Puebla Autonomous University
Mexico
National Technical University of Ukraine "KPI"
Ukraine

1. Introduction

IMPATT (IMPact Avalanche ionization and Transit Time) diodes are principal active elements for use in millimetric generators. Semiconductor structures suitable for fabrication of continuos-mode IMPATT diodes have been well known for a long time (Scharfetter & Gummel, 1969; Howes & Morgan, 1976). They have been utilized successfully in many applications in microwave engineering. The possibilities of using the same structures for pulsed-mode microwave generators were realized too. Considering, that the increase of the output power of millimetric generators is one of the main problems of microwave electronics; it is important to optimize the diode's active layer to obtain the generator maximum power output. From the beginning (Read, 1958) the main idea to obtain the negative resistance was defined on the basis of the phase difference being produced between RF voltage and RF current due to delay in the avalanche build-up process and the transit time of charge carriers. The single drift region (SDR) and the double drift region (DDR) IMPATT diodes are very well known (Scharfetter & Gummel, 1969; Howes & Morgan, 1976; Fong & Kuno, 1979; Chang, 1990) and used successfully for the microwave power generation in millimeter region. The transit time delay of both types of diodes is the essential factor of the necessary phase conditions to obtain negative resistance. The typical DDR diode structure is shown on Fig. 1(a) by curve 1, where N is the concentration of donors and acceptors, l is the length of diode active layer. In this type of diodes, the electrical field is strongly distorted when the avalanche current density is sufficiently high. This large space charge density is one of the main reasons for the sharp electrical field gradient along the charge drift path. Because of this field gradient, the space charge avalanche ruins itself and consequently the optimum phase relations degrade between microwave potential and current. This factor is especially important when the IMPATT diode is fed at the maximum current density, which is exactly the case at the pulsed-mode operation. The idea to use a complex doping profile semiconductor structure for microwave diode was originally proposed in the first analysis of IMPATT diode by (Read, 1958). This proposed ideal structure has never been realized till now. However, a modern semiconductor technology provides new possibilities for the fabrication of sub micron semiconductor structures with complex doping profiles. This stimulates the search for IMPATT-diode special structure's optimization.

(a)

(b)
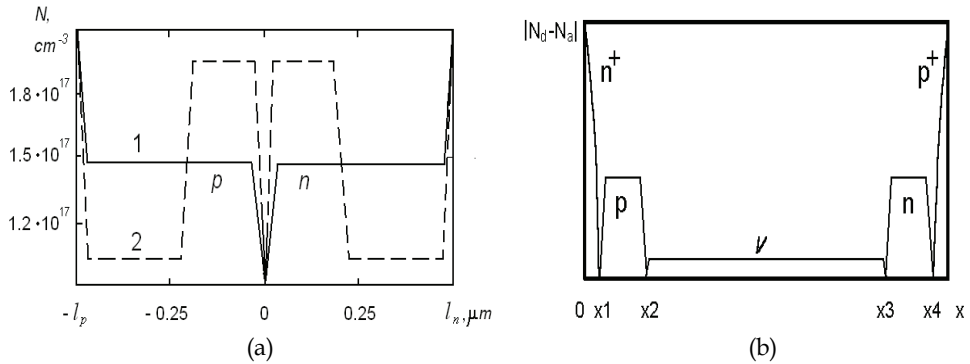
Fig. 1. (a) Doping profile for two types of DDR IMPATT diodes: 1- constant doping profile; 2- quasi-Read-type doping profile; (b) Doping profile for DAR type of IMPATT diode

The other proposed DDR type of IMPATT diode doping profile is shown on the Fig. 1(a) by the curve 2. This type of semiconductor structure can be named as quasi-Read-type structure. This type of doping profile provides a concentration of electrical field within the p-n junction. This measure helps to decrease the destruction of the avalanche space charge and therefore permits to improve the phase stability between the diode current and voltage. The DDR type of IMPATT diode produces one frequency band only in practice because a very strong losses for high frequency bands. The typical small signal admittance characteristic of DDR diode is shown in Fig. 2.
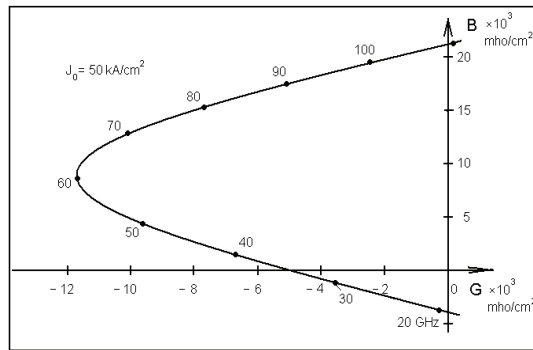
Fig. 2. Complex small signal DDR diode admittance (conductance -G versus susceptance B) for different frequencies

However a diode that has two avalanche regions can produce an avalanche delay which alone can satisfy conditions necessary to generate microwave power. In this case the phase delay of the drift zone becomes subsidiary. The DAR diode can be defined for instance by means of the structure n+p$v$np+ in Fig. 1(b).

The DAR diode has two avalanche regions around n+p and np+ junctions and one common drift region. This type of diode was suggested in (Som et al., 1974). The characteristics of this diode were analyzed in DC and RF modes (Datta & Pal, 1982; Datta et al., 1982; Pati et al., 1991; Panda et al., 1995). The authors affirm that the avalanche delay produced by each of

the thin avalanche regions becomes nearly $\pi/2$, making the total avalanche delay equal to π that is sufficient to produce negative diode resistance. The electric field distribution along the axis x for this type of the diode is shown in Fig. 3, curve 2. Curve 1 approximates the electric field distribution for the DDR with constant doping profile diode for comparison.
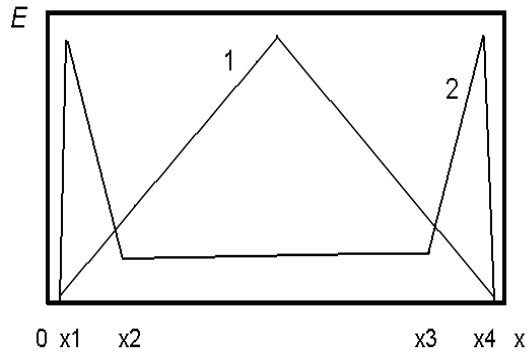
Fig. 3. Electric field distribution for DDR IMPATT diode with constant doping profile (1) and DAR IMPATT diode (2)

There is one avalanche zone for DDR diode and two avalanche zones for DAR diode. The authors of the above cited works affirm that the drift zone transit time delay is not a critical parameter for DAR diode because of the total avalanche delay equal to π. Some advantages of the DAR structure were prognosticated due to this fact. The analysis provided in these works gave the interesting and at the same time very surprising results concerning main features of the DAR diodes. One of the important conclusions of these works concerns of the drift zone width v influence to the diode frequency characteristics. It is noted that the diode active properties are produced practically for any drift zone width and this width has an influence on the number of the frequency bands. The larger drift zone provokes more number of frequency bands. Some of these results were obtained by means of the small signal model (Som et al., 1974; Datta & Pal, 1982; Datta et al., 1982). Other results (Pati et al., 1991; Panda et al., 1995) were obtained on the basis of simplified nonlinear model. Unfortunately since the first invention of DAR IMPATT diode there are no any experimental data concerning the fabrication and/or experimental characteristics investigation of such a type of diode. It is possible that the theoretic prediction and the results obtained by the simulation are not sufficient yet to fabricate the DAR diode structure with desired characteristics. We suppose that we can analyze the DAR diode structure on the basis of precise nonlinear drift-diffusion model of the IMPATT diode which had been developed for DDR type of IMPATT diode (Zemliak et al. 1997; Zemliak & De La Cruz, 2002) and was used successfully for the active layer optimization of this type of the diode.

Historically, many analytical and numerical models have been developed for the various operational modes of IMPATT diodes (Tager & Vald-Perlov, 1968; Scharfetter & Gummel, 1969; Zemliak, 1981; Kafka & Hess, 1981; El-Gabaly et al., 1984; Zemliak & Zinchenko, 1989; Dalle & Rolland, 1989; Zemliak & Roman, 1991; Vasilevskii, 1992; Stoiljkovic et al., 1992; Curow, 1994; Joshi et al., 1995; Tornblad et al., 1996; Zemliak et al., 1997). However, they are some problems with numerical scheme stability for any complex doping profile of IMPATT

diode. We have modified the local-field electrical model (Zemliak & De La Cruz, 2002) to calculate the functional dependence of equation coefficients from electric field, and using all these data finally we derive the IMPATT-diode dynamic characteristics.

2. Numerical models

Two different numerical models are described in this section. The first model is useful for the precise analysis of the internal structure of IMPATT diode and it describes all the most important phenomena into the semiconductor structure. The other model is approximate and it is useful during the process of the internal structure optimization.

2.1 Precise numerical model
The numerical model developed for the analysis of various generator operation modes. This model is based on the system of continuity equations for semiconductor structure:

$$\frac{\partial n(x,t)}{\partial t} = \frac{\partial J_n(x,t)}{\partial x} + \alpha_n \left| J_n(x,t) \right| + \alpha_p \left| J_p(x,t) \right|$$

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial J_p(x,t)}{\partial x} + \alpha_n \left| J_n(x,t) \right| + \alpha_p \left| J_p(x,t) \right|$$

$$J_n(x,t) = n(x,t)\, V_n + D_n \frac{\partial n(x,t)}{\partial x}$$

$$J_p(x,t) = p(x,t)\, V_p - D_p \frac{\partial p(x,t)}{\partial x}$$

(1)

where n, p are the concentrations of electrons and holes; J_n, J_p are the current densities; α_n, α_p are the ionization coefficients; V_n, V_p are the drift velocities; D_n, D_p are the diffusion coefficients. Ionization coefficients, drift velocities and diffusion coefficients are functions of two arguments; the spaces coordinate x and the times coordinate t.

The ionization coefficients, drift velocities and diffusion coefficients are functions of electric field and temperature in all points of semiconductor structure. The dependence of these coefficients on electric field and temperature can be approximated using the approach described in (Grant, 1973):

$$\alpha_n(E,T) = \begin{cases} 2.6 \cdot 10^6\ e^{-\left[\left(1.4 \cdot 10^6 + 1.3 \cdot 10^3 \cdot T \right)/E \right]} & E < 2.4 \cdot 10^5 \\ 6.2 \cdot 10^5\ e^{-\left[\left(1.05 \cdot 10^6 + 1.3 \cdot 10^3 \cdot T \right)/E \right]} & 2.4 \cdot 10^5 < E < 5.3 \cdot 10^5 \\ 5.0 \cdot 10^5\ e^{-\left[\left(0.96 \cdot 10^6 + 1.3 \cdot 10^3 \cdot T \right)/E \right]} & E > 5.3 \cdot 10^5 \end{cases}$$

$$\alpha_p(E,T) = \begin{cases} 2.0 \cdot 10^6\ e^{-\left[\left(1.95 \cdot 10^6 + 1.1 \cdot 10^3 \cdot T \right)/E \right]} & 2.0 \cdot 10^5 < E < 5.3 \cdot 10^5 \\ 5.6 \cdot 10^5\ e^{-\left[\left(1.296 \cdot 10^6 + 1.1 \cdot 10^3 \cdot T \right)/E \right]} & E > 5.3 \cdot 10^5 \end{cases}$$

The temperature T is expressed in °C and electrical field E is expressed in *V/cm.* The drift velocities and diffusion coefficients were calculated by means of approximations given in (Jacoboni et al., 1977; Canali et al., 1975; Nava et al., 1979).

The boundary conditions for the system (1) can be written as follows:

$$n(0,t) = N_D(0); \qquad p(l_0,t) = N_A(l_0);$$
$$J_n(l_0,t) = J_{ns}; \qquad J_p(0,t) = J_{ps}. \tag{2}$$

where J_{ns}, J_{ps} are electron current and hole current for inversely biased $p\text{-}n$ junction; $N_D(0)$, $N_A(l_0)$ are concentrations of donors and acceptors at two space points $x = 0$ and $x = l_0$, where l_0 is the length of the active layer of semiconductor structure.

Electrical field distribution into semiconductor structure can be obtained from Poisson equation. As electron and hole concentrations are functions of the time, therefore, this equation is time dependent too and time is the equation parameter. Poisson equation for this problem has the following form:

$$\frac{\partial E(x,t)}{\partial x} = -\frac{\partial^2 U(x,t)}{\partial x^2} = N_D(x) - N_A(x) + p(x,t) - n(x,t) \tag{3}$$

where $N_D(x)$, $N_A(x)$ are the concentrations of the donors and acceptors accordingly, $U(x,t)$ is the potential, $E(x,t)$ is the electric field. The boundary conditions for this equation are follows:

$$U(0,t) = 0; \;\; U(l_0,t) = U_0 + \sum_{m=1}^{M} U_m \, \sin \, (\omega mt + \varphi_m) \tag{4}$$

where U_0 is the DC voltage on diode contacts, U_m is the amplitude of the harmonic number m, ω is the fundamental frequency, φ_m is the phase of harmonic number m, M is the number of harmonics.

Equations (1)-(4) adequately describe the physical processes in the IMPATT diode in a wide frequency band. However, numerical solution of this system is very difficult because of the sharp dependence of equation coefficients on electric field. The evident numerical schemes have poor stability and require a lot of computing time for the good calculation accuracy obtaining. It is more advantageous to use an implicit numerical scheme, that has a significant property of absolute stability. The computational efficiency and the numerical algorithm accuracy are improved by applying space and time coordinates symmetric approximation.

After the approximation of the functions and its differentials, the system (1) is transformed to the implicit modified Crank-Nicholson numerical scheme. This modification consists of two numerical systems, each of them having the three-diagonal matrix. These systems have the following form:

$$-(a_n - b_n)n_{i-1}^{k+1} + (1 + 2a_n)n_i^{k+1} - (a_n + b_n)n_{i+1}^{k+1} = a_n n_{i-1}^k + (1 - 2a_n)n_i^k + a_n n_{i+1}^k$$
$$+b_n\left(n_{i+1}^k - n_{i-1}^k\right) + \alpha_n\left|\tau \cdot V_n n_i^k + r \cdot D_n\left(n_{i+1}^k - n_{i-1}^k\right)\right| + \alpha_p\left|\tau \cdot V_p p_i^k - r \cdot D_p\left(p_{i+1}^k - p_{i-1}^k\right)\right|$$
$$-(a_p + b_p)p_{i-1}^{k+1} + (1 + 2a_p)p_i^{k+1} - (a_p - b_p)p_{i+1}^{k+1} = a_p p_{i-1}^k + (1 - 2a_p)p_i^k + a_p p_{i+1}^k \tag{5}$$
$$-b_p\left(p_{i+1}^k - p_{i-1}^k\right) + \alpha_p\left|\tau \cdot V_p p_i^k - r \cdot D_p\left(p_{i+1}^k - p_{i-1}^k\right)\right| + \alpha_n\left|\tau \cdot V_n n_i^k + r \cdot D_n\left(n_{i+1}^k - n_{i-1}^k\right)\right|$$

$i = 1, 2, \dots I_1 - 1; \qquad k = 0, 1, 2, \dots \infty$

where $a_{n,p} = \dfrac{\tau\, D_{n,p}}{2h^2}$, $b_{n,p} = \dfrac{\tau\, V_{n,p}}{4h}$, $r = \dfrac{\tau}{2h}$, i is the space coordinate node number, k is the time coordinate node number, h is the space step, τ is the time step, l_1 is the space coordinate node number.

The approximation of the Poisson equation is performed using the ordinary finite difference scheme at every time step k:

$$U_{i-1}^{k} - 2U_{i}^{k} + U_{i+1}^{k} = h^2 \left(N_{D\,i} - N_{A\,i} + p_{i}^{k} - n_{i}^{k} \right) \tag{6}$$

The numerical algorithm for the IMPATT diode characteristics calculation consists of the following stages: 1) The voltage is calculated at the diode contacts for every time step. 2) The initial charge distribution is calculated. 3) The electric potential is calculated at every space point from Poisson equation by the factorization method (Krylov et al., 1977). The electrical field distribution along the diode active layer is calculated. 4) The ionization coefficients and drift parameters are calculated in numerical net nodes for the current time step. 5) The system (5) is solved by matrix factorization method and electron and hole concentration distribution is calculated for the new time step. After this, the calculation cycle is repeated for all time steps from the beginning to the step 3. This process is continued until the convergence is achieved. The current of the external electronic circuit is determined. Then all harmonics of external current are calculated by the Fourier transformation $\left(J_0;\ J_m = |J_m| \exp(j\phi_m) \right)$; the admittance is calculated for any harmonic number m $\left(Y_m = J_m / U_m \right)$ and the power characteristics for all harmonics can be calculated by the following formulas: $\left(P_m = -\dfrac{1}{2}\ \mathrm{Re}(Y_m)\ |U_m|^2;\ \eta_m = \dfrac{P_m}{J_0\, U_0} \right)$

2.2 Approximate numerical model

Other numerical model is more suitable for the previous analysis and for the diode internal structure optimization. This model can reduce the total computer time expences for the structure optimization process.

The numerical method for the solution of the system (1) is based on the classical Fourier series utilization. This approach transforms of the boundary problem for the system of differential partial equations to an ordinary differential equation system. The model describes the physical processes in IMPATT diode by the stationary-operation mode and provides the possibilities to reduce the demands for a computer time that is necessary for the output parameters calculation.

We suppose that all functions of the system (1) can be presented in a form of Fourier series:

$$n(x,t) = \sum_{m=-\infty}^{\infty} n_m(x) \cdot \exp(jm\omega t);$$

$$p(x,t) = \sum_{m=-\infty}^{\infty} p_m(x) \cdot \exp(jm\omega t);$$

$$J_n(x,t) = \sum_{m=-\infty}^{\infty} (I_n)_m(x) \cdot \exp(jm\omega t);$$

$$J_p(x,t) = \sum_{m=-\infty}^{\infty} (I_p)_m (x) \cdot \exp(jm\omega t) \, ;$$

$$\alpha_n(x,t) = \sum_{m=-\infty}^{\infty} (\alpha_n)_m (x) \cdot \exp(jm\omega t) \, ; \qquad (7)$$

$$\alpha_p(x,t) = \sum_{m=-\infty}^{\infty} (\alpha_p)_m (x) \cdot \exp(jm\omega t) \, ;$$

$$V_n(x,t) = \sum_{m=-\infty}^{\infty} (v_n)_m (x) \cdot \exp(jm\omega t) \, ;$$

$$V_p(x,t) = \sum_{m=-\infty}^{\infty} (v_p)_m (x) \cdot \exp(jm\omega t) \, ;$$

$$D_n(x,t) = \sum_{m=-\infty}^{\infty} (d_n)_m (x) \cdot \exp(jm\omega t) \, ;$$

$$D_p(x,t) = \sum_{m=-\infty}^{\infty} (d_p)_m (x) \cdot \exp(jm\omega t) \, .$$

In such a case the principal system (1) can be reduced to a system of the ordinary differential equations for the complex charge density and for the current amplitudes:

$$\frac{dn_m}{dx} = \sum_{k=-\infty}^{\infty} \left\{ -\left(\frac{v_n}{d_n}\right)_k n_{m-k} + \frac{(I_n)_{m-k}}{(d_n)_k} \right\}$$

$$\frac{dp_m}{dx} = \sum_{k=-\infty}^{\infty} \left\{ \left(\frac{v_p}{d_p}\right)_k p_{m-k} - \frac{(I_p)_{m-k}}{(d_p)_k} \right\}$$

$$\frac{d(I_n)_m}{dx} = jm\omega \, n_m - \sum_{k=-\infty}^{\infty} \left\{ (\alpha_n)_k (I_n)_{m-k} + (\alpha_p)_k (I_p)_{m-k} \right\} \qquad (8)$$

$$\frac{d(I_p)_m}{dx} = -jm\omega \, p_m + \sum_{k=-\infty}^{\infty} \left\{ (\alpha_n)_k (I_n)_{m-k} + (\alpha_p)_k (I_p)_{m-k} \right\}$$

$$m = 1,2,\dots\infty$$

where $(\alpha_n)_m, (\alpha_p)_m$ are the electron and hole ionization coefficient amplitudes, $(v_n)_m, (v_p)_m$ are the electron and hole velocity amplitudes, $(d_n)_m, (d_p)_m$ are the electron and hole diffusion coefficient amplitudes, n_m, p_m are the electron and hole concentration amplitudes, $(I_n)_m, (I_p)_m$ are the electron and hole current amplitudes.

A number of harmonics m in these series can be reduced down to the number M, which defines the accuracy of the solution and necessary computer time. The system (8) can be presented in matrix form as:

$$Y' = AY \qquad (9)$$

The charge diffusion and sharp dependence of the ionization coefficients on the electrical field determine the great module of eigenvalues of the matrix A. For this case, a shooting method, which reduces a boundary problem to Cauchy problem, is not suitable because coordinate basis degenerates in the solution process and therefore is not stable. The boundary problem (9) is solved on the basis of the functional matrix correlation (Bakhvalov et al., 1987):

$$B^t(x)\ Y(x) = G(x) \tag{10}$$

where B^t is the factorization matrix; G is the boundary condition vector. The unknown matrixes of equation (10) are satisfied in the following differential equation system:

$$\begin{aligned} B' + A^t B = 0 \\ G' = 0 \end{aligned} \tag{11}$$

The fundamental matrix F is used to obtain the process stability of the integration of equations (11). This matrix is determined as $F(x) = \exp\{A^t(x_k)h_k\}$, where h_k is the space step. Transition to the next coordinate node is made using the term $B(x_k + h_k) = F(x_k)\ B(x_k)$. The degradation of the coordinate basis B can be overcome using the Gram-Schmidt ortogonalization procedure for equation (10) on each integration step.

The algorithm for the analysis of IMPATT diode includes the following steps: 1) the initial charge distribution in the diode is calculated; 2) the electric field harmonics are determined from the Poisson equation; 3) ionization and drift parameters are determined from the Fourier analysis, and the matrix of the system of equations on the coordinate net is formed; 4) the boundary problem is solved for the system of continuity equations and the charge and current amplitudes are determined. The harmonics of the external circuit current are calculated. After this, the calculation cycle is repeated from the beginning to point 4) until the external current is determined with sufficient convergence. Then all output parameters of the IMPATT diode are determined.

The main advantage of this harmonic method is the reducing the total computer time for the calculation of stationary mode of the IMPATT diode. In Fig. 4 are shown computer time Tc in relative units and relative error Er as the functions of the harmonic number M.
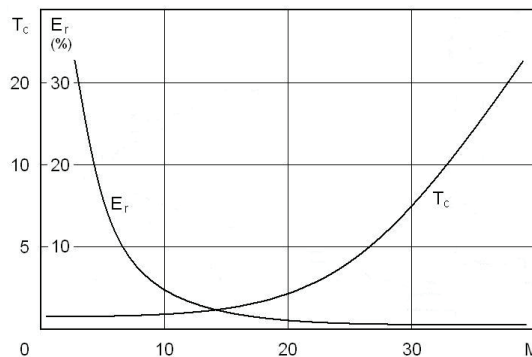
Fig. 4. Computer time Tc in relative units and relative error Er as the functions of the harmonic number M

These data are corresponded to the nonlinear modes with an average level of the non-linearity. For this case we determine error as the relative difference of the diode admittance value that we obtain by this harmonic method and by more precise numerical method of the section 2.1. It is clear that the harmonic number M more than 12-15 is sufficient to obtain a good accuracy of the diode parameters. At the same time we have a significantly reducing of the total computer time. Computer time for one probe of diode analysis is the principal characteristic of the optimization procedure. That is the main reason why this approximate model is elaborated. For example the total computer time for the diode analysis by precise numerical model is corresponded to the number of harmonic $M = 40$.

3. Optimization procedure

The optimization algorithm was designed as the combination of one of kind of direct method and a gradient method. This is one of the modification of well-known algorithm, which is successfully used for function with complicate structure. This method is more precisely successful for the optimization of millimetric wave devices because the objective function of that type of devices as the function of its arguments has a very complex behavior similar to a one "valley" in N-dimensional space. The objective function can be determined as the maximum electronic power, for example. The number of free variables depends on the diode structure. For the DDR diode structure optimization with constant doping profile we need to define 4 parameters: two lengths and two doping profile levels. For the DDR quasi-Read diode structure optimization we need to define 8 parameters: four lengths and four doping profile levels. We have been formed the principal vector of variables y for some parameters of semiconductor structure.

The optimization algorithm consists of the next steps:

1. Given as input two different approximations of two initial points: y^0 and y^1.
2. At these points, we start by gradient method, and have performed some steps. As a result, we have two new points Y^0 and Y^1.

$$y^{0n+1} = y^{0n} - \delta_n \cdot \nabla F\left(y^{0n}\right), \qquad y^{1n+1} = y^{1n} - \delta_n \cdot \nabla F\left(y^{1n}\right), \qquad n = 0,1,\ldots N-1,$$

$$Y^0 = y^{0N}, \qquad Y^1 = y^{1N},$$

    where F is the objective function, δ_n is the parameter of the gradient method.
3. We draw a line through two these points, and performed a large step along this line. We have a new point y^{s+1}:

$$y^{s+1} = Y^s + \alpha \left(Y^s - Y^{s-1}\right), \;\; s=1,$$

    where α is the parameter of the line step.
4. Than we perform some steps from this point by the gradient method, and obtain a new point Y^s.

$$y^{s\,n+1} = y^{s\,n} - \delta_n \cdot \nabla F\left(y^{s\,n}\right), \;\; s=s+1, \;\; Y^s = y^{s\,N}.$$

Then steps 3 and 4 are repeated with the next values of the index s $(s = 2, 3, \ldots)$.

The optimization process that is presented above cannot find the global minimum of the objective function, but only a local one. To obtain the confidence that we have the better solution of the optimum procedure, it is necessary to investigate N-dimensional space with different initial points. In that case, it is possible to investigate N-dimensional volume in more detail. During the optimization process, it is very important to localize the subspace of the N-dimension optimization space for more detail analysis. N-dimensional space volume of independent parameters is determined approximately on the base of model described in section 2.2 for the first stage of optimization procedure. In that case, a Fourier series approximation of principal functions is used and because of this approximate model, we have a ten times acceleration. After that, on the basis of the precise model described in section 2.1 we have analyzed the internal structure of the different types of silicon diode.

4. DDR diode analysis

4.1 94 GHz diode
In Fig.5 (a), (b) the characteristics of power-level and efficiency for the constant profile DDR diode and complex profile DDR diode for 94 GHz are presented as functions of feeding current density I_0 for the optimum structures and for others that are near the optimum. Parameters of these structures are presented in Table 1 and Table 2.

n	L 1 $(10^{-4}\ cm)$	L 2 $(10^{-4}\ cm)$	N 1 $(10^{17}\ cm^{-3})$	P 1 $(10^{17}\ cm^{-3})$
1	0.35	0.33	1.5	1.5
2	0.35	0.33	1.65	1.68
3	0.35	0.33	1.7	1.7
4	0.35	0.33	1.9	1.9
5	0.35	0.33	2.1	2.1

Table 1. Internal structure parameters for diode with constant doping profile for 94 GHz

n	L 1 $(10^{-4}\ cm)$	L 2 $(10^{-4}\ cm)$	L 3 $(10^{-4}\ cm)$	L 4 $(10^{-4}\ cm)$	N 1 $(10^{17}\ cm^{-3})$	N 2 $(10^{17}\ cm^{-3})$	P 2 $(10^{17}\ cm^{-3})$	P 1 $(10^{17}\ cm^{-3})$
1	0.086	0.283	0.266	0.084	1.3	2	2	1.3
2	0.065	0.212	0.203	0.063	1.3	2	2	1.3
3	0.072	0.236	0.222	0.072	1.3	2	2	1.3
4	0.072	0.236	0.222	0.072	1.56	2.4	2.4	1.56

Table 2. Internal structure parameters for diode with complex doping profile for 94 GHz

The parameters for optimization procedure for DDR diode with constant doping profile are defined as: *L1* is the length of *n* region, *L2* is the length of *p* region, *N1* and *P1* are the doping profile level for *n* and *p* regions accordingly. Structure 4 in Fig. 5(a) has the maximum power level 436 kW/cm^2 and the optimal current density value I_0 =140 kA/cm^2. In that case, the efficiency is equal to 11.2 % for the maximum power point. Structure 5 has a maximum efficiency as the function of the current I_0, but for the optimum power point this value no larger than for structures 2, 3 and 4. Besides, for this structure it is necessary to increase the current value until 153 kA/cm^2 to obtain of the optimum power point. Semiconductor structures with a complex doping profile are analyzed to improve the power level and efficiency of pulsed-mode IMPATT diode (Fig. 5(b)). In that case eight parameters have been varied: *L1, L2, L3, L4, N1, N2, P2, P1. L1* is the length of *n* region with low level of doping, *L2*

is the length of n region with high level of doping, $L3$ is the length of p region with high level of doping, $L4$ is the length of p region with low level of doping, $N1$ and $N2$ are the low and high levels of doping for n region, and $P1$ and $P2$ are the low and high levels of doping for p region. Structure 3 is the optimal one and has an efficiency of about 15% and 446 kW/cm^2 power level at 123 kA/cm^2. Others structures are near this optimum one but have a lower power level and efficiency. The extension of doping level high parts (structure 1) or increasing this level (structure 4) results to moving the power curve to the greater current density.
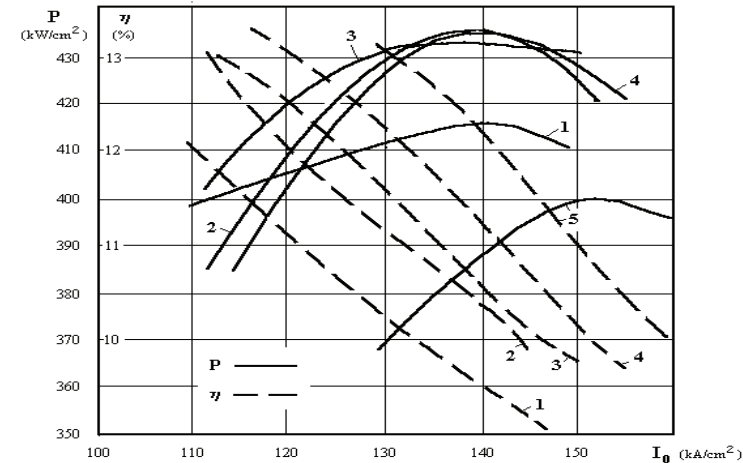
(a)
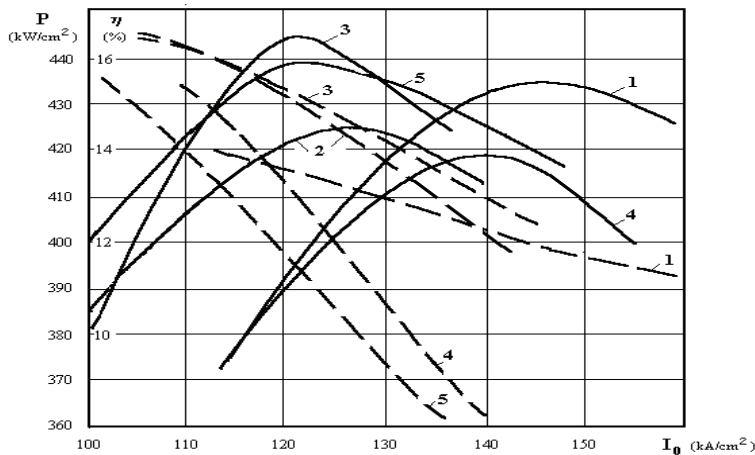
(b)

Fig. 5. Output power P and efficiency coefficient η as functions of the feeding current density I_0 for optimum and near optimum structures for (a) constant and (b) complex doping profile DDR diode

Comparison of the optimal characteristics for two different types of the structures as a constant doping profile (curve 4, Fig. 5a) and complex doping profile (curve 3, Fig. 5b) shows that the maximum output power level is almost equal for two these optimal structures (436 kW/cm^2 and 446 kW/cm^2), but efficiency coefficient has more difference (11.2% and 14.4%). The most important fact is a significant decrease of optimal value of permanent current density for the complex doping structure. The optimal current density value for this case is equal to 140 kA/cm^2 for the constant doping structure, but for the complex doping structure is equal to 123 kA/cm^2. Therefore the complex doping profile structure has better energy characteristics.
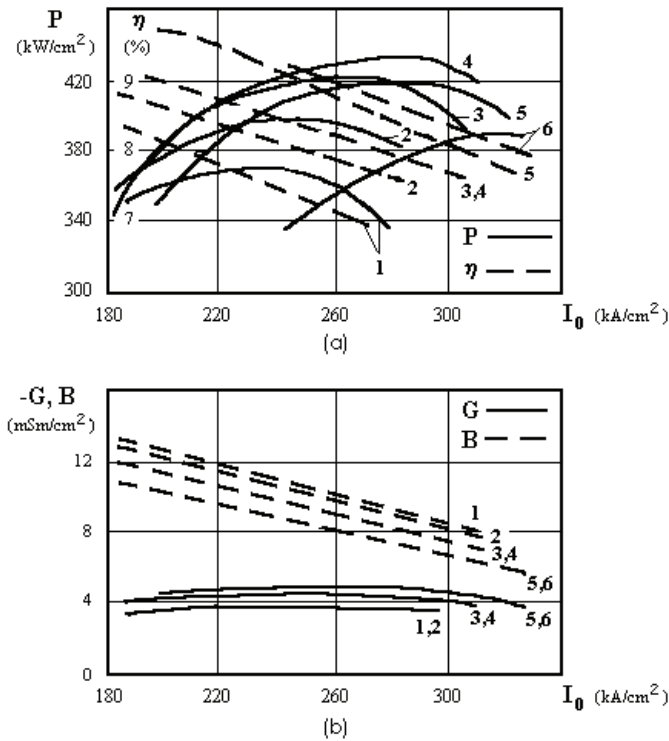
Fig. 6. (a) Output power P and efficiency coefficient η, (b) real G and imaginary B parts of the total admittance as functions of feeding current density I_0 for optimal and near optimal structures with constant doping profile level

4.2 140 GHz diode

In Fig.6 (a), (b) the characteristics of power-level, efficiency, and the real and imaginary parts of the complex admittance of the 140 GHz diode are presented as functions of the feeding current density I_0 for the optimum structures and for others that are near the optimum. Parameters of the constant doping level and length values for this figure are presented in Table 3.

n	L 1 (10^{-4} cm)	L 2 (10^{-4} cm)	N 1 (10^{17} cm^{-3})	P 1 (10^{17} cm^{-3})
1	0.22	0.19	3	3
2	0.24	0.19	3.5	3.5
3	0.21	0.19	3.5	4.5
4	0.21	0.19	4	4
5	0.21	0.19	4.5	4.5
6	0.21	0.19	5	5

Table 3. Internal structure parameters for diode with constant doping profile for 140 GHz

All technological parameters are defined similar to previous section. Structure 4 has a maximum power level of 430 kW/cm^2 and an optimal current density value of I_0 = 285 kA/cm^2. In that case, the efficiency is equal to 8.0 % for the maximum power point. Structure 5 has a maximum of the negative real admittance and efficiency (8.5 %), but has a smaller power level because the doping level is high, and therefore the permanent voltage and first harmonic amplitude voltage are smaller. Structure 6 has a maximum efficiency as the function of the current I_0, but for the optimum power point, this value is less than for structures 4 and 5. Besides, for this structure, it is necessary to increase the current value to 320 kA/cm^2 to obtain the optimum power point. Structures 5 and 6 have a maximum value of the real part of the total admittance, but have a greater doping level, and therefore a smaller value of the permanent and variable voltage and output power.

Semiconductor structures with a complex doping profile are analyzed for improving the power level and efficiency of the IMPATT diode with a maximum level of permanent current density. Parameters of these structures are presented in Table 4. In Fig. 7 (a), (b) the dependencies of power level, efficiency and admittance are presented as functions of feeding current density I_0 for the optimum structure and for near optimum ones.

n	L 1 (10^{-4} cm)	L 2 (10^{-4} cm)	L 3 (10^{-4} cm)	L 4 (10^{-4} cm)	N 1 (10^{17} cm^{-3})	N 2 (10^{17} cm^{-3})	P 2 (10^{17} cm^{-3})	P 1 (10^{17} cm^{-3})
1	0.09	0.08	0.11	0.06	1.6	4.7	4.1	1.6
2	0.09	0.08	0.11	0.06	1.5	4.7	4.7	1.5
3	0.09	0.08	0.11	0.06	1.6	4.7	4.1	1.4
4	0.09	0.08	0.11	0.06	1.6	5.2	4.6	1.6
5	0.08	0.09	0.12	0.05	1.6	4.7	4.1	1.6
6	0.07	0.08	0.11	0.04	1.6	4.7	4.1	1.6

Table 4. Internal structure parameters for diode with complex doping profile for 140 GHz

Structure 1 is the optimal one. In this case, the power level is 457 kW/cm^2, and the optimal current density value is 235 kA/cm^2. Others structures are near this optimum one, but have a lower power level and efficiency. The extension of the high doping level parts (structures 4 and 5) results in moving the power curve to a greater current density. These two types of semiconductor structures have a greater value of active admittance than the optimal one, but have a smaller microwave voltage amplitude and power level.

It is very important to compare the optimal characteristics for the two different types of structures as the constant doping profile (curve 4, Fig. 6(a)) and the complex doping profile (curve 1, Fig. 7(a)). A comparative analysis shows that the maximum output power level is quasiequal for two these optimal structures (436 kW/cm^2 and 452 kW/cm^2), but the efficiency coefficient has more difference (8.5% and 10.7%). The most important fact is a significant decrease of the optimal value of the permanent current density for the complex

doping structure. For the permanent doping structure, the optimal current density value is $285\ kA/cm^2$, but for the complex doping structure, it is $235\ kA/cm^2$.

(a)

(b)
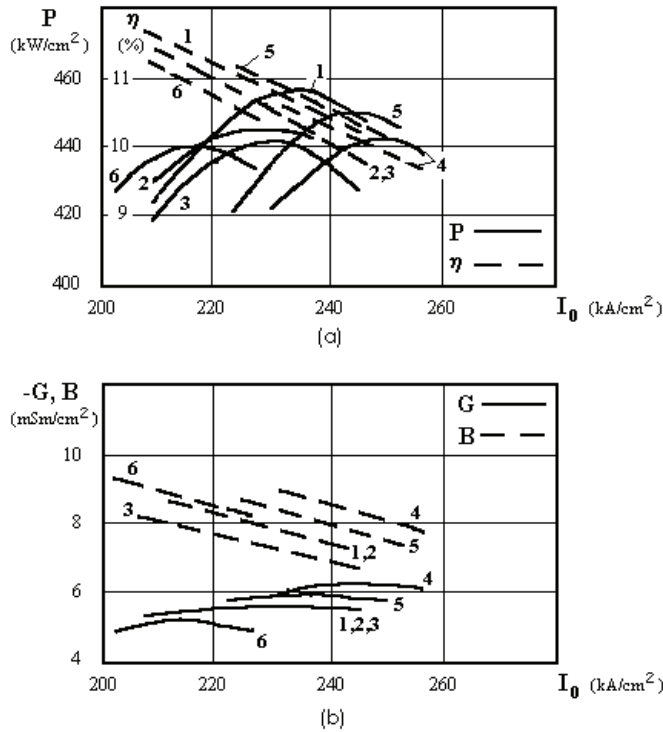
Fig. 7. (a) Output power P and efficiency coefficient η and (b) real G and imaginary B parts of the total admittance as functions of feeding current density I_0 for optimal and near optimal structures with complex doping profile level

Therefore, the complex doping profile structure has better energy characteristics, and allows the possibility to exploiting the diode under easier conditions.

One of the important problems for the real type of the complex doping profile diode optimization is the sensitivity analysis of energy characteristics for various geometrical sizes and doping levels. The total number of the analyzed structures is very large, because of a large number of combinations of the 8 parameters. Some results of the investigation of an optimal structure by changing the doping profile levels $N1$, $N2$, $P2$, $P1$ and lengths $L1$, $L2$, $L3$, $L4$ within 20% around the optimal structure are presented in Tables 5 and 6, respectively. Presented examples give the possibility to evaluate correctly the technology inaccuracy influence to the energy characteristics deterioration. In these tables, we present the maximum achievable output power density, the permanent current density that corresponds to this optimum, and the real and imaginary parts of the complex admittance. The diode doping profile level increase within 20% with respect to the optimal structure leads to a small decrease of the output power. On the other hand, the diode doping profile level decrease leads to a great decrease of output power. Some structures (6, 7 of Table 5)

have a real admittance module greater than the optimal one (number 1), but because the microwave voltage amplitude is decreased, the output power level for these structures is smaller.

n	$N1$	$N2$	$P2$	$P1$	P_{max}	I_{opt}	G	B
	$(10^{17}\ cm^{-3})$	$(10^{17}\ cm^{-3})$	$(10^{17}\ cm^{-3})$	$(10^{17}\ cm^{-3})$	(kW/cm^2)	(kA/cm^2)	(mSm/cm^2)	(mSm/cm^2)
1	1.6	4.7	4.1	1.6	455	235	-5.24	7.83
2	1.8	4.7	4.1	1.8	451	241	-5.36	7.51
3	2	4.7	4.1	2	451	242	-5.46	7.21
4	1.5	4.7	4.7	1.5	448	230	-5.25	7.51
5	1.4	4.1	4.1	1.4	443	227	-5.21	7.35
6	1.6	5	4.4	1.6	448	241	-5.71	8.07
7	1.6	5.2	4.6	1.6	447	251	-6.04	8.38
8	1.6	4.5	3.9	1.6	443	225	-4.78	8.51
9	1.6	4.2	3.7	1.6	440	227	-4.43	8.67

Table 5. Doping profile level variation within 20% around the optimal structure and diode output characteristics corresponding to these structures

n	$L1$	$L2$	$L3$	$L4$	P_{max}	I_{opt}	G	B
	$(10^{-4}\ cm)$	$(10^{-4}\ cm)$	$(10^{-4}\ cm)$	$(10^{-4}\ cm)$	(kW/cm^2)	(kA/cm^2)	(mSm/cm^2)	(mSm/cm^2)
1	0.09	0.08	0.11	0.06	455	235	-5.24	7.83
2	0.08	0.09	0.12	0.05	451	245	-5.51	7.91
3	0.07	0.1	0.13	0.04	447	253	-5.61	8.11
4	0.1	0.07	0.1	0.07	450	230	-4.97	8.11
5	0.11	0.06	0.09	0.08	445	225	-4.85	8.44
6	0.1	0.08	0.11	0.07	450	232	-5.17	7.41
7	0.11	0.08	0.11	0.08	447	234	-4.91	6.81
8	0.08	0.08	0.11	0.05	441	244	-5.34	7.41
9	0.07	0.08	0.11	0.04	434	246	-5.36	7.56

Table 6. Length variation within 20% around the optimal structure and diode output characteristics corresponding to these structures

An analysis of the results that are presented in Table 6 shows that the variation of the total length $L=x9-x1$ around the optimal value leads to a great deterioration of the energy characteristics (structures 8 and 9). On the other hand, the redistribution of separate part's dimensions between the high and low doping profile parts within 20% has not led to a great decrease of the output power level.

5. DAR diode analysis

5.1 Numerical scheme convergence
The analysis of numerical scheme (5)-(6) for different DDR IMPATT diode showed a good convergence of the numerical model. The convergence was obtained during 6 – 8 periods.
The analysis of numerical model for the DAR type of the doping profile (Fig. 2) gave unexpected but understandable results. The numerical scheme convergence for this type of the doping profile is very slow. The quantitative results of the numerical scheme convergence for the principal diode characteristic, DAR diode conductance as the period number function are shown in Fig. 8. The necessary number of the consequent periods depends on the operating frequency and can change from 30 – 50 for the frequency region 15 – 60 GHz up to 150 – 250 periods for 200 – 300 GHz. This very slow convergence is stipulated by the asynchronies movement of the electron and hole avalanches along the

same transit time region v. It occurs owing to the different drift velocities of the carriers. This type of the numerical convergence provokes a large number of necessary periods and a large computer time. We need to declare that the bad convergence result is the feature of the diode mathematical numerical scheme and is not the physical diode property. The physical stability is discussed in detail below.
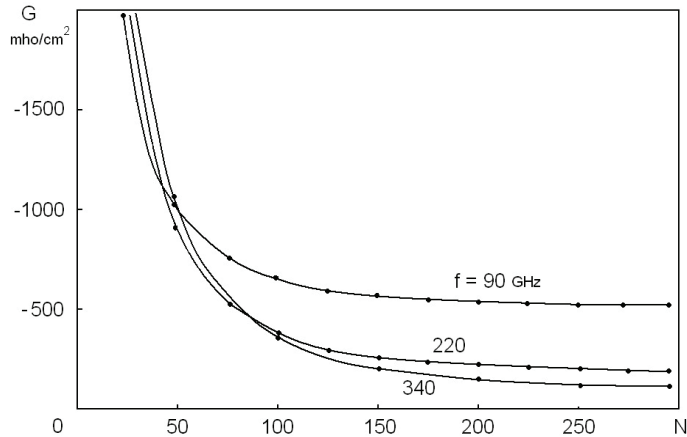
Fig. 8. Calculated conductance as function of period number N

5.2 Results

The analysis of DAR IMPATT diode provided some yeas ago shown interesting and at the same time very surprising results concerning main features of the DAR diodes. One of the important conclusions of these works concerns of the drift zone width v influence to the diode frequency characteristics. It is noted that the diode active properties are produced practically for any drift zone width and this width has an influence on the number of the frequency bands. The larger drift zone provokes more number of frequency bands. Some of these results were obtained by means of the small signal model (Som et al., 1974; Datta & Pal, 1982; Datta et al., 1982). Other results (Pati et al., 1991; Panda et al., 1995) were obtained on the basis of simplified nonlinear model. We suppose that it is necessary to analyze this diode by means of precise model described in section 2.1 (Zemliak & De La Cruz, 2006).

The DAR diode doping profile was defined the same as in paper (Panda et al., 1995) for primary analysis to provide the adequate comparison between results which were obtained by two different approaches. Then the accurate analysis for DAR IMPATT diode has been made for different values of p, n and v region width and the different donor and acceptor concentration level. The analysis showed that the active properties of the diode practically are not displayed for more or less significant width of the region v. The same doping profile as in (Panda et al., 1995) gives the negative conductance for very narrow frequency band only as shown in Fig. 9 in conductance versus susceptance plot.

The solid line of this figure gives dependency for drift layer width $W_v = 0.6$ μm and the dash line for $W_v = 1.5$ μm. First dependency displays the diode active properties for one narrow frequency band from 50 GHz up to 85 GHz. Second admittance dependency for $W_v = 1.5$ μm gives very narrow one frequency band from 40 GHz up to 62 GHz with a vary small value of negative conductance G. In general the admittance behavior has a damp oscillation

character but only first peak lies in negative semi plane. The negative conductance disappears completely for W_v >1.5 μm. All these results have been obtained in assumption of a sufficiently small value of a series resistance R_s = 0.5$_{10}$-6 Ohm · cm². This value was used for all further analysis too.
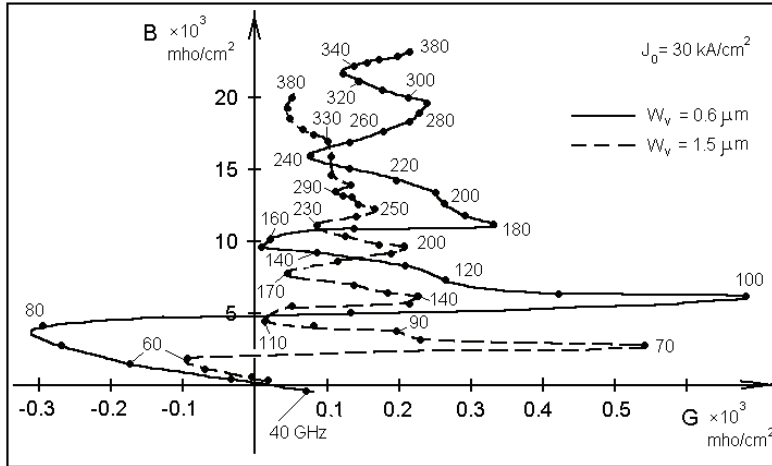
Fig. 9. Complex small signal DAR diode admittance (conductance -G versus susceptance B) for different frequencies and two values of drift layer widths W_v

The main reason of obtained characteristics behavior is the same as for the slow mechanism convergence of the numerical model. The electron and hole avalanches have different transit velocities but they move along the same drift region v. It provokes different time delay for the carriers during the transit region movement. The larger width of the region v makes delay time more different and the active properties are reduced. That is why we need to reduce the width W_v to obtain necessary negative admittance. This conclusion is contrary to results of the papers (Pati et al., 1991; Panda et al., 1995). The main results obtained by these authors showed the DAR diode active features presence in some frequency bands for different values of v region widths from 0.5μm to 2.0μm. Our results display the active features of the DAR diode the same profile for some frequency bands in case when the v-region width less than 0.5μm only. The obtained difference could be explained probably by means of approximate numerical model used in paper (Panda et al., 1995). One modified Runge-Kutta method was used to mathematical model solve as shown in this paper. However it is known that any explicit numerical method like Runge-Kutta does not have the necessary stability to solve the sufficiently difficult problem (1) with a very sharp dependency of ionization coefficients.

One positive idea to increase negative admittance of the diode consists in non-symmetric doping profile utilization too. This profile gives some compensation to the asynchronies mechanism. Taking into account these considerations non-symmetric doping profile diode was analyzed in a wide frequency band. One of the perspective diode structures that was analyzed detail is defined by means of following parameters: the doping level of the n-zone is equal to 0.5$_{10}$17 cm^{-3}, the doping level of the p-zone is equal to 0.2$_{10}$17 cm^{-3}, the widths of the two corresponding areas are equal to 0.1μm and 0.2μm, accordingly, the width of the

drift v-region is equal to 0.32μm, the width of each p-n junction was given as 0.02μm from the technological aspects. This structure provides concentration of electrical field within the two p-n junctions and asynchronies mechanism is not displayed drastically yet.

In Fig. 10 the small signal complex admittance i.e. the conductance versus susceptance is presented for the wide frequencies band for DAR diode and for the current density J_0 = 30 kA/cm². The DC voltage U_0 is equal to 26.59 V with a small variation from one frequency to other to obtain this value of current density. The first harmonic voltage amplitude is equal to 0.1 V.

There are some differences of the DAR diode frequency characteristics from the classical DDR IMPATT diodes. First of all the DAR diode type has three active bands in the millimeter range (Fig.10) and the DDR diode has only one band. The first active band of the DAR diode is very wide and covers frequency region from 12 to 138 GHz. The second and the third bands give the perspective to use this structure for the high frequency generation in the millimeter range too. We can decide that two superior bands appear from the positive conductance G semi plane (look Fig. 9) as a result of the special conditions making for these bands. This effect gives possibility to use superior frequency bands, at least the second band, for the microwave power generation of the sufficient level. The dependences of conductance -G as the function of the first harmonic amplitude U_1 are shown in Fig. 11 for three frequency bands and for the same value of the current density J_0 = 30 kA/cm².
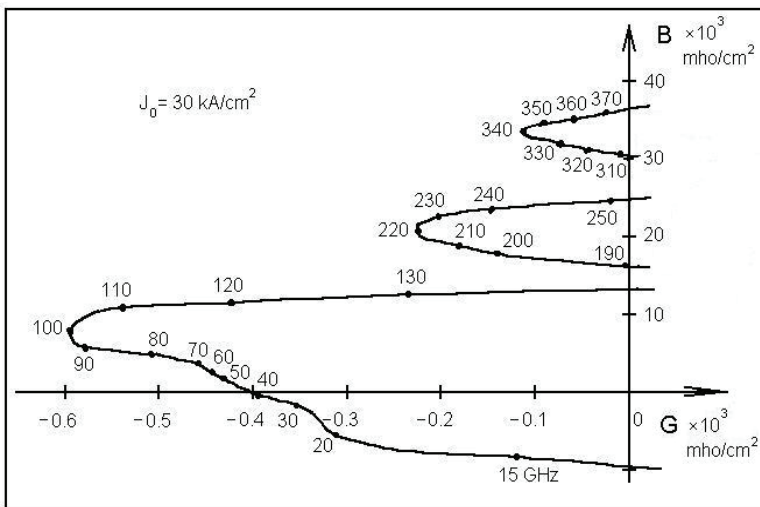
Fig. 10. Complex small signal DAR diode admittance (conductance -G versus susceptance B) for different frequencies and W_v = 0.32μm

It is clear that the first frequency band characteristic (f = 90 GHz) has a better behavior. The maximum value of the conductance -G is large and achieves nearly the 600 mho/cm² under the small signal. The amplitude dependency for the first band is very soft and this provides a significant value of the generated power. Nevertheless the second and the third bands (for 220 GHz and for 340 GHz) have the perspective too. The output power dependences for two frequency bands are presented in Fig. 12 as functions of the first harmonic amplitude.
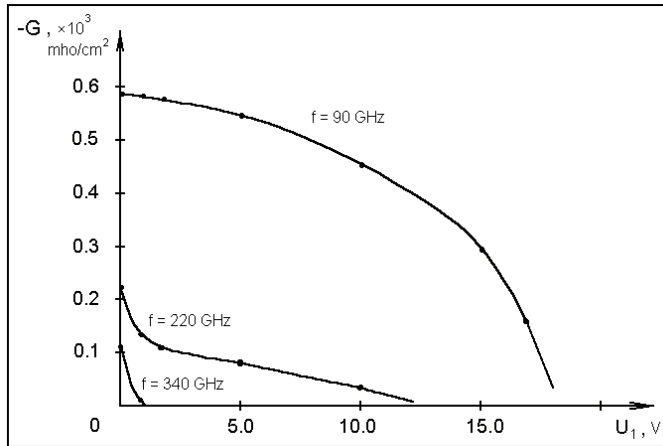
Fig. 11. Conductance G dependence as functions of first harmonic amplitude U_1 for different frequency bands
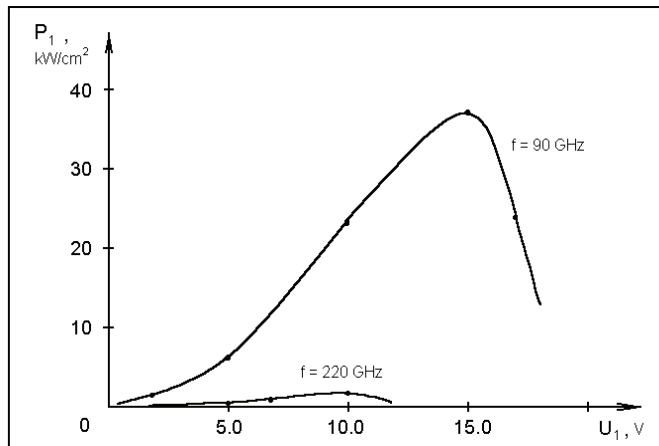
Fig. 12. Output power P dependence as functions of first harmonic amplitude U_1 for different frequency bands

The maximum power density is equal to 37 kW/cm^2 for the first frequency band (90 GHz), and 1.4 kW/cm^2 for the second one (220 GHz). One principal limit of output power for second and third bands is based on the sharp amplitude dependency as shown in Fig. 11. However possible optimization of the diode internal structure for selected frequency band can improve these characteristics and permits raise the power and the efficiency.

The DAR diode internal structure optimization has been provided below for the second frequency band near 220 GHz and for the third frequency band near 330 GHz separately. The optimization algorithm is combined by one kind of direct method and a gradient method and was described in section 3.

The cost function of the second frequency band optimization process was selected as output power for frequency 220 GHz. It means that the energy characteristics for the first and the

third frequency bands have been obtained as functions of a secondary interest without a special improvement. The set of the variables for the optimization procedure was composed from five technological parameters of the diode structure: two doping levels for p and n regions and three widths of p, n and v regions. The optimal values of these parameters are following: doping level of the n-zone is equal to $0.42_{10}17$ cm^{-3}, the doping level of the p-zone is equal to $0.28_{10}17$ cm^{-3}, the widths of the two corresponding areas are equal to 0.1μm and 0.2μm, accordingly, and the width of the drift v-region is equal to 0.34μm, The internal structure optimization of second frequency band has been made for the feeding current density 30 kA/cm^2. However it is interesting to calculate the diode power characteristics for other current density too. The complete analysis was done for three current density values: 30 kA/cm^2, 50 kA/cm^2 and 70 kA/cm^2. Although the structure optimization was provided for the large signal, the small signal diode admittance dependency is an interest too. These small signal characteristics are shown in Fig. 13 for all possible frequency bands and three values of feeding current density.
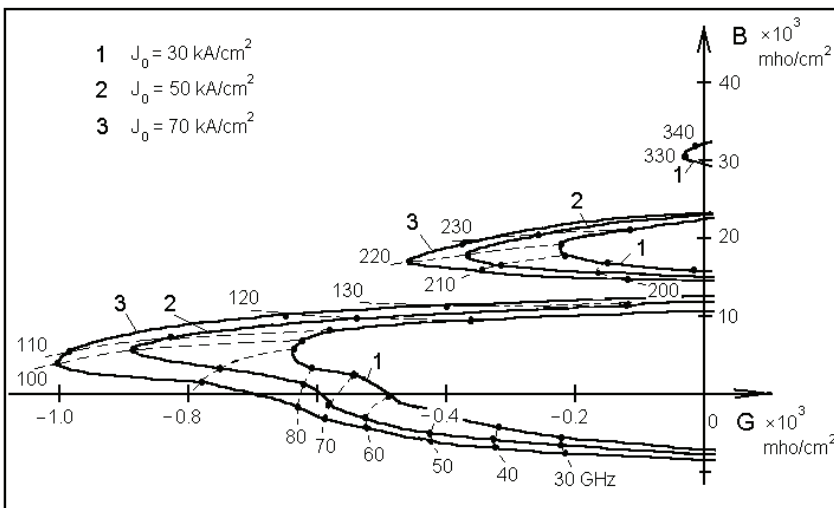
Fig. 13. Complex small signal DAR diode admittance optimized for second frequency band for different value of feeding current density

The active diode properties for two first bands are improved when the current density increases. Because the technological parameters have been optimized for 220 GHz more positive effect was obtained for this frequency. At the same time the third frequency band practically disappears. The maximum value of diode conductance for more favorable frequency 330 GHz is equal to -30 mho/cm^2 for the structure optimized for the second band. On the other hand the diode conductance is equal to -120 mho/cm^2 for 340 GHz for before analyzed structure in Fig. 10. Further current density increasing leads to complete disappearance of the active properties for this frequency.

The characteristics obtained for 220 GHz under the large signal serve as the main result of the optimization process. The amplitude characteristics for this frequency and for three values of feeding current density are shown in Fig. 14.
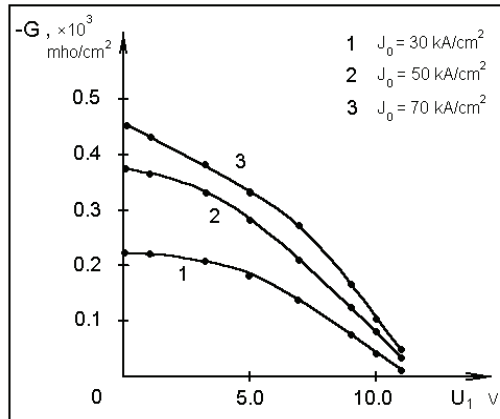
Fig. 14. Conductance G dependence as functions of first harmonic amplitude U_1 for f = 220 GHz and for three values of feeding current density

Because the diode structure optimization was provided for current density 30 kA/cm^2 the amplitude characteristic that corresponds to this current has better behavior in comparison to others. This characteristic is softer. Characteristics for current densities 50 kA/cm^2 and 70 kA/cm^2 are sharper but correspond to larger conductance $-G$. As a result this property gives a larger output generated power. The characteristics that correspond to two last current density values can be made better if the optimization process realize for each of this current value. The output power dependencies as a function of first harmonic amplitude U_1 for f = 220 GHz and for three values of feeding current density are shown in Fig. 15.
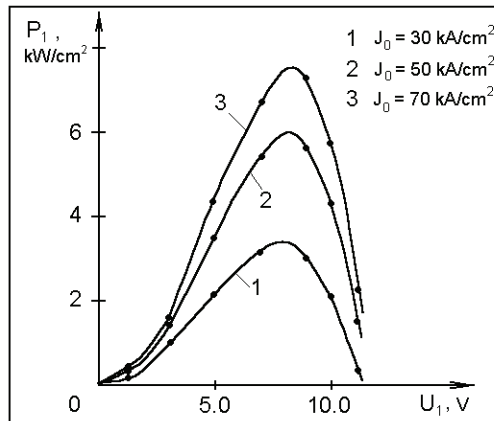
Fig. 15. Output generated power P dependence as functions of first harmonic amplitude U_1 for f = 220 GHz and for three values of feeding current density

We can state that a sufficient improvement of power characteristics is observed for this diode structure in comparison with before analyzed structure. The maximum values of generated power are equal to 3.3 kW/cm^2 for J_0 = 30 kA/cm^2, 6.0 kW/cm^2 for J_0 = 50

kA/cm² and 7.5 kW/cm² for J_0 = 70 kA/cm² accordingly. These results were obtained taking into account the series resistance R_s = 0.5$_{10}$-6 Ohm · cm². The increase of this resistance up to 1.0$_{10}$-6 Ohm · cm² gives reduction of the conductance and the generated power from 5 times for J_0 =30 kA/cm² to 2 times for J_0 = 70 kA/cm². However it is possible provide the diode structure optimization for large value of the current density. In this case we can obtain a significant level of output generated power again.

The power level optimization for the third frequency band has been provided below. The cost function of the third frequency band optimization process was selected as output power for the frequency 330 GHz and for two values of the feeding current density as 50 kA/cm² and 70 kA/cm². The results of the analysis and optimization of small signal admittance for third frequency band are shown in Fig. 16 for two current density values: 50 kA/cm² and 70 kA/cm².

The set of the variables for the optimization procedure was composed from five technological parameters of the diode structure: two doping levels for *p* and *n* regions and three widths of *p*, *n* and *v* regions.
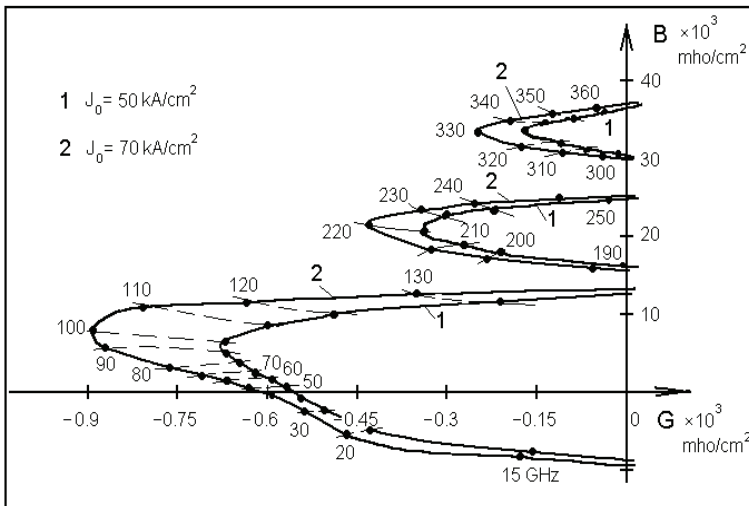
Fig. 16. Complex small signal DAR diode admittance optimized for third frequency band for two values of feeding current

The optimal values of these parameters are next: doping levels of *n* and *p* zone are equal to 0.48 $_{10}$17 cm^{-3} and 0.36$_{10}$17 cm^{-3} accordingly, the widths of the two corresponding areas are equal to 0.09μm and 0.18μm, and the width of the drift *v*-region is equal to 0.32μm.

The active diode properties for all frequency bands are improved when the current density increases. More positive effect was obtained for the frequency 330 GHz because the optimization for this frequency.

The characteristics obtained for 330 GHz under a large signal serve as the main result. The amplitude characteristics of the conductance for this frequency are shown in Fig. 17 for two values of the current density.
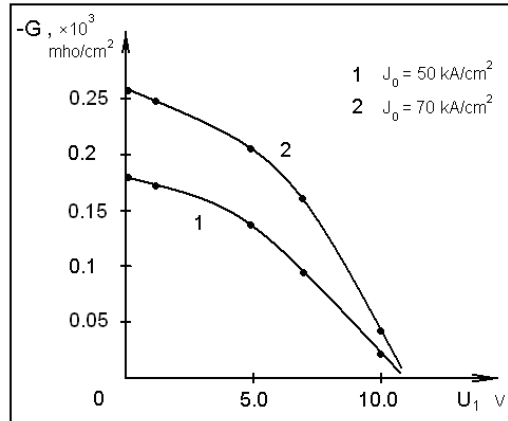
Fig. 17. Conductance *G* dependence as functions of first harmonic amplitude U_1 for f = 330 GHz and for two values of feeding current density

The conductance characteristic is softer for current density 50 kA/cm² because the diode structure optimization was provided for this current. The characteristics for 70 kA/cm² are sharper but correspond to the larger conductance –*G*.

The output power dependencies as a function of first harmonic amplitude U_1 for f = 330 GHz and for two values of feeding current are shown in Fig. 18.
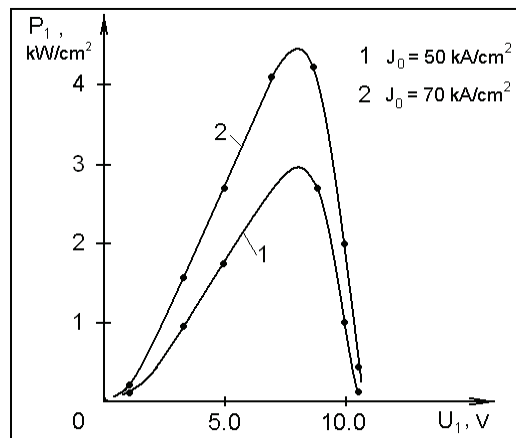
Fig. 18. Output generated power *P* dependence as functions of first harmonic amplitude U_1 for f = 330 GHz and for two values of feeding current density

These amplitude characteristics show the possibility to obtain a sufficient level of output power near the 4 kW/cm².

6. Conclusion

The numerical scheme that has been developed for the analysis of the different types of IMPATT diodes is suitable for the DDR and DAR complex doping profile investigation but in case of the DAR diode the numerical scheme convergence is slower.

Some new features of the DAR diode were obtained by the analysis on the basis of nonlinear model. The principal obtained results show that the diode does not have the active properties in some frequency bands for the sufficiently large drift region. To obtain the negative conductance for some frequency bands we need to reduce the drift layer widths significantly. Nevertheless the diode has a wide first frequency band generation and two superior frequency bands with sufficient output power level. The DAR diode analysis gives us the possibility to increase the output power level for the second and third frequency bands. The diode structure optimization gives the possibility to improve the admittance characteristics for high frequency bands.

7. References

Bakhvalov, N.S.; Zhidkov, N.P. & Kobelkov, G.M. (2008). *Numerical Methods*, Binom, ISBN 978-5-94774-815-4, Moscow.

Canali, C.; Jacoboni, C., Ottaviani, G. & Alberigi Quaranta, A. (1975). High Field Diffusion of Electrons in Silicon. *Appl. Phys. Lett.*, Vol. 27, p. 278.

Chang K. (Ed), (1990). *Handbook of Microwave and Optical Components*, Vol. 1, John Wile & Sons, N.Y..

Curow, M. (1994). Proposed GaAs IMPATT Devices Structure for D-band Applications. *Electron. Lett.*, Vol.30, pp. 1629-1631.

Dalle, C. & Rolland, (1989). Drift-Diffusion Versus Energy Model for Millimetric-Wave IMPATT Diodes Modelling. *Int. J. Numer. Modelling*, Vol. 2, pp. 61-73.

Datta D.N. & Pal B.B. (1982). Generalized small signal analysis of a DAR IMPATT diode. *Solid-State Electron.*, Vol. 25, No. 6, pp. 435-439.

Datta D.N., Pati S.P., Banerjee J.P., Pal B.B. & Roy S.K. (1982). Computer analysis of DC field and current-density profiles of DAR IMPATT diode. *IEEE Trans Electron Devices,* Vol. ED-29, pp. 1813-1817.

El-Gabaly, M.A., Mains, R.K. & Haddad, G.I. (1984). Effects of Doping Profile on GaAs Double-Drift IMPATT Diodes at 33 and 44 GHz Using the Energy-Momentum Transport Model. *IEEE Trans.*, Vol. MTT-32, No.10, pp.1353-1361.

Fong, T.T. & Kuno, H.J. (1979). Millimeter-Wave Pulsed IMPATT Sourse. *IEEE Trans.*, Vol. MTT-27, No.5, pp. 492-499.

Grant, W.N. (1973). Electron and Hole Ionization Rates in Epitaxial Silicon at High Electric Fields. *Solid-State Electronics*, Vol.16, No. 10, pp. 1189-1203.

Howes, M.J. & Morgan, D.V. (Eds.) (1978). *Microwave Devices.Devices Circuit Interactions*, John Wiley & Sons, N.Y.

Jacoboni, C., Canali, C., Ottaviani, G. & Alberigi Quaranta, A. (1977). A Review of Some Charge Transport Properties of Silicon. *Solid-State Electron.*, Vol. 20, pp. 77-89.

Joshi, R.P., Pathak, S. & Mcadoo, J.A. (1995). Hot-Electron and Thermal Effects on the Dynamic Characteristics of Single-Transit SiC Impact-Ionization Avalanche Transit-Time Diodes. *J. Appl. Phys.*, Vol. 78, pp. 3492-3497.

Kafka, H.J. & Hess, K. (1981). A Carrier Temperature Model Simulation of a Double-Drift IMPATT Diode. *IEEE Trans.*, ED-28, No.7, pp. 831-834.

Krylov, V.I., Bobkov, V.V. & Monastyrski, P.I. (1977). *Numerical Methods*, Nauka, Moscow.

Nava, F., Canali, C., Reggiani, L., Gasquet, D., Vaissiere, J.C. & Nougier, J.P. (1979). On Diffusivity of Holes in Silicon. *J. Appl. Phys.*, Vol. 50, p. 922.

Panda A.K., Dash G.N. & Pati S.P. (1995). Computer-aided studies on the wide-band microwave characteristics of a silicon double avalanche region diode. *Semicond Sci Technol.*, Vol. 10, pp.854-864.

Pati, S.P., Banerjee, J.P. & Roy, S.K. (1991). High frequency numerical analysis of double avalanche region IMPATT diode, *Semicond Sci Technol*, No. 6, pp. 777-783.

Read, W.T. (1958). A Proposed High-Frequency Negative-Resistance Diode. *Bell System Tech. J.*, Vol. 37, pp. 401-406.

Scharfetter, D.L. & Gummel, H.K. (1969). Large-Signal Analysis of a Silicon Read Diode Oscillator. *IEEE Trans.*, Vol. ED-16, No.1, pp. 64-77.

Som B., Pal B.B. & Roy S.K. (1974). A small signal analysis of an IMPATT device having two avalanche layers interspaced by a drift layer. *Solid-State Electron.*, Vol. 17, pp. 1029-1038.

Stoiljkovic, V., Howes, M.J. & Postoyalko, V. (1992). Nonisothermal Drift-Diffusion Model of Avalanche Diodes. *J. Appl. Phys.*, Vol.72, pp. 5493-5495.

Tager A.S., & Vald-Perlov, V.M. (1968). *Avalanche Diodes and Application on Microwave Endineering*, Sov. Radio, Moscow.

Tornblad, O., Lindefelt, U. & Breitholtz, B. (1996). Heat Generation in Si Bipolar Power Devices: the Relative Importance of Various Contributions. *Solid State Electronics*, Vol.39, No.10, pp. 1463-1471.

Vasilevskii, K.V. (1992). Calculation of the Dynamic Characteristics of a Silicon Carbide IMPATT Diode, *Sov. Phys. Semicond.*, Vol.26, pp. 994-999.

Zemliak, A.M. (1981). Difference Circuit Stability Analysis for IMPATT-Diode Design. *Izvestiya VUZ Radioelectronica*, Vol.24, No.8, pp. 88-89.

Zemliak, A.M. & Zinchenko, S.A. (1989). Non-Linear Analysis of IMPATT Diodes. *Vestnik K.P.I., Radiotechnika*, Vol.26, pp. 10-14.

Zemliak, A.M. & Roman, A.E. (1991). IMPATT Diode for the Pulsed-Mode. *Izvestiya VUZ Radioelectronica*, Vol. 34, No. 10, pp.18-23.

Zemliak, A., Khotiaintsev, S. & Celaya, C. (1997). Complex Nonlinear Model for the Pulsed-Mode IMPATT Diode. *Instrumentation and Development*, Vol. 3, No. 8, pp. 45-52.

Zemliak A., Celaya C. & Garcia R. (1998). Active layer parameter optimization for high-power Si 2 mm pulsed IMPATT diode. *Microwave and Opt Technol Lett.*, Vol. 19, No. 1, pp. 4-9.

Zemliak A. & De La Cruz R. (2002). An analysis of the active layer optimization of high power pulsed IMPATT diodes. *Comput & Systems.*, No. 6, pp. 99-107.

Zemliak, A.M. & De La Cruz, R. (2006). Numerical analysis of a double avalanche region
    IMPATT diode on the basis of nonlinear model. *Microelectronics Reliability*, Vol. 46,
    No. 2-4, pp. 293-300.

Ohmic Contacts for High Power and High Temperature Microelectronics

Lilyana Kolaklieva and Roumen Kakanakov
Central Laboratory of Applied Physics, Bulgarian Academy of Sciences
Bulgaria

1. Introduction

The increased requirements to the microelectronics regarding the device potential for work at high temperatures, high powers, and high frequencies and in harsh environments engendered the increased interest to the wide band-gap semiconductors. They are considered as a third generation materials in the semiconductor industry, after Si and Ge, and A^3B^5 compounds and their solid solutions. Several materials of the wide band-gap semiconductor group such as SiC, III-V nitrides (GaN, AlN, c-BN), ZnSe, and diamond are very important for the device industry. The unique combination of physical properties in these materials allows development of devices, which could be applied in fields where the devices of the first and second generations cannot be used. Whereas Si and GaAs are chemically stable at 400 ^{0}C and 650 ^{0}C, respectively, SiC and III-V nitrides are stable up to 1000 ^{0}C (Meyer & Metzger, 1996). This high thermal stability allows development of new class high temperature and high power devices with maximal working temperature of 600 ^{0}C, which is three and four times higher than this one of GaAs and Si devices, respectively.

Among the wide band-gap semiconductors, SiC and GaN have been most successfully applied in the device fabrication. These semiconductors offer a higher electric breakdown field (4-20 times), a higher thermal conductivity (3-13 times), and a larger saturated electron drift velocity (2-2.5 times) in comparison with silicon. These features make them very useful materials in development of high temperature and high power devices. The advantages of SiC and III-V nitrides technologies allowed manufacture of SiC-based and GaN-based devices such as unipolar high-voltage power FETs (MOSFET, JFET and HEMT), bipolar power diodes (p-n and p-i-n) and transistors (BJT, IGBT and HBT).

The existing applications present many challenges in obtaining high-performance ohmic contacts because they are limiting for device functioning. The ohmic contacts are a critical factor that could restrict the high power and high temperature device application. The high operating temperatures may cause diffusion processes in the contact layer and reactions between the contact components, which could lead to changes of the contact properties during operation at high temperatures, and deterioration of the devices. If the contact resistivity is not sufficiently low inadmissible high voltage drop could arise due to the high current density in the contact of the high power devices. Hence, the following requirements to the ohmic contacts are decisive for application in high power and high temperature microelectronics:

- *Low contact resistivity* – in general, the make of low resistivity ohmic contacts is difficult for wide band-gap semiconductors due to the difficulty in doping and, in the case of p-type materials, due to the wide forbidden band-gap.
- *High temperature stability* – this problem is very important in the wide band-gap semiconductors. In the Si and GaAs devices the maximal working temperature is limited by the material stability, because of that the problem of the contact stability is important but not critical. The great potential of SiC and III-V nitrides regarding the capacity for a work at temperatures up to 600 ^{0}C and higher, set strong requirements to the thermal stability and reliability of the contacts.
- *Reproducibility* – this requirement is important in the case of the device production. Therefore the contact technology should allow the achievement not only good performance, but good reproducibility.

The listed requirements point that the operation of high temperature and high power SiC and GaN-based devices under severe conditions demands development of electrically, thermally and chemically stable metal contacts.

2. Theoretical base of the ohmic contacts

The metal-semiconductor contact is one of the main elements of the semiconductor device structure, which parameters may significantly affect the device working characteristics.

When the metal comes into a contact with the semiconductor, a potential barrier is formed at the interface. Usually, this barrier has rectifying properties and it is named "Schottky barrier". Two types of metal-semiconductor contacts are known: ohmic and Schottky contacts. The Schottky contacts are metal-semiconductor contacts in which a Schottky barrier is formed. The ohmic contacts are metal-semiconductor contacts, which have linear and symmetrical I-V characteristic and a negligible contact resistance as compared to the bulk or series resistance of the semiconductor. They realize the connection between the chip and package in the semiconductor devices.

The presence of ohmic properties is determined by the shape and slope of the I-V characteristic. The main parameter characterized the ohmic contact is the resistivity (specific resistance), which is defined as (Yu, 1970; Sze, 1981)

$$\rho_c \equiv \left(\frac{\partial J}{\partial V} \right)_{V=0}^{-1} \tag{1}$$

According to the definition, the theoretical expressions for contact resistivity could be determined from the I-V characteristics taking into account the current transport mechanism through the contact (Yu, 1970). Four basic mechanisms are considered determinative for the current transport in the metal n-type semiconductor contact when a forward voltage is applied: 1) Emission of electrons from the semiconductor into the metal over the top of the barrier (thermionic emission). 2) Quantum-mechanical tunnelling through the barrier (field emission). 3) Recombination in the space-charge region. 4) Hole injection from the metal into the semiconductor. Depending on the carrier concentration the current transport through the contact is realized mainly by thermionic emission (TE) or field emission (FE). At a low semiconductor doping level ($N_D < 10^{17}$ cm^{-3}) the thermionic emission is prevailing. In the case of moderate doped semiconductors (10^{17} cm$^{-3} \leq ND \leq 1020$ cm-3) the depletion layer width decreases; the barrier becomes thinner and a part of electrons tunnel through it. With highly

doped semiconductors (ND>1020 cm-3) and low temperatures the current transport is determined by the field emission through the barrier only. These processes are defined by a characteristic energy E00 (Padovani & Stratton, 1966):

$$E_{oo} = \frac{q\hbar}{2}\sqrt{\frac{N_D}{m_n^* \varepsilon_s}} \; ,$$ (2)

where ℏ is a Plank's constant h divided by 2π; m_n^* is an effective electron mass in the semiconductor. The factor kT/E00 (where k is a Boltzmann's constant, and T is the absolute temperature) is a criterion for the interrelation between thermionic and field emission processes. The characteristic energy E00 correlates with the tunnelling probability and increases with the semiconductor doping level due to the decrease of the depletion layer width. The thermionic emission occurs at kT/ E00>>1. When kT/ E00≅1 these two processes are comparable and the current transport mechanism is named thermionic-field emission. The field emission is predominating at kT/ E00<<1.

With thermionic emission the contact resistivity depends on the potential barrier height only (Yu, 1970):

$$\rho_c = \frac{k}{q A^* T} \exp\left(\frac{q\varphi_{Bn}}{kT}\right).$$ (3)

In this case, metals formed a potential barrier with low height should be chosen to obtain contacts with low resistivity. With moderately doped semiconductors (thermionic-field emission), the resistivity is determined by both, the barrier height and the doping level of the semiconductor:

$$\rho_c = \left[\frac{k}{qA^*T}\right]\frac{kT}{\sqrt{\pi(\varphi_{Bn}+V_n)E_{oo}}}\cosh\left(\frac{E_{oo}}{kT}\right)\left[\sqrt{\coth\left(\frac{E_{oo}}{kT}\right)}\right]\exp\left[\frac{\varphi_{Bn}+V_n}{E_o}-\frac{V_n}{kT}\right],$$ (4)

where E0 is a measure of the probability for tunnelling through the potential barrier and

$$E_o = E_{oo}\coth\left(\frac{E_{oo}}{kT}\right).$$

The impurity concentration determines the contact resistivity with highly doped semiconductors. In this case the contact resistivity is changed exponentially by a factor of $\varphi_{Bn}/\sqrt{N_D}$ and it is determined by the equation:

$$\rho_c = \left[\frac{A\pi q}{kT\sin(\pi c_1 kT)}\exp\left(-\frac{\varphi_{Bn}}{E_{oo}}\right)-\frac{Ac_1 q}{(c_1 kT)^2}\exp\left(\frac{-\varphi_{Bn}}{E_{oo}}-c_1 V_n\right)\right]^{-1},$$ (5)

where A=A*T2 and

$$c_1 = \frac{1}{2E_{oo}}\ln\left[\frac{4\varphi_{Bn}}{V_n}\right].$$

The quality and reliability of the ohmic contacts have been evaluated by the behaviour of the main characteristic parameter, the contact resistivity. According to the definition of the contact resistivity it could be present as the contact resistance R_c multiplied by the contact area S:

$$\rho_c = R_c.S \tag{6}$$

Several methods for contact resistivity measurement are known: two probes method, differential method, extrapolation method, method of the interface probes, four probes method and Transmission line model method (TLM). TLM method is the mostly used method because it combines a low measurement error with a possibility of I-V characteristic linearity determination and low sizes of the test structures. Depending on the contact shape in the test structure the TLM method has two modifications, linear (Berger, 1972) and circular (Marlow & Das, 1982). The linear TLM method allows promptly determination of the contact resistivity despite that formation of mesa structures is needed. Formation of mesa structures is not necessary with the circular TLM method; however it is applicable at very low sheet resistance of the metals only. The values of the contact resistivity presented herein are determined using a linear TLM method.

As follows from the theory, two basic approaches could be used to create ohmic contacts: by increasing the semiconductor doping level and/or by decreasing the barrier height (Fig. 1).
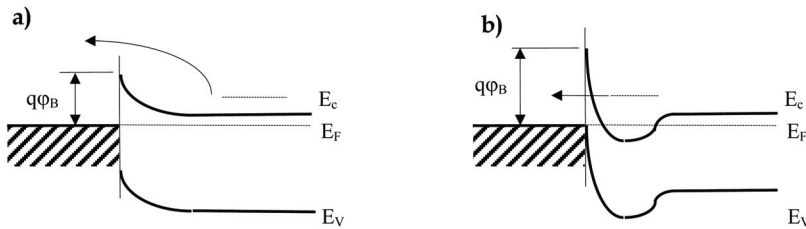
Fig. 1. A zone diagram of an ohmic contact with a) low barrier height and b) high doping level.

It is difficult to obtain low resistivity ohmic contacts to p-type wide band-gap semiconductors due to the high electron affinity and high width of the band-gap. For instance, the electron affinity of SiC and GaN is 3.3 eV and 1.84 eV, respectively. For the mostly used SiC polytypes, the band-gap width is in the interval 2.3-3.2 eV, while for GaN it is 3.44 eV for wurtzite polytype and 3.2 eV of the zinc-blended structure. Hence, a very high Schottky barrier is formed at the interface metal/p-type (SiC, GaN). A metal does not generally exist with a work function enough to yield a low barrier at the interface. In such cases the technique for making ohmic contacts involves the establishment of a heavy doped surface layer such as metal/p^+-p contact by various methods, such as shallow diffusion, alloy regrowth, in-diffusion of a dopant contained in the contact material. Annealing is the mostly used method for obtaining low resistivity contacts. After the metal film deposition the contacts are heated at the corresponding eutectic temperature for an optimal time in an ambient of an inert gas. As a result they are alloyed into the semiconductor or compounds lowering the barrier height are formed at the interface. The development of such methods of the modern microelectronics as molecular-beam epitaxy (MBE), metal-organic chemical vapour deposition (MOCVD) epitaxy and ion implantation allow obtaining high doping

level ($\geq 10^{20}$ cm^{-3}) of the epitaxial layers during the growth. By this technique "in-situ" ohmic contacts can be obtained without annealing.

3. Ohmic contacts to SiC

SiC is a material, which exists in over 130 polytypes. Among them only 6H, 4H and 3C are of interest for microelectronics. Between these three polytypes 4H-SiC has been mainly used for microelectronics devices due to the best combination of widest band-gap, highest breakdown voltage and highest electron mobility.

3.1 N-type contacts to SiC

Ohmic contacts with contact resistivity in an order of 10^{-6} Ω.cm^2 have been successfully developed firstly to n-type SiC. Different metals and its compounds such as Cr, Ni, TiN, TiW, W, Ti, Mo, Ta and etc., have been reported as suitable for ohmic contacts with resistivity in the range of ($10^{-2} \div 10^{-6}$) Ω.cm^2 (Porter & Davis, 1995; Crofton et al., 1997). In this section the properties of Ni-based contacts to n-type SiC are discussed because Ni has found to be the most appropriate metal for the device application. Ni and some metals such as Mo, Co, and W are found to form silicides in the metallization interface. During annealing of these contacts the reactions occurring at the interface are accompanied by liberation of carbon whose accumulation in the contact layer deteriorates the contact reliability during the device operation. In order to eliminate this unfavourable effect, a contact system consisting of multilayered Ni/Si films in the ratio 2Ni:Si, has been proposed instead of the pure nickel (Marinova et al., 1997). Herein, Ni/n-SiC and multilayered Ni/Si/n-SiC and Si/Ni/n-SiC contacts formed on substrates with a concentration of 1.8x10^{18} cm^{-3} (6H-SiC) and 8x10^{18} cm^{-3} (4H-SiC) are compared regarding the electrical, thermal and chemical properties. Studies on the electrical characteristics of these contacts have shown that after annealing at 950 ^{0}C for 10 minutes, Ni and Ni/Si layers form low-resistivity ohmic contacts to n-type SiC. The value of the resistivity depends strongly on the substrate doping concentration (Fig. 2). The resistivity of the contacts formed on substrates with the same doping concentration does not differ significantly by the contact composition. Values as measured by a linear TLM method have been determined in the interval (1.6÷2.9)x10^{-5} Ω.cm^2 for the contacts formed on substrates with carrier concentration of 1.8x10^{18} cm^{-3}. Increase the doping concentration to 8x10^{18} cm^{-3} affects on contact resistivity decrease by an order of magnitude and a value of 2.7x10^{-5} Ω.cm^2 has been measured with these substrates. The calculations show that the contacts with the same substrate doping level of 1.8x10^{18} cm^{-3} have a depletion layer width (potential barrier width, respectively) within the range (1.08 ÷ 1.11)x10^{-6} cm (Kassamakova-Kolaklieva, 1999). The higher doping concentration causes narrowing the depletion layer width to 6.06x10^{-7} cm, which results in decrease of the contact resistivity by an order of magnitude.

The mechanism of current transport through the Ni/SiC, Ni/Si/SiC and Si/Ni/SiC can be determined on the basis of the kT/E_{00} ratio. With the contact systems under consideration, a kT/E_{00} ratio of about 1 has been calculated with a doping concentration ranging from 1.7x10^{18} cm^{-3} to 1x10^{19} cm^{-3}. This result determines the thermionic-field emission as the main mechanism of current transport through the contacts. The good agreement between the experimentally obtained values of the resistivity and the theoretical dependence on the doping level calculated with a potential barrier height of (0.20 ÷ 0.35) eV confirms the
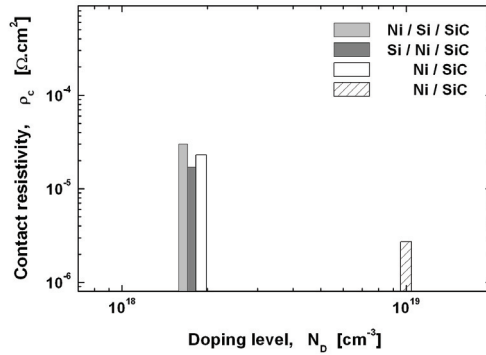
Fig. 2. Dependence of the resistivity of Ni-based contacts to n-type SiC on the substrate doping and the initial contact composition.

thermionic-field character of the current transport in the Ni, Ni/Si and Si/Ni ohmic contacts to n-type SiC (Fig. 3) (Kassamakova-Kolaklieva, 1999).
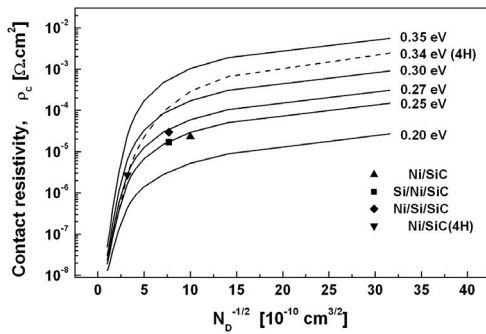
Fig. 3. Comparison of experimentally obtained resistivity values of Ni-based contacts to n-type 6H- and 4H-SiC with theoretical values calculated with different potential barrier heights.

X-ray photoelectron depth analyses (XPS) performed in order to understand the origin of ohmic properties in Ni-based contacts to n-SiC, have shown that the as-deposited polycrystalline nickel layer is homogeneous and a smooth surface is observed. The interface is chemically abrupt with a very thin amorphous layer, probably due to the ion bombardment prior to evaporation. Fig. 4a shows the XPS profile of a Ni/SiC contact after annealing at 950 ^{0}C. The Ni2p/Si2p peak ratio, as well as the binding energy of these peaks (respectively 853.2 eV and 99.4 eV), indicate the formation of a nickel silicide with a composition close to Ni_2Si (Grunthaner et al., 1980). Carbon in graphite state (Cls at 284.2 eV) is present in the whole contact layer with a maximal concentration at the interface. At the interface, the Ni2p peak remains at the same position while the maximum of the Si2p and C1s peaks are shifted towards the binding energies corresponding to SiC. The TEM

cross section of the annealed specimen presented in Fig. 5a confirms that the entire nickel layer has reacted to form a nickel silicide. The contact layer contains a lot of Kirkendall voids and its thickness has been increased substantially. The interface is shifted into the SiC, part of which has been consumed to supply Si for Ni_2Si formation. In the area of the original interface, an extremely high number of voids can be found. Quantitative EDS analysis indicates a composition close to Ni_2Si and strong carbon incorporation. Diffraction patterns from different grains could be indexed as the δ-Ni_2Si orthorhombic phase. These results suggest the following mechanism to describe the Ni/SiC contact formation after annealing at optimal temperature of 950 ^{0}C: (1) SiC dissociates due to the strong reactivity of nickel above 400 ^{0}C; (2) at 950 ^{0}C, the Ni_2Si stable phase is formed leading to carbon accumulation both at the interface and in the metal layer (Waldrop & Grant, 1993); and (3) a part of dissociated Si atoms diffuse through the nickel layer and simultaneously Ni atoms diffuse towards SiC until the complete consumption of the deposited nickel layer is realized.
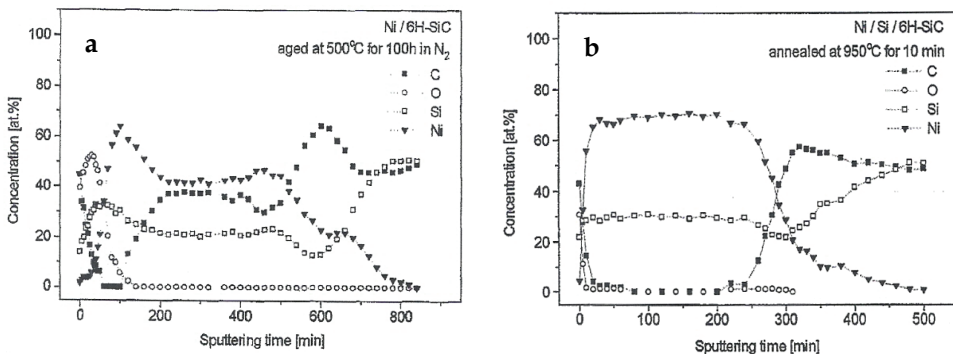
Fig. 4. XPS depth profiles of Ni/SiC (a) and Ni/Si/SiC (b) contacts after annealing at 950 ^{0}C for 10 minutes.

The depth distribution of the elements shows that similar profiles are observed for the both contacts, Ni/Si and Si/Ni after annealing at 950 ^{0}C for 10 min, when silicon was introduced in the nickel layer in order to prevent the SiC dissociation during the contact formation. XPS profile and TEM micrograph image of the Ni/Si/SiC contact are presented in Fig. 4b and Fig. 5b, respectively. A Ni_2Si layer is obtained as indicated by the binding energies and the ratio of Ni2p to Si2p signal intensity. There is no carbon contained in the silicide layer. At the interface, carbon is still observed but the amount is lower than for the previous Ni contacts. The bright field TEM image of the contact obtained with Si as first deposited layer reveals that the contact layer is uniform, polycrystalline and the δ-Ni_2Si orthorhombic phase is identified. Some Kirkendall voids are still present at the interface but not in the contact layer itself. In the case of the nickel interfacial layer, the contact morphology is similar while a greater number of voids is observed. These results suggest that intentional silicon incorporation in the nickel layer modifies the diffusion processes, which are responsible for the contact formation. In the case of Ni/Si multilayers, Ni and Si mutual diffusion occurs, leading to the stable phase of the nickel silicide, Ni_2Si. The reaction between Ni and SiC at the interface is limited since almost all the nickel is already bonded to the silicon atoms. As a consequence, the SiC decomposition is reduced and only a small amount of carbon is released.
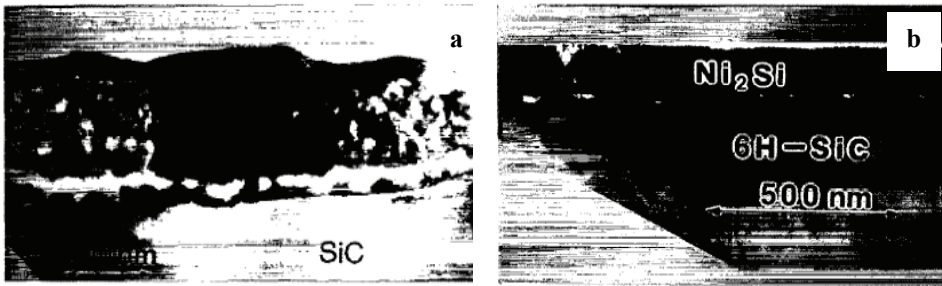
Fig. 5. TEM micrographs of the Ni/SiC (a) and Ni/Si/SiC (b) interfaces after annealing at 950 ^{0}C for 10 minutes.

3.2 P-type contacts to 4H-SiC

As it was mentioned, the combination of high electron affinity with wide band-gap in p-type SiC causes high barrier formation at the interface metal/p-SiC, which strongly hampers obtaining low resistivity ohmic contacts. In the case of p-type 4H-SiC a value of the barrier height has been estimated to be over 6 eV. Hence, suitable for p-type ohmic contacts to 4H-SiC could be metals having a work function in order of 7 eV, which is unrealistic. The metals used in microelectronics have work functions between 4 eV and 5.5 eV. For that reason, it is impossible to create ohmic contacts with low resistivity relaying the barrier decrease by selection of a suitable metal only.

A large variety of alloys, metals and compositions have been proposed as suitable for ohmic contacts to p-type SiC. (Porter & Davis, 1995; Crofton et al., 1997) Depending on the contact composition they could be classified in two main groups, 1) Al-based contacts and 2) Al-free contacts. The Al-based contacts consist of Al, its alloys or multilayers one of which is Al. It has been considered that during annealing of these contacts, Al diffuses into SiC, which causes increase of the p-type concentration in the interface layer. As a result, the depletion layer width is decreased and the p-type carriers can tunnel effectively through the potential barrier. Usually, for Al-free contacts refractory metals having a high work function and/or forming compounds, which create low and thin interface barrier, are used.

3.2.1 Al-based contacts to 4H-SiC

Among the all elements (Al, B, Ga, In, Be) used in the present semiconductor technology, Al is the most suitable dopant in the growth of p-type SiC. For that reason, the compositions containing Al have been considered as very appropriate material for p-type ohmic contacts to SiC. Al is the first metal suggested for an ohmic contact to p-SiC. However, the application of pure Al metallization is restricted by formation of pits during annealing which worsen the contact morphology and conductance. In analogy with the silicon technology, Al/Si(1-2 wt.%) composition has been proposed to avoid this problem. Nevertheless, the low diffusion coefficient of Al into SiC requires high annealing temperatures, which enhances the tendency of metal film oxidation and consequently the contact resistivity increases. Therefore, less oxidized refractory metals have been added into the Al-based contact systems. The mostly used metal for this purpose is Ti (Crofton et al.,

1993). The titanium presence in the metallization scheme prevents the Al oxidation and allows its diffusion into SiC. The used thick upper Ti layer acts as a barrier for the Al volatilization observed during the contact annealing. Although other Al-based contacts have been proposed subsequently, the Ti/Al contact still remain the most applied p-type contact in the SiC devices. These contacts could be obtained using an Al-Ti alloy or by subsequent deposition of Al and Ti multilayers (layered contacts).

The electrical characteristics of four typical Al-based contacts are presented in Figs. 6. and 7. (Kakanakov et al., 2001; Kolaklieva et al., 2004; Kolaklieva et al., 2007). The contacts are formed on p-type 4H-SiC epitaxial layers with a thickness of 1 μm and a carrier concentration of 3×10^{19} cm^{-3}. The Au/Al/Si and Au/Ti/Al contacts are multilayered and the films are successively deposited. The thicknesses of the Si, Ti and Al component films were nanoscaled and chosen according to the ratios: Si (2 wt.%) in Al/Si, Ti (70 wt.% and 30 wt.%) in Ti/Al and Al (30 wt.% and 70 wt.%) in Ti/Al before annealing. The total thickness of these films is 100 nm. The same thickness has an AlSi(2 wt.%)Ti(0.15 wt.%) alloy contact. In all contact types a 100 nm thick Au film is deposited as a cap layer.
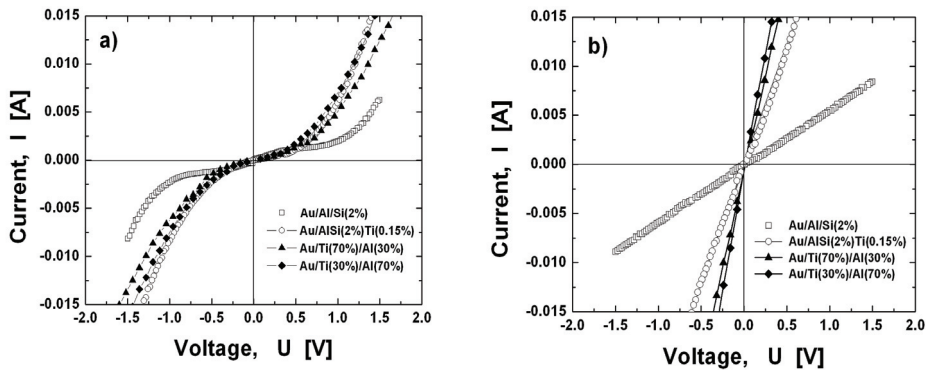
Fig. 6. I-V characteristics of Al-based contacts to p-type SiC: a) as-deposited; b) annealed at an optimal temperature.

I-V characteristics of all as-deposited Al-based metallizations have a shape typical of the Schottky barrier, which determines the rectifying behaviour of the unannealed contacts. They do not differ significantly, which is expectable because of the same substrate doping concentration and the same metal film at the interface. Small difference corresponding to higher potential barrier is observed with the Au/Al/Si contact because of the different element forming the interface with SiC. The initial contact composition and composite ratio influence on the annealing process and dependence of the contact resistivity on the temperature. Fig. 6b presents I-V characteristics of the contacts after annealing at optimal temperature, at which lowest contact resistivity has been observed (Fig. 7). The I-V characteristics of the annealed contacts exhibit various slopes implying different contact resistivities. The smaller slope corresponds to the higher resistivity value, which is confirmed by the results from the investigation of the contact resistivity dependence on the annealing temperature. The addition even of a little titanium amount to the contact composition causes an increase in the annealing temperature at which the lowest contact

resistivity has been obtained, from 700 ^{0}C (for Au/Al/Si) to 1000 ^{0}C (for Au/Ti/Al). The presence of Ti in the contact composition also affects on the resistivity decrease to a value of $1.2x10^{-5}$ Ω.cm^2 compared to that of the Au/Al/Si contacts ($2.5x10^{-4}$ Ω.cm^2).
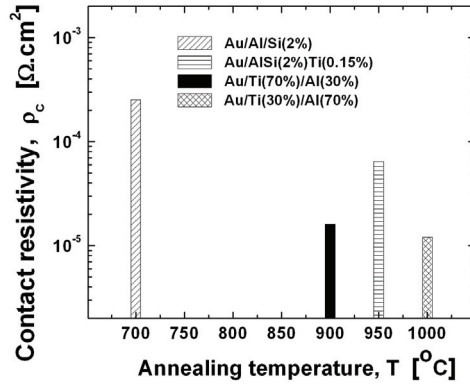
Fig. 7. Contact resistivity obtained for Al-based contacts annealed at optimal temperatures: Au/Al/Si – $2.5x10^{-4}$ Ω.cm^2; Au/AlSITi – $6.4x10^{-5}$ Ω.cm^2; Au/Ti(30%)/Al(70%) - 1.410^{-5} Ω.cm^2 and Au/Ti(70%)/Al(30%) - $1.2x10^{-5}$ Ω.cm^2.

The surface morphology of Al-based contacts obtained by Atomic Force Microscopy (AFM) demonstrates strong dependence on the contact composition and weight percentage of each contact component as well as the annealing conditions (Tabl. 1) (Kassamakova et al., 2001; Kolaklieva et al., 2007). AFM images taken from (2x2) μm^2 (for Au/Al/Si and Au/AlSiTi) and (10x10) μm^2 (for Au/Ti/Al) area reveal the granular structure of the as-deposited contacts. The higher Al amount in the as-deposited contacts causes raising the surface roughness, which originates from the specific feature of Al to form drops during evaporation. After annealing at the optimal temperature the surface roughness increases with the annealing temperature and the Al amount in the contact.

Contact type	Al/Si(2%)	AlSi(2%) Ti(0.15%)	Au/Ti(70%)/ Al(30%)	Au/Ti(30%)/ Al(70%)
R_{MS} (as – deposited contact)	3 nm (2x2) μm^2	9 nm (2x2) μm^2	8 nm (10x10) μm^2	22 nm (10x10) μm^2
R_{MS} (annealed at optimal temperature contact)	13 nm (2x2) μm^2	16 nm (2x2) μm^2	28 nm (10x10) μm^2	101 nm (10x10) μm^2

Table 1. Surface roughness of as-deposited and annealed at optimal temperature Al-based contacts.

XPS depth analyses of the contact composition and interface chemistry of the as-deposited Al-based contacts reveal an abrupt metal/SiC interface. No interdiffusion between the as-deposited multilayers is observed and well-expressed borders between them are detected. The different compositions and resulting different annealing temperatures led to remarkable

differences in element distribution and interface chemistry of the contacts (Fig. 8). The deposition of silicon between the Al layer and the SiC substrate, and the relatively low annealing temperature reduce the interdiffusion/chemical reaction processes. As a result a significantly more abrupt interface in the Al/Si/SiC contact after annealing at 700 ^{0}C has been observed (Fig. 8a). The analysis of the photoelectron spectra shows that interdiffusion and chemical reactions during annealing at 950 ^{0}C lead to the transformation of the initial AlSiTi alloy layer (fig. 8b). Due to the catalytic effect of Al at elevated temperatures SiC dissociation occurs at the metal/SiC interface. A part of C reacts with Al to form Al_4C_3 (BE 283 eV), while the other part is registered as graphite (BE 284.5 eV) in the contact layer (0-200 min sputtering). The broadened to the lower binding energy C1s peak (BE 283 eV) after 100 min to 150 min sputtering could be assumed as overlapping peaks of C in TiC and Al_4C_3. A weak Si2p peak determined in the contact layer corresponds to the Si-Si bond. After 200 min sputtering aluminium in metal state only has been detected. The intensity of Al2p peak typical for Al in metal state does not change significantly in the sputtering interval 200-500 min, which suggests diffusion of Al atoms into SiC and widening the interface region (Kassamakova et al., 2001).

Likewise, annealing of the Au/Ti/Al contact changes essentially the element distribution in the contact layer and at the interface as the process determines by the Ti/Al ratio (Figs. 8c and 8d; The different sputtering time is due to the different sputtering ratio, not different contact thickness). The XPS depth profiles allow dividing the contact structure into three regions: surface, film and interface. In both contacts strong Al diffusion to the surface
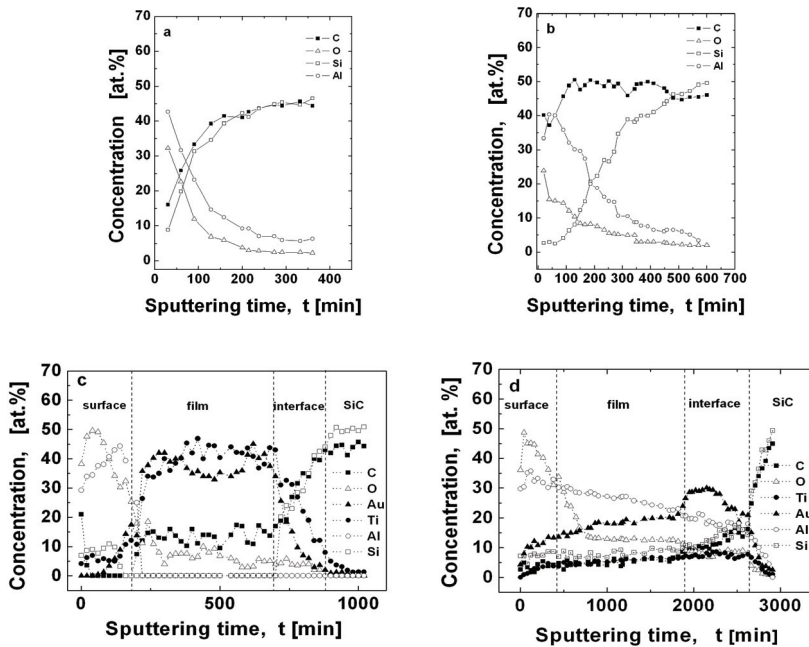
Fig. 8. XPS depth profiles of Al-based contacts: a) Al/Si/SiC; b) AlSiTi/SiC; c) Au/Ti(30%)/Al(70%)/SiC; d) Au/Ti(30%)/Al(70%)/SiC.

promoted by the thermal treatment is obtained. With prolonged sputtering, the Ti/Al=(70/30) wt.% film (Fig. 8c) shows a simultaneous increase in gold and titanium concentrations. The shift of the binding energy of gold in this region up to 84.7 eV is probably connected with the change in chemical surroundings of the Au atoms by Ti atoms and it could be associated with formation of an Au(35at%)+Ti(42at%) alloy. The position of the C1s peak reveals presence of TiC in the film region. Quite a different element distribution is observed for the Ti/Al=(30/70) wt.% contact (Fig. 8d). The film region consists mainly of Au and Al as a decrease in Al and increase in Au concentrations is detected.

The concentrations of Ti, Si and C remain almost constant and formation of Ti_3SiC_2 compound is possible. Simultaneously a shift of the Au4f peak core level to a higher binding energy (84.2 eV) has been detected up to 1200 min sputtering. This could be associated with change in chemical surroundings of the Au atoms by Al. The formation of an Au-Al alloy is not excluded. Between 1200 and 1800 min sputtering, Au in metal state and formation of TiC is detected. The interface region of the Ti/Al=(70/30) wt.% contact has been found to be narrow. The annealing at 900 ^{0}C causes dissociation of the SiC surface and dissociated carbon interacts with Ti forming TiC. The interface region for the contact with a Ti/Al=(30/70) wt.% ratio annealed at 1000 ^{0}C is remarkably wider. Along the interface increase of Au, Si and C concentrations has been observed, accompanied by a decrease in Al concentrations. The shift of the Au4f binding energy to the higher values has been determined, which could be connected with silicide formation. Appearance of TiC as a result of the SiC surface dissociation has been also detected. No carbon in graphite state and nonbonded silicon has been obtained (Kolaklieva et al., 2007).

3.2.1 Pd-based contacts to 4H-SiC

Palladium is a metal appropriate for ohmic contacts to p-type semiconductors due to its high work function (5.12 eV). Palladium ohmic contacts are successfully used in GaAs devices. It is known that the silicides of noble metals such as Ir, Pt and Pd are suitable for ohmic contacts to p-type Si. Palladium reacts with SiC at relatively low temperatures (~ 500 ^{0}C) and forms silicides, which are considered to be contributory to the barrier height decrease. Besides, recently Pd is a widely used metal in SiC chemical and gas sensors intending to operate at high temperatures. In earlier works Pd is reported as a very promising metal for low resistivity ohmic contacts to p-type SiC (Kassamakova et al., 1999, Kassamakova-Kolaklieva et al., 2003).

The properties of Au/Pd and Au/Pd/Ti/Pd ohmic contacts will be compared in this section. They are deposited on the substrates as the Al-based contacts. As-deposited Au/Pd and Au/Pd/Ti/Pd contacts show Schottky barrier behaviour. The dependence of the contact resistivity on the annealing temperature is different for both contact compositions (Fig. 9). Ohmic properties for the Au/Pd contacts are observed after annealing at 600 ^{0}C and a resistivity of 7.2×10^{-4} $\Omega.cm^2$ has been measured at this temperature. The contact resistivity decreased smoothly with the temperature increase up to 850 ^{0}C. After annealing at this temperature a lowest resistivity of 4.2×10^{-5} $\Omega.cm^2$ was obtained for these contacts. The dependence of the resistivity on the annealing temperature is steeper for the Au/Pd/Ti/Pd contacts. Annealing at temperatures of 600 ^{0}C and 650 ^{0}C does not change the Schottky behaviour. They become ohmic after annealing at 700 ^{0}C, but the contact resistivity is still high, 3.3×10^{-3} $\Omega.cm^2$. A lowest reproducible resistivity of 2.9×10^{-5} $\Omega.cm^2$ has been
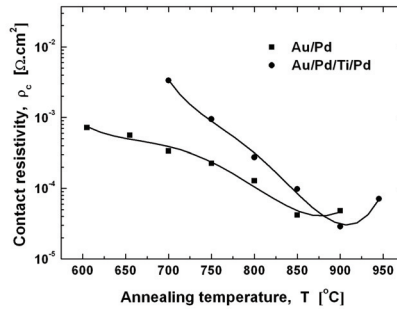
Fig. 9. Dependence of the resistivity of Pd- based contacts on the annealing temperature.

obtained after annealing at temperature of 900 ^{0}C. This result shows that addition of the refractory titanium into the contact composition shifts the optimal annealing temperature to the higher values. Further increase of the annealing temperature of both contact types causes a resistivity increase. Consequently, the annealing temperature of 850 ^{0}C and 900 ^{0}C can be accepted as optimal for ohmic properties formation of the Au/Pd and Au/Pd/Ti/Pd contact composition, respectively.

The different annealing techniques cause different surface morphology of the Pd-based contacts. The surface morphology of the as-deposited contacts follows the surface features (terraces) of the SiC substrate with a surface roughness of around 1.2 nm and a mean grain size of around 100 nm. After RTA annealing at temperature optimal for both contact types the surface reveals an altered granular structure of the metals. Both grain size and surface roughness are increasing to values of 1-3 µm and 50-100 nm, respectively (Fig. 10 a, b). The annealing in a resistance furnace leads to improved surface morphology and contact properties. The surface roughness (RMS=13 nm) and the mean grain size (150 nm) are smaller (Fig 10 c). However, in this contact structure a strong interdiffusion occurs as can be seen in the AFM image of a scan across the border of a contact pad (Fig. 10 d), where the initial step height between SiC surface and contact pad almost vanishes.
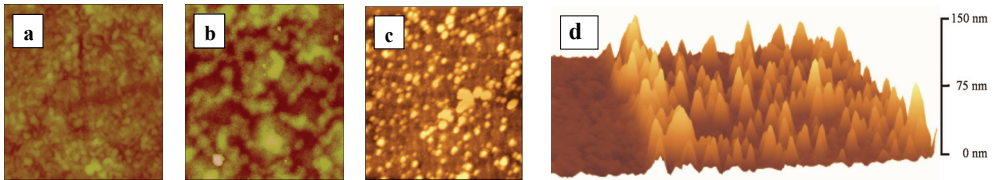
Fig. 10. AFM images: ((25x25) µm², z-scale 1 µm) of RTA annealed Au/Pd/SiC contacts at 850 ^{0}C (a) and Au/Pd/Ti/Pd/SiC at 900 ^{0}C (b); of RFA annealed Au/Pd/Ti/Pd/SiC contacts at 900 ^{0}C - 2D image (5x5 µm², z-scale 50 nm) (c); and 3D image (10x10) µm² across the border between an annealed contact pad and the SiC substrate (d).

The XPS depth profile of the as-deposited Pd/SiC and Au/Pd/Ti/Pd/SiC contacts show a steep interface metal/SiC as well as steep interfaces between the metals forming the contact composition (Kassamakova et al., 1999; Kolaklieva et al., 2004). After annealing at an optimal temperature the contact composition changes completely (Fig.11). Annealing of the

Pd/SiC contact at 700 ^{0}C initializes dissociation of SiC surface in the presence of Pd atoms. The released Si atoms interact with palladium to form palladium silicide while the dissolved carbon atoms start to accumulate at the interface. The XPS spectra have established the presence of the two palladium silicides Pd_3Si and Pd_2Si together with carbon in graphite state distributed in the whole contact film. As a result, the SiC interface is shifted into the SiC bulk, since a part of the original interface is consumed to supply Si for the Pd_3Si formation. After annealing of the Au/Pd/Ti/Pd contact, a new contact composition has been obtained. The contact layer consists of Au in a metal state, unreacted Pd, palladium rich silicide (Pd_3Si) and TiC, while the interface layer is composed of a less Pd-rich silicide (Pd_2Si). As in the Pd/SiC contact a part of the original interface is consumed due to the partial dissociation of SiC to Si and C. Again, the free Si atoms interact with Pd to form Pd_2Si in the interface near region and Pd_3Si in the more remote contact layer, while the dissolved C atoms react with Ti and TiC is formed. Due to the presence of Ti in the contact composition, the carbon resulting from SiC dissociation during annealing is completely consumed. It should be noted that in contrast to the Pd/SiC contact, no carbon in graphite state has been observed in the annealed Au/Pd/Ti/Pd contact. The absence of free C in the annealed contact causes improvement of the contact stability during the long-term treatments and at high operating temperatures. The presence of Au and Pd in metal state contributes to the good contact conductivity.
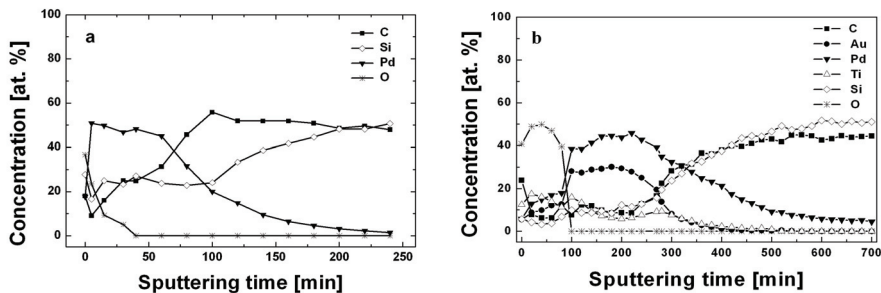
Fig. 11. XPS depth profiles of Pd-based contacts: a) Pd/SiC annealed at 700 ^{0}C and (b) Au/Pd/Ti/Pd/SiC annealed at 900 ^{0}C.

3.3 Thermal stability of n- and p-type ohmic contacts to SiC.

By contrast with the Si and GaAs devices, which operating temperature is limited by the electronic properties of the semiconductor material, the maximum operating temperature of SiC and III-nitride devices is limited by stability of the contacts. Some device parameters such as response time, output power and etc. depend strongly on the ohmic contact resistivity and its stability at high operating temperatures. Therefore the contact reliability at high temperature treatment is considered as the critical factor determining their power application.

The thermal stability of the contacts consists in their parameters remaining unchanged under the effect of the temperature. This property is investigated on the basis of the behaviour of a physical or electrical parameter characterising the contact under the effect of the temperature. For ohmic contacts such parameter is the resistivity. Usually, the thermal stability of ohmic contacts is investigated for long time treatment at fixed temperatures

(ageing test) and by the dependence of the resistivity on the dynamically increasing temperature (temperature-dependence test).

In this section the thermal properties of Ni-based, Al-based and Pd-based ohmic contacts to SiC are presented (Kakanakov et al., 2004; Kolaklieva et al., 2004; Kassamakova-Kolaklieva et al., 2003). The effect of the long term ageing of the contacts on the electrical properties has been studied by heating at 500 °C, 600 °C and 700 °C for 100 hours at each temperature. In fixed time intervals the contacts are cooled to room temperature and the contact resistivity is measured. The results from this study are summarized in Fig.12. All contacts show non-essential change of the resistivity during 100 hours ageing at 500 °C. Both Pd-based contact types have demonstrated good thermal stability at 500 °C heating for 100 hours. Increase of the ageing temperature to 600 °C results in different contact behaviour. A significant effect of the thermal treatment at this temperature is observed on the electrical properties of the Au/Pd contacts. After 24 hours heating their contact resistivity increases to a value of 1.4×10^{-4} $\Omega.cm2$. Further heating at this temperature does not deteriorate them. On the contrary, the Au/Pd/Ti/Pd contacts show excellent thermal stability during ageing at 600 0C and 700 0C. The improved thermal stability of Au/Pd/Ti/Pd ohmic contacts can be explained by formation of a thermodynamically stable contact configuration during annealing. The annealing of the Au/Pd contacts results in formation of Pd_2Si at the interface. Pd_2Si is the Pd-richest silicide, which is in thermodynamic equilibrium with SiC. Therefore it is considered as a metallization to SiC stable during prolonged thermal treatments. However, the formation of palladium silicides during annealing leads to the accumulation of free C within the contact layer, which is responsible for the observed instability of Au/Pd contacts during the long term ageing at higher temperatures. During annealing of the Au/Pd/Ti/Pd contacts two processes run: formation of Pd_2Si at the interface and reaction between the titanium and the free carbon in the contact layer. The latter leads to the formation of the thermodynamically stable TiC compound phase and reduction (or total use up) of the free C in the contact layer, which results in improving of the thermal stability of the contacts.
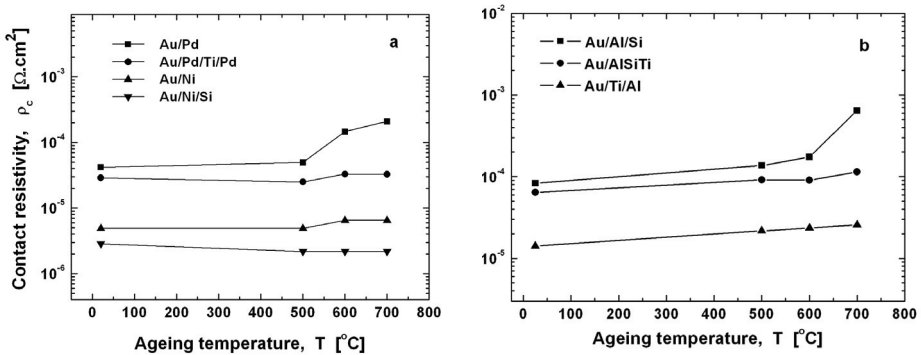
Fig. 12. Dependence of the contact resistivity on the long-term temperature treatment of: (a) Ni-based and Pd-based contacts, and (b) Al-based contacts.

Increase of the ageing temperature to 600 °C causes a very small rise of the resistivity of the Au/Al/Si contact. The resistivity of both contacts, Au/AlSiTi and Au/Ti/Al, remain

practically the same during the whole time interval at this temperature. During heating at 700 ^{0}C, the Au/Al/Si contact resistivity increases continuously to a value of 6.4x10^{-4} Ω.cm2 measured after the 100th hour. Slight increase of the resistivity from 9.1x10-5 Ω.cm2 to 1.2x10-4 Ω.cm2 is noticed for the Au/AlSiTi contact with the same test. No practical changes in the contact resistivity are detected when the Au/Ti/Al contact is subjected to ageing at 700 ^{0}C for 100 hours. The addition of Ti to the contact composition improves its thermal and power properties. This effect is less pronounced in the Au/AlSiTi contacts because of the very small Ti amount in the contact composition. Due to the higher Ti concentration the carbon resulted from the SiC dissociation during annealing is completely consumed and TiC is formed in the contact layer. The absence of C in graphite state is the main factor, which ensures the stability of Au/Ti/Al contact during the ageing up to 700 ^{0}C.

The resistivity of Ni-based contacts remains practically the same in the whole time interval at these temperatures. Small instability has been observed with Au/Ni contacts after ageing at 600 ^{0}C, but the resistivity remains still low. The observed excellent thermal stability of these contacts is due to the formation of the chemically stable interface with the semiconductor and a stable contact composition of Ni_2Si.

In the temperature-dependence test the measurements have been proceeded at a temperature increasing smoothly from 25 ^{0}C to 450 ^{0}C in air. This study gives information on the contact reliability at the corresponding operating temperature as the contact resistivity has been measured during the heating. For the temperature-current treatment, a current with a pre-set density of 10^3 A/cm^2 is supplied for a fixed time at a constant temperature (up to 450 ^{0}C). This test has been also performed in air and contact resistivity is measured at the corresponding temperature. The results from the two tests are presented in Fig. 13.
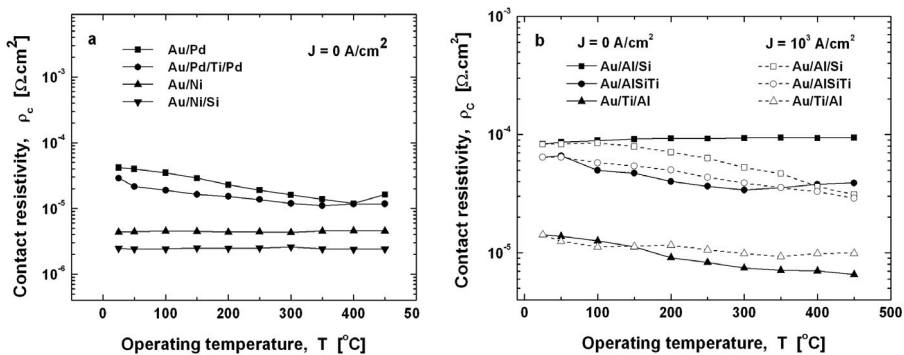
Fig. 13. Dependence of the contact resistivity on the operating temperature and supplied power of: (a) Ni-based and Pd-based contacts, and (b) Al-based contacts

Au/Pd/Ti/Pd contacts have demonstrated better stability at operating temperatures in the interval 25 ^{0}C – 450 ^{0}C in air. For the Au/Pd contacts the contact resistivity decreases twofold as the temperature increased from 25 ^{0}C to 450 ^{0}C. Similarly, the contact resistivity of the Au/Ti/Al contact decreases with temperature, however at a slow rate. A slow rate decrease is also observed with the Au/AlSiTi contacts from 25 ^{0}C to 300 ^{0}C. Further

temperature increase to 450 ^{0}C causes increase of the resistivity of these contacts. However, the resistivity value measured at 450 ^{0}C is still lower than this one determined at 25 ^{0}C. The resistivity of the Au/Al/Si contact remains practically the same at all temperatures from 25 ^{0}C to 450 ^{0}C. All Al-based contacts have shown a resistivity decrease when a current with a density of J=10^3 A/cm^2 is supplied during the heating. The Ni-based contacts do not change the resistivity during this treatment. After the test is completed and the samples are cooled down the contact resistivity is measured again at 25 ^{0}C. The contact resistivity obtained does not differ from the values measured for each contact type before the test.

4. Ohmic contacts for HEMTs based on GaN/AlGaN heterostructures

For the last years III-nitrides have been received great attention as a material having big potential for short-wave optoelectronic as well as RF and power microelectronic device applications. High electron mobility transistors (HEMTs) based on AlGaN/GaN heterostructures are very appropriate for high frequency and high power devices because of the intrinsic material properties such as wide band gap, high breakdown field, and high electron saturated velocity. The low resistivity, excellent reliability at elevated temperatures and good reproducibility of the ohmic contacts are critical factors, which limit the optimum HEMT performance. Besides these requirements, the smooth surface morphology is essential to facilitate sharp edge acuity for short channel devices. Large variety of metal schemes have been proposed and studied as ohmic contacts to AlGaN/GaN HEMTs. Among them Ti/Al-based system has become the conventional widely used ohmic contacts. Such metal scheme could be described as Ti/Al/X(Ni, Ti, Mo, Pd, Pt)/Au.

Multilayered Ti/Al/Ti/Au metal films are one of the mostly used metallizations for obtaining ohmic contacts to HEMTs (Fig. 14a) (Kolaklieva et al., 2008). In the device technology, it is known that Al tends to ball up during contact annealing. This behaviour results in a rough surface morphology of the Ti/Al-based contacts. The first Ti layer being in intimate contact with the GaN or AlGaN interface takes essential role in ohmic properties formation during annealing. Besides, during annealing of these contacts Al reacts with Ti forming Ti$_x$Al$_{1-x}$ alloys, whose presence in the contact contributes to the contact conductivity. Therefore, investigations have been carried out toward a search for the appropriate initial ratio between the former Ti layer and subsequent Al film (Ti/Al) (Fig. 14b), which enables obtaining low resistivity ohmic contacts with a smooth surface.
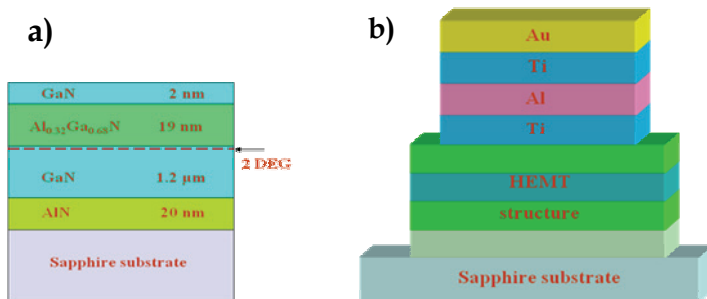
Fig. 14. Schemes of: a) a HEMT structure, and b) an as-deposited contact.

I-V characteristics of all as-deposited Ti/Al/Ti/Au metallizations coincide completely because of the same carrier concentration of the upper GaN layer and the same Ti interface metal layer (Fig. 15a) (Kolaklieva et al., 2009). They have a shape typical of the Schottky barrier, which determines the rectifying behaviour of the contacts. After annealing at temperatures higher than 700 ^{0}C the I-V characteristics become linear indicating ohmic contact properties. The I-V characteristics of the Ti/Al (30/70 wt.%) and Ti/Al (50/50 wt.%) contacts coincide completely (Fig. 15 b). This result is expectable because these contacts show the same resistivity after annealing at optimal temperature (Fig. 16a). The I-V characteristic of the Ti/Al (70/30 wt.%) contact exhibits smaller slope implying higher resistivity, which is confirmed by the TLM measurements (Fig. 16 a). For the Ti/Al (30/70 wt.%) and Ti/Al (50/50 wt.%) contacts, ohmic properties have been obtained after annealing at a temperature as low as 700 ^{0}C, but the contact resistivity is still high, especially for the contact with higher Ti content. For the Ti/Al (70/30 wt.%) contact, ohmic properties have been observed after annealing at 750 ^{0}C. The behaviour of the three contact compositions does not differ essentially in character. There is a tendency to shift to higher
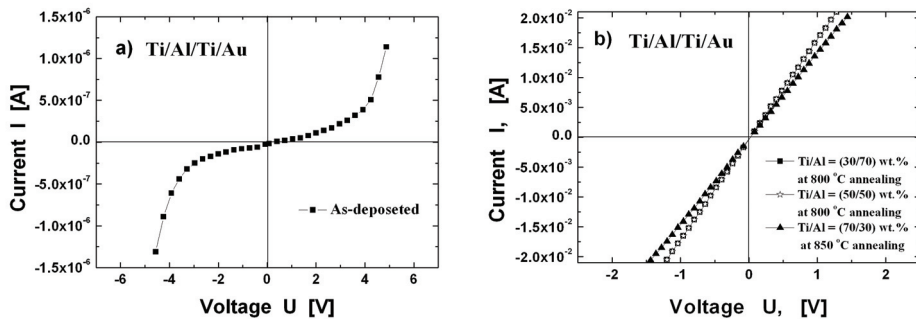
Fig. 15. I-V characteristics of as-deposited (a) and annealed at optimal temperature (b) Ti/Al/Ti/Au contacts with a different Ti:Al ratio.
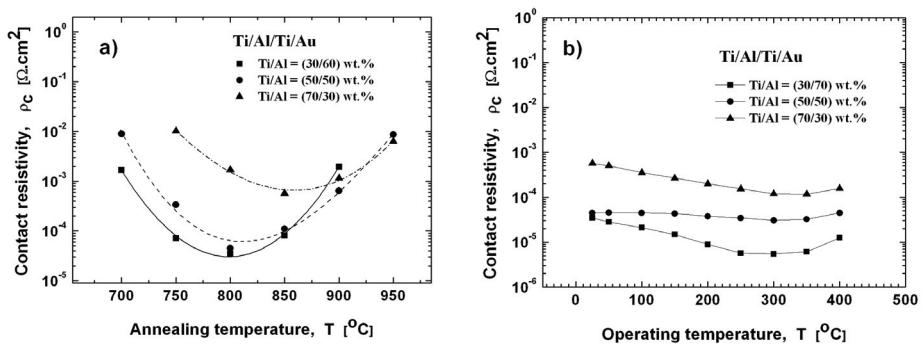
Fig. 16. Dependence of the resistivity of Ti/Al/Ti/Au contacts with a different Ti/Al ratio on the annealing temperature (a) and operating temperature (b).

optimal annealing temperatures with increasing Ti content in the former-Ti/Al layer, which is expectable. The contact resistivity of the Ti/Al (30/70 wt.%) and Ti/Al (50/50 wt.%) contacts decreases smoothly to 800 ^{0}C, at which temperature it reaches a minimum value of 4.2×10^{-5} $\Omega.cm^2$ and 4.4×10^{-5} $\Omega.cm^2$, respectively. For the Ti/Al (70/30 wt.%) contact, the lowest resistivity of 5.7×10^{-4} $\Omega.cm^2$ is measured after annealing at 850 ^{0}C. Further increase of the annealing temperature causes increase of the contact resistivity. This resistivity increase could be explained by out-diffusion of Ti and Al to the Au layer and their oxidation at the contact surface, which processes are intensified at high temperatures. The presence of aluminium oxide at the surface has been detected by XPS analysis, which confirms this suggestion.

The investigation on the thermal properties of the three types of contact compositions has been performed in air at a temperature increasing smoothly from 25 ^{0}C to 400 ^{0}C. Obviously, different initial contact composition causes different thermal behaviour (Fig. 16 b). The best stability shows the contact with Ti/Al ratio of 50/50 wt.%. Its resistivity practically does not change up to 350 ^{0}C. Both other contact compositions exhibit smooth decrease of the contact resistivity with temperature increase. A fourfold resistivity drop is found to occur over the whole temperature interval for the contact with Ti/Al ratio of 70/30 wt.%, while six fold resistivity drop of the Ti/Al (30/70 wt.%) contact follows heating under the same conditions. This result shows that higher Ti content causes enhanced stability at operating temperatures up to 400 ^{0}C in air.

AFM measurements (Fig. 17) reveal that the surface strongly roughens upon annealing and randomly distributed hillocks appear in dependence on the Ti/Al ratio. It is found that the root mean square (R_{MS}) roughness and the grain size depend on the Al amount in the contact layer. Higher Al percentage in the former-Ti/Al layer causes rising the roughness. R_{MS} surface roughness of 17.3 nm and 15.9 nm is determined for Ti/Al (30/70 wt.%) and Ti/Al (50/50 wt.%) contacts, respectively, after annealing at 800 ^{0}C. Lowering the Al content affects on decrease of the grain size from 180 nm to 140 nm as well. Further increase of the Ti/Al ratio leads to a lower roughness of the surface and a smaller grain size of the contact system, even after annealing at temperatures as high as 850 ^{0}C. R_{MS} of 12.8 nm and grain size in the interval 110-130 nm are measured with Ti/Al (70/30 wt.%) contacts. The results obtained from AFM examination of contacts with a varying Ti/Al ratio in the former layer have shown that decrease of the Al content improves the surface morphology. The same effect of the Al content has been observed in ohmic contacts to SiC.
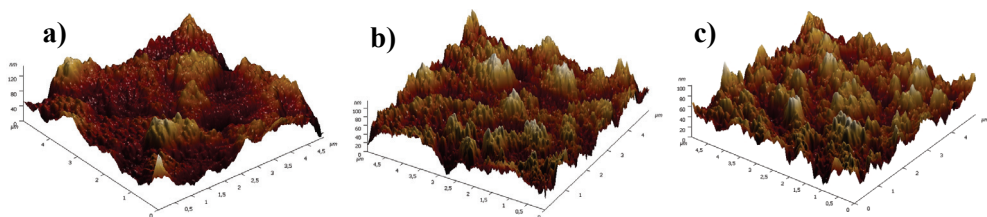
Fig. 17. AFM 3D image of (5x5) μm^2 surface area of a Ti/Al/Ti/Au contacts annealed at optimal temperature with a Ti/Al ratio of: (a) - (30/70) wt.%, (b) - (50/50) wt.% and (c) - (70/30) wt.%.

The different initial Ti/Al ratio and the resulting different annealing temperatures lead to remarkable differences in element distribution and interface chemistry of both ohmic contacts as well. The element depth distributions for the Ti/Al (50/50 wt.%) contact after annealing at 800 ^{0}C and Ti/Al (70/30 wt.%) contact after annealing at 850 ^{0}C are presented in Fig. 18. The profiles reveal intermixing of Al, Ti, and Au layers. In both contacts, strong Al diffusion to the surface induced by the thermal treatment is observed. The surface region of the Ti/Al (50/50 wt.%) contact consists mainly of Al and Au. Going into the depth a gradual decrease in Al and increase in Au concentrations is detected. The binding energy of Au4f7/2 at 84.6 eV is close to that obtained for AlAu$_2$ alloy. A significant amount of N and smaller amounts of Ga and Ti are found in the region below the gold layers. This is clearly a result of N and Ga outward diffusion towards the surface. Since the measured binding energies of N1s and Ti2p peaks (396.8 eV and 454.8 eV, respectively) correspond to that obtained for TiN, it might be suggested that the diffused N reacts with Ti to form TiN. The depth profile also reveals that during the annealing Al diffuses through the Ti and GaN layers to the interface with AlGaN. The binding energy of Al2s peak here is 119.0 eV, which corresponds to Al in the metal state. At the interface with the AlGaN layer the Al2s peak is broadened and exhibits second maximum at 122.0 eV, which is characteristic of AlGaN. In the surface layers of the Ti/Al (70/30 wt.%) contact predominantly Al in the form of Al$_2$O$_3$ is detected (Fig. 18b). Its concentration sharply decreases going into the depth of the layers. This is followed by a strong increase of the gold concentration, which suggests that the thicker Ti layer is more effective barrier against gold diffusion to the interface. The binding energy value of the Au4f7/2 peak near to the region rich in Al is 84.6 eV but decreases to 84.1 eV, into the depth of the contact. The higher annealing temperature results in enhanced outward diffusion of N and Ga toward the surface. The diffused nitrogen reacts with Ti and forms TiN that is evidenced by the measured binding energies of N1s and Ti2p peaks. The most significant difference as compared to Ti/Al (50/50 wt.%) contact is the higher concentration of Ga in this region (20% vs. 10 %), which is probably due to the higher diffusion rate of gallium at 850 ^{0}C.
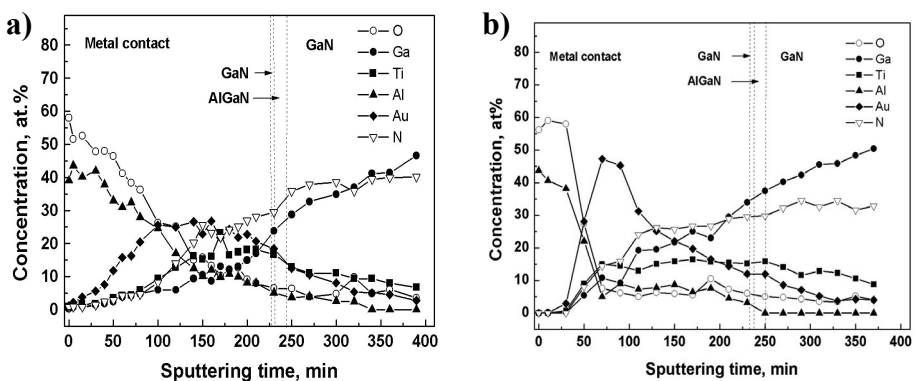
Fig. 18. XPS depth profiles of Ti/Al/Ti/Au contacts annealed at optimal temperature with a Ti/Al ratio of: (a) – (50/50) wt. % and (b) – (70/30) wt. %.

The AFM analysis shows improvement of the surface morphology and narrowing the contact periphery with a decrease of the Al amount in the former-Ti/Al layer. The lowest R_{MS} = 12.8 nm of the surface has been achieved for the Ti/Al (70/30 wt.%) contact after annealing at 850 ^{0}C. However, the higher annealing temperature enhanced the interdiffusion of the components and the tendency to oxidation of Ti and Al. As a result this contact composition exhibits the worst contact resistivity. Consequently, a compromise regarding the choice of the appropriate composition for ohmic contact to GaN/GaAlN HEMT structures should be made.

5. Summary

The study of ohmic contacts to wide band-gap semiconductors proves that when metal/semiconductor contacts are deposited, they commonly result in rectifying Schottky contacts which barrier height inhibits current flow across the metal/semiconductor interface. There are four primary variables which control the Schottky barrier height at metal/semiconductor interfaces: the work function ϕ_m of the metal; the crystalline or amorphous structure at the metal-semiconductor interface; the diffusion of metal atoms across the interface into the semiconductor; and, the outermost electronic configuration of the metal atoms. Otherwise, there are several constants and properties characterising the wide band-gap semiconductors which postulate the specific approach used for formation of ohmic properties of the metal/semiconductor interface: the high electron affinity, the wide forbidden zone, and low diffusion coefficient of the most metals. Consequently, it is almost impossible to form ohmic properties, relying only to the choice of a metal with suitable work function and metal diffusion into the semiconductor during annealing. Therefore in the case of ohmic contacts to wide band-gap semiconductors metallization schemes have been chosen so as to form intermediate layer at the interface, which could decrease the barrier height and/or narrow the depletion layer at the semiconductor interface. In these cases, heat treatment results interfacial compounds, such as metal/compound/ semiconductor contacts. In these contacts, the metal/semiconductor interface is eliminated and replaced by new interfaces, a metal/compound and a compound/semiconductor interface. The resulting barrier height ϕ_B is not longer dependent on the surface properties of the semiconductor or metal work function. Instead, it depends upon the difference in electron affinity and work function between the metal/compound and compound/ semiconductor. As a result, contacts can be reproducibly formed with a predictable ϕ_B. In the case of Ni-based and Pd-based contacts to SiC such compound is nickel silicide and palladium silicide, respectively.

On the basis of XPS data the following mechanism of chemical reactions occurring during the formation of ohmic properties may be proposed. In the case of Ni/SiC the contact formation is initiated by the dissociation of SiC surface, due to the strong reactivity of Ni at 950 ^{0}C. The nickel atoms at the interface interact with a part of dissociated Si atoms and Ni_2Si is formed. Simultaneously, at the interface nickel atoms diffuse through the mixed Ni_2Si+C layer towards the SiC. Thus, the supply of Ni atoms at the SiC interface continues and the above reactions are repeated to the complete consumption of the deposited nickel layer. Carbon accumulates, both at the interface and in the contact layer. The presence of carbon in the contact layer and at the interface could become a potential source of contact

degradation at very high temperatures. When Ni/Si multilayers (instead of pure Ni) are deposited on SiC, the contact formation is preceded by Ni and Si mutual diffusion in the deposited layer yielding Ni_2Si. The presence of Ni atoms at the interface is a reason for dissociation of SiC to Si and C, after which Ni atoms are bonded to the free Si atoms and form Ni_2Si along with carbon in the graphite state. A smaller amount of carbon is observed at the interface. Low carbon segregation at the interface and an abrupt interface characterise this contact. The mechanism of Ni-based ohmic contact formation is illustrated in Fig. 19. The calculations are made on the base of the measured forward I-V characteristic for the as-deposited contact and the thermionic-field emission transport mechanism in the annealed contacts at doping concentration of 1×10^{19} cm^{-3}, T=298 K and an effective electron mass $m^*_n = 0.206 m_0$ (Kassamakova-Kolaklieva, 1999).
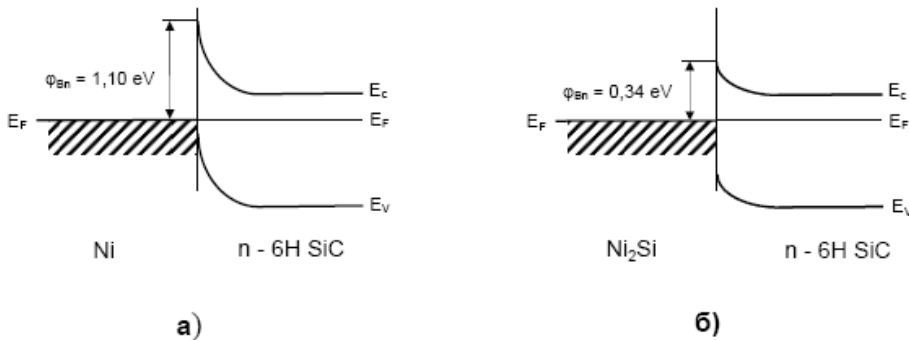
Fig. 19. Energy band diagram of unannealed (a) and annealed at 950 ^{0}C (b) Ni/n-type 4H-SiC interface.

Annealing of the interface in Pd-based contacts also causes partial dissociation of SiC to Si and C. As a result of this process, the SiC interface is shifted into the SiC bulk since a part of the original interface is consumed. The free Si atoms interact with Pd to form Pd_2Si in the interface near region and Pd_3Si in the more remote contact layer. The formation of these compounds at the interface and in the contact layer, respectively, has been observed for all Pd-based contacts. Consequently, the presence of Pd_2Si at the interface leads to reduction of the barrier height and appearance of ohmic properties, i.e. again lowering the barrier height is realised by silicide formation at the interface (Fig. 20) (Kassamakova-Kolaklieva, 1999).
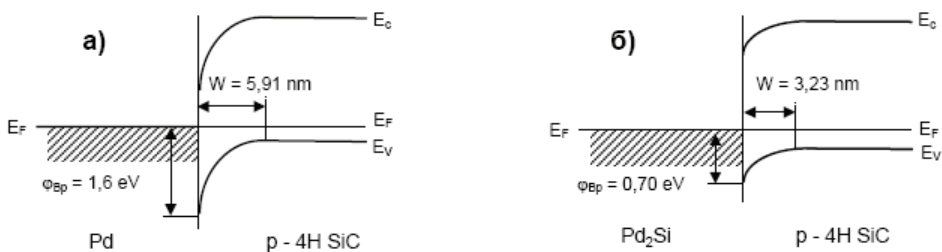
Fig. 20. Energy band diagram of unannealed (a) and annealed (b) Pd/p-type 4H-SiC interface.

The origin of ohmic properties of Al-based ohmic contacts to 4H-SiC depends strongly on the contact composition and annealing temperature. There is no the same mechanism for ohmic properties formation. The low annealing temperature of the Al/Si/SiC contacts decreases the interdiffusion/chemical reaction processes because the dissociation of SiC surface is poor at 700 ^{0}C. In addition, the Si layer, deposited on the substrate surface, acts as a barrier for aluminium diffusion. As a result, Al in metal state only is established in the XPS spectra of the Al/Si contacts annealed at this temperature. After ageing of Al/Si contacts at 600 ^{0}C for 48 hours areas without a metal film on the contact pads could be seen, suggesting that a part of undiffused Al from the annealed contact layer evaporates during the long term heating, resulting in temperature instability. The increase of the annealing temperature in the AlSiTi contact stimulates a higher interdiffusion/chemical reaction of Al with SiC. Due to the catalytic effect of Al at elevated temperatures SiC dissociation occurs at the metal/SiC interface. The undiffused Al atoms of the contact layer react entirely with the carbon forming a stable compound, Al_4C_3. Indeed, the presence of chemical stable Al_4C_3 compound and the absence of Al in metal state are prerequisite for the improved thermal stability of AlSiTi contacts at high ageing temperatures (Kassamakova et al., 2001). In the case of Au/Ti/Al contacts strong dependence of the contact structure on the Ti:Al ratio and annealing temperature, respectively, has been found out. The TEM analysis reveals that titanium and aluminium silicides and carbides are formed after annealing at 900 ^{0}C irrespective of the Ti:Al ratio. However, the Ti:Al ratio affect the kind of silicides and carbides created. In the contact with a Ti:Al ratio of 70:30 Ti_3SiC_2 and TiSi are formed. Although Ti is not in the contact with SiC in the as-deposited structure, it could diffuse through the melted aluminium very fast and reacts with SiC, which is resolved at presence of the molten Al. As a result, the rich on carbon Ti_3SiC_2 phase is formed. The excess Si reacts with Ti to form TiSi and Ti_5Si_3 depending on the Ti amount in the initial contact film. Higher Al content in the initial contact, lower Ti:Al ratio respectively, hinders the formation of ternary Ti_3SiC_2 compound and favours the reactions leading to the formation of binary compounds. Obviously, the higher Al amount makes it more reactive to the carbon than Ti and AlC_4 is detected. In the case of the Au/Ti(70%)/Al(30%) contact the origin of ohmic properties is the formation of ternary Ti_3SiC_2 compound at the interface, which is known to exhibit advantageous metallic properties. However, this compound is not detected in the annealed Au/Ti(30%)/Al(70%) contacts. XPS analysis of this contact has revealed a slight diffusion of Al into the SiC surface after annealing at 1000 ^{0}C. It could be supposed, in analogy with the Ti-Al alloyed contacts with the same Al percentage content and annealed at the same temperature (Crofton et al., 1993) that in the annealed Ti/Al layered contacts Al is also distributed like spikes near the SiC surface. Resistivity improvement of the Au/Ti(30%)/Al(70%) contacts after annealing at 1000 ^{0}C is due to the Al spikes into SiC. Hence, the origin of the ohmic properties improvement could be explained by the formation of Ti_3SiC_2 compound and enhanced carrier transport by the presence of metal spikes into SiC depending on the initial contact composition and as consequence the optimal annealing temperature (Kolaklieva et al., 2007).

In the case of Ti/Al-based contacts the first Ti layer being in intimate contact with the GaN (or AlGaN) interface takes essential role in ohmic properties formation during annealing. The formation of Ti_xN at the interface is considered important for ohmic behaviour

obtaining. TixN can be grown at the interface between the multilayered metallization by interfacial reactions at temperatures ranging from 250°C (furnace anneal) to 900°C (rapid thermal anneal). The presence of TiN at the interface, with a theoretically predictable work function of 3.74 eV and reasonable electrical conductivities, decreases the barrier height and ohmic properties have been obtained. The formation of TiN at the interface metal /GaN creates nitrogen vacancies in the GaN substrate. These vacancies act as shallow donors, which enhance the doping level at the interface and decrease the width of the depletion layer resulting in decrease of the contact resistivity.

It should be pointed out that besides the interfacial compound, additional alloys and compounds are formed in the contact layer during annealing, which presence aids the better contact conductivity. Obviously, their composition determines by the contact composition before annealing, semiconductor composition and the annealing temperature. Nevertheless, the interfacial reactions are critical to the formation of ohmic contacts on semiconductors, whether they have a large or a small band-gap.

6. References

Berger, B., (1972), Models for contacts to planar devices. *Solid-State Electronics,* Vol. 15, No. 2, 145-148, ISSN: 0038-1101.

Crofton, J., Barnes, P., Williams, J. & A. Edmond, J., (1993), Contact resistance measurements on p-type 6H–SiC. *Applied Physics Lett*ers, Vol. 62, No. 4, p. 384-386, ISSN: 00036951.

Crofton, J., Porter, L. & Williams, J., (1997), The Physics of Ohmic Contacts to SiC, *Physica Status Solidi (b)*, Vol. 202, No. 1, 581-603, ISSN: 0370-1972.

Grunthaner, P., Grunthaner, F. & Mayer, J., (1980), XPS study of the chemical structure of the nickel/silicon interface. *J. Vacuum Science & Technology,* Vol. 17, No. 5, 925-929.

Kakanakov, R., Kasamakova – Kolaklieva, L., Hristeva, N., Lepoeva, G., Gomes, J., Avramova, I. & Marinova, Ts., (2004), High temperature and high power stability investigation of Al-based ohmic contacts to p-type 4H-SiC. *Materials Science Forum,* Vols. 457-460, 877-880, ISSN 0255-5476.

Kakanakov, R., Kasamakova, L., Kasamakov, I., Zekentes, K. & Kuznetsov, N., (2001), Improved Al/Si ohmic contacts to p-type 4H-SiC. *Materials Science and Engineering*, Vol. B80, No. 1-3, 374-377, ISSN: 1862-6300.

Kassamakova, L., Kakanakov, R., Kassamakov, I., Zekentes, K., Tsagaraki, K., & Atanasova, G., (2001), Origin of the excellent thermal stability of Al/Si-based ohmic contacts to p-type LPE 4H-SiC. *Materials Science Forum,* Vols. 353-356, 251-254, ISSN 0255-5476.

Kassamakova, L., Kakanakov, R., Kassamakov, I., Nordell, N., Savage, S., Hjörvarssön, B., Svedberg, E., Åbom, L., & Madsen, L., (1999), Temperature stable Pd ohmic contacts to p-type 4H-SiC formed at low temperatures, *IEEE Trans. on Electr. Dev.,* Vol. 46, 605-611, ISSN: 0018-9383.

Kassamakova, L., Kakanakov, R., Nordell, N., Savage, S., Kakanakova-Georgieva, A., & Marinova, Ts., (1999), Study of the electrical, thermal and chemical properties of Pd ohmic contacts to p-type 4H-SiC depending on annealing conditions. *Materials Science and Engineering*, Vol. B56, 291-295, ISSN: 1862-6300.

Kassamakova-Kolaklieva, L., (1999), Development and investigation of temperature stable ohmic and Schottky contacts for high power devices, *Ph.D. Thesis*.

Kassamakova-Kolaklieva, L., Kakanakov, R., Hristeva, N., Lepoeva, G., Cimalla, V., Kuznetsov, N. & Zekentes, K., (2003), Pd-based ohmic contacts to LPE 4H-SiC with improved thermal stability. *Materials Science Forum,* Vols. 433-436, 713-716, ISSN 0255-5476.

Kolaklieva, L., Kakanakov, R., Avramova, I., & Marinova, Ts., (2007), Nanolayered Au/Ti/Al Ohmic Contacts To P-Type SiC: Electrical, Morphological And Chemical Performances Depending On The Contact Composition. *Materials Science Forum*, Vols. 556-557, 725-728, ISSN 0255-5476.

Kolaklieva, L., Kakanakov, R., Cimalla, V., Maroldt, St., Niebelschütz, F., Tonisch, K. & Ambacher, O., (2008), The Role of Ti/Al Ratio in Nanolayered Ohmic Contacts for GaN/AlGaN HEMTs. *Proc. of 26th International Conference on Microelectronics*, *Nič, Serbia, 11-14 May 2008,* 221-224, Electron Devices Society of the Institute of Electrical and Electronics Engineers, Inc., ISBN: 978-1-4244-1882-4.

Kolaklieva, L., Kakanakov, R., Lepoeva, B. Gomes, J. & Marinova, Ts., (2004), Au/Ti/Al contacts to SiC for power applications: electrical, chemical and thermal properties. *Proc. of 24rd International Conference on Microelectronics*, *Nič, Serbia and Montenegro, 16-19 May 2004*, Vol. 2, 421-424, Electron Devices Society of the Institute of Electrical and Electronics Engineers, Inc., ISBN: 0-7803-8166-1.

Kolaklieva, L., Kakanakov, R., Marinova, Ts. & Lepoeva, G., (2005), Effect of the metal composition on the electrical and thermal properties of Au/Pd/Ti/Pd contacts to p-type SiC. *Materials Science Forum*, Vols. 483-485, 749-752, ISSN 0255-5476.

Kolaklieva, L., Kakanakov, R., Stefanov, P., Cimalla, V., Ambacher, O., Tonisch, K., Niebelschütz, M. & Niebelschütz, F., (2009), Composition and Interface Chemistry Dependence in Ohmic Contacts to GaN HEMT Structures on the Ti/Al Ratio and Annealing Conditions. *Materials Science Forum,* Vols. 615-617, 951-954, ISSN 0255-5476.

Marinova, Ts., Kakanakova-Georgieva, A., Krastev, V., Kakanakov, R., Neshev, M., Kassamakova, L., Noblanc, O., Arnodo, C., Cassette, S., Brylinski, C., Pecz, B., Radnocy, G. & Vinze, Gy., (1997) Nickel based ohmic contacts on SiC. *Materials Science and Engineering,* Vol. B46, No. 1-3, 223-226, ISSN: 1862-6300.

Marlow, G., & Das, M., (1982), *Solid-State Electronics,* Vol. 25, No. 2, 91-94, ISSN: 0038-1101.

Meyer, M. & Metzger, R., (1996), Flying high: the commercial satellite industry converts to compound semiconductor solar sells. *Compound Semiconductor,* Vol. 2, No. 6, 22-24, ISSN: 1096-598X.

Padovani, F. & Stratton, R., (1966), Field and thermionic-field emission in Schottky barriers, *Solid-State Electronics,* Vol. 9, No. 7, 695-707, ISSN: 0038-1101.

Porter, L. & Davis, R., (1995), A critical review of ohmic and rectifying contacts for silicon carbide. *Materials Science and Engineering, Vol.* B34, No. 2-3, 83-105, ISSN: 1862-6300.

Sze, S., (1981), *Physics of Semiconductor Devices,* John Wiley & Sons, Inc., ISBN: 0-471-05661-8, USA.

Waldrop, J. & Grant, R., (1993), Schottky barrier height and interference chemistry of annealed metal contacts to alpha 6H-SiC: Crystal face dependence. *Applied Physics Letters,* Vol. 62, No. 21, 2685-2687, ISSN: 00036951.

Yu, A., (1970), Electron tunnelling and contact resistance of metal-silicon contact barriers. *Solid-State Electronics,* Vol. 13, No. 2, 239-247, ISSN: 0038-1101.

Implications of Negative Bias Temperature Instability in Power MOS Transistors

Danijel Danković, Ivica Manić, Snežana Djorić-Veljković,
Vojkan Davidović, Snežana Golubović and Ninoslav Stojadinović
University of Niš
Serbia

1. Introduction

As the device dimensions in metal-oxide-silicon (MOS) technologies have been continuously scaled down, a phenomenon called negative bias temperature instability (NBTI), which refers to the generation of positive oxide charge and interface traps in MOS structures under negative gate bias at elevated temperature, has been gaining in importance as one of the most critical mechanisms of MOS field effect transistor (MOSFET) degradation. NBTI effects are manifested as the changes in device threshold voltage (V_T), transconductance (g_m) and drain current (I_D), and have been observed mostly in p-channel MOSFETs operated under negative gate oxide fields in the range 2 - 6 MV/cm at temperatures around 100°C or higher (Huard et al., 2006; Stathis & Zafar, 2006; Schroder, 2005; Alam & Mahapatra, 2005; Schroder & Babcock, 2003; Kimizuka et al., 1999; Ogawa et al., 1995). The phenomenon itself had been known for many years, but only recently has been recognised as a serious reliability issue in state-of-the-art MOS integrated circuits. Several factors associated with device scaling have been found to enhance NBTI: *i*) operating voltages have not been reduced as aggressively as gate oxide thickness, leading to higher oxide electric fields and increased chip temperatures; *ii*) threshold voltage scaling has not kept pace with operating voltage, resulting in larger degradation of drain current for the same shift in threshold voltage; and *iii*) addition of nitrogen during the oxidation process has helped to reduce the thin gate oxide leakage, but the side effect was to increase NBTI (Stathis & Zafar, 2006).

Considering the effects of NBTI related degradation on device electrical parameters, NBT stress-induced threshold voltage shift (ΔV_T) seems to be the most critical one, and a couple of basic questions, which are to be addressed now, are why the NBTI appears to be of great concern only in p-channel devices, and why the negative bias causes more considerable degradation than positive bias. The bias temperature stress-induced V_T shifts are generally known to be the consequence of underlying buildup of interface traps and oxide-trapped charge due to stress-initiated electrochemical processes involving oxide and interface defects, holes and/or electrons, and variety of species associated with presence of hydrogen as the most common impurity in MOS devices (see e.g. (Schroder & Babcock, 2003)). An interface trap is an interfacial trivalent silicon atom with an unsaturated (unpaired) valence electron at the SiO_2/Si interface. Unsaturated Si atoms are additionally found in SiO_2 itself, along with other oxide defects, the most important being the oxygen vacancies. Both oxygen

vacancies and unsaturated Si atoms in the oxide are concentrated mostly near the interface and they both act as the trapping centers responsible for buildup of oxide-trapped charge. Interface traps readily exchange charge, either electrons or holes, with the substrate and they introduce either positive or negative net charge at interface, which depends on gate bias: the net charge in interface traps is negative in n-channel devices, which are normally biased with positive gate voltage, but is positive in p-channel devices as they require negative gate bias to be turned on. On the other hand, charge found trapped in the centers in the oxide is generally positive in both n- and p-channel MOS transistors and cannot be quickly removed by altering the gate bias polarity. The absolute values of threshold voltage shifts due to stress-induced oxide-trapped charge and interface traps in n- and p-channel MOS transistors, respectively, can be expressed as (Ma & Dressendorfer, 1989):

$$\Delta V_{Tn} = \frac{q\Delta N_{ot}}{C_{ox}} - \frac{q\Delta N_{it}}{C_{ox}},\tag{1}$$

$$\Delta V_{Tp} = \frac{q\Delta N_{ot}}{C_{ox}} + \frac{q\Delta N_{it}}{C_{ox}},\tag{2}$$

where q denotes elementary charge, C_{ox} is gate oxide capacitance per unit area, while ΔN_{ot} and ΔN_{it} are stress-induced changes in the area densities of oxide-trapped charge and interface traps, respectively. The amounts of NBT stress-induced oxide-trapped charge and interface traps in n- and p-channel devices are generally similar (Stathis & Zafar, 2006), but above consideration clearly shows that the net effect on threshold voltage, ΔV_T, must be greater for p-channel devices, because in this case the positive oxide charge and positive interface charge are additive. As for the question on the role of stress bias polarity, it seems well established that holes are necessary to initiate and/or enhance the bias temperature stress degradation (Huard et al., 2006; Stathis & Zafar, 2006; Schroder, 2005; Alam & Mahapatra, 2005; Schroder & Babcock, 2003; Kimizuka et al., 1999; Ogawa et al., 1995), which provides straight answer since only negative gate bias can provide holes at the SiO$_2$/Si interface. Moreover, this is an additional reason why the greatest impact of NBTI occurs in p-channel transistors since only those devices experience a uniform negative gate bias condition during typical CMOS circuit operation.

Several models of microscopic mechanisms responsible for the observed degradation have been proposed (Huard et al., 2006; Stathis & Zafar, 2006; Schroder, 2005; Alam & Mahapatra, 2005; Schroder & Babcock, 2003; Ogawa et al., 1995), but in spite of very extensive studies in recent years, the mechanisms of NBTI phenomenon are still not fully understood, so technology optimization to minimize NBTI is still far from being achieved. With reduction in gate oxide thickness, NBT stress-induced threshold voltage shifts are getting more critical and can put serious limit to a lifetime of p-channel devices having gate oxide thinner than 3.5 nm (Kimizuka et al., 1999), so accurate models and well established procedure for lifetime estimation are needed to make good prediction of device reliable operation.

Though the gate oxide in nanometre scale technologies is continuously being thinned down, there is still high interest in ultra-thick oxides owing to widespread use of MOS technologies for the realisation of power devices. Vertical double-diffused MOSFET (VDMOSFET) is an attractive device for application in high-frequency switching power supplies owing to its superior switching characteristics which enable operation in a megahertz frequency range

(Baliga, 1987; Benda et al., 1999). High-frequency operation allows the use of small-size passive components (transformers, coils, capacitors) and thus enables the reduction of overall weight and volume, making the power VDMOSFETs especially suited for application in power supply units for communication satellites, but they are also widely used as the fast switching devices in home appliances and automotive, industrial and military electronics. Degradation of power MOSFETs under various stresses (irradiation, high field, and hot carriers) has been subject of extensive research (see e.g. (Stojadinović et al., 2006) and references cited therein), but very few authors seem to have addressed the NBTI in these devices (Demesmaeker et al., 1997; Gamerith & Polzl, 2002; Stojadinović et al., 2005; Danković et al., 2006; Danković et al., 2007; Danković et al., 2008, Manić et al., 2009). However, power devices are routinely operated at high current and voltage levels, which lead to both self heating and increased gate oxide fields, and thus favour NBTI. Accordingly, NBTI could be critical for normal operation of power MOSFETs though they have very thick gate oxides.

Given the above considerations, this chapter is to cover the NBTI implications on reliability of commercially available power VDMOSFETs. In the next section, we will describe the experimental procedure for accelerated NBT stressing applied in our study and analyse typical results for the threshold voltage shifts observed in stressed devices. Applicability of some empirical expressions for fitting the dependences of stress-induced threshold voltage shifts on stress conditions (voltage, temperature, time) to our experimental data will be discussed as well. Third section is to describe in details the results of the procedure applied to fit the experimental data and estimate the device lifetime by means of several fitting and extrapolation models. Impacts of stress conditions, failure criteria, models used for fitting and extrapolation, and intermittent annealing on lifetime projection will be discussed as well. The extrapolation models available in the literature offer only extrapolation along the voltage (or electric field) axis and provide lifetime estimates only for the temperatures applied during the accelerated stressing, so in the next section we propose a new approach, which requires double extrapolation along both voltage and temperature axes, but can estimate the device lifetime for any reasonable combination of operating voltages and temperatures, including those falling within the ranges normally found in usual device applications. Finally, most important findings presented in the chapter will be summarized in the conclusion section.

2. NBT stress-induced threshold voltage shifts

Devices used in our study were commercial p-channel power VDMOSFETs IRF9520, encapsulated in TO-220 plastic cases, with current and voltage ratings of 6.8 A and 100 V, respectively. Devices were built in standard silicon-gate technology with 100 nm thick gate oxide, and had the threshold voltage V_T = -3 V. Several sets of devices have been stressed up to 2000 hours by applying negative voltages in the range 30 – 45 V to the gate, with drain and source terminals grounded, at temperatures ranging from 125 to 175°C. A conventional methodology, based on periodic breaks during the stress to measure the device transfer I-V characteristics, was applied to characterize the NBT stress effects. Threshold voltage values were estimated from the above-threshold transfer characteristics as the intersections of extrapolated linear region of $\sqrt{I_D} - V_{GS}$ curves with V_{GS} - axis.

Typical transfer *I-V* characteristics of p-channel power VDMOSFETs measured during the NBT stressing are shown in Fig. 1. It can be seen that, as the stressing progresses, the characteristics are being shifted along the V_{GS} axis towards the higher voltage values, which is the consequence of stress-induced buildup of oxide-trapped charge. The shifts are more significant in the early phase of stressing and gradually become smaller with tendency to saturate in the advanced stress phase. At the same time, the slope of the curves slightly decreases, indicating that interface traps are being generated as well.
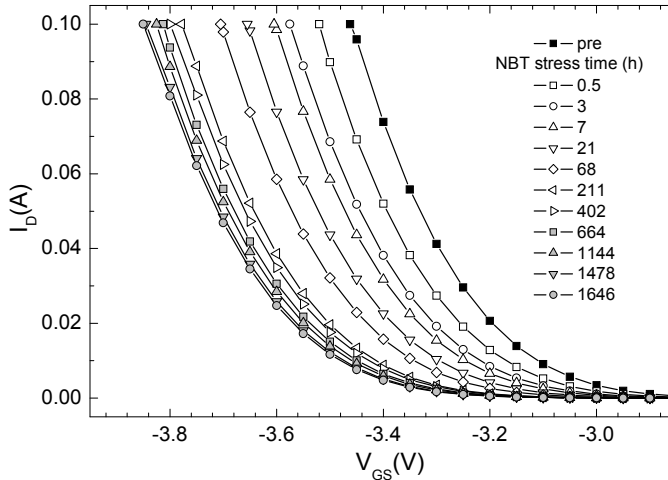
Fig. 1. I_D-V_{GS} characteristics of p-channel power VDMOSFETs during NBT stressing with V_G = - 40 V at 150°C.

In line with observed shift of transfer characteristics along the voltage axis, NBT stressing was found to cause significant threshold voltage shifts in our devices. Two characteristic sets of data, for IRF9520 devices stressed with - 40 V at different temperatures and with different voltages at 150°C, are shown in Figs. 2 and 3, respectively. Apparently, more pronounced shifts are observed in devices stressed at higher temperatures (Fig. 2) and/or with higher stress voltages (Fig. 3). In all cases, ΔV_T time dependences have been found to follow the t^n power law, with three distinct phases (as indicated by the dashed lines), which can be clearly distinguished depending on the value of parameter *n* (Stojadinović et al., 2005; Danković et al., 2006; Danković et al., 2006a). In the first (early) stress phase, *n* strongly depends on both stress bias and temperature, varying from 1.14 to 0.4. In the second phase, *n* is almost independent on bias and temperature, and ΔV_T follows the well-known $t^{0.25}$ law (Jeppson & Svensson, 1977; Ogawa et al., 1995; Schroder, 2005; Huard et al., 2006; Stathis & Zafar, 2006). The second phase begins earlier in devices stressed with higher voltages and/or at higher temperatures so the first phase might even disappear if more severe stress conditions had been applied. Finally, in the third phase, *n* becomes bias and temperature dependent again and gradually decreases from 0.25 to 0.14, whereas ΔV_T tends to saturate. The ΔV_T in saturation after near 2000 hours of stressing was found to vary from about 4.4 % in devices stressed at 125°C with - 30 V) up to 19.8 % in those stressed at 175°C with - 45 V (Stojadinović et al., 2005).
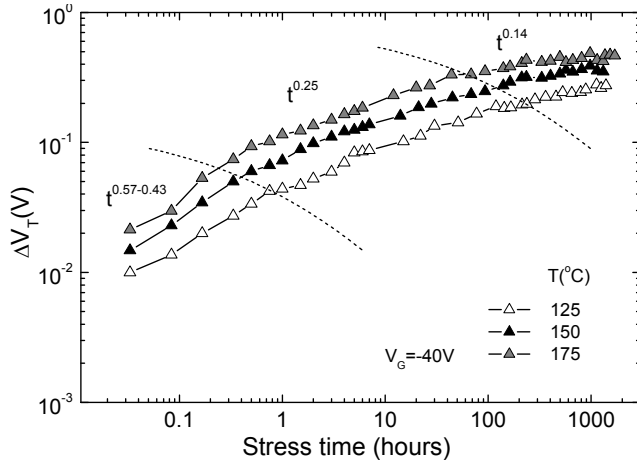
Fig. 2. Threshold voltage shifts in p-channel power VDMOSFETs during the NBT stressing with V_G = - 40 V at different temperatures.
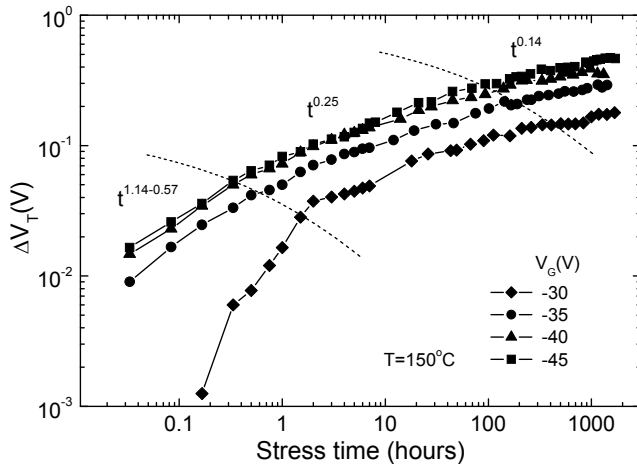
Fig. 3. Threshold voltage shifts in p-channel power VDMOSFETs during the NBT stressing with different gate voltages at 150°C.

In modelling the NBT stress-induced threshold voltage shifts, the effects of stress time and temperature are commonly accounted for by the power law time dependence and Arrhenius temperature acceleration terms, respectively, whereas the dependence on gate oxide electric field (or gate voltage applied) can be modelled by incorporating either the exponential, modified exponential, or power law terms. Accordingly, the experimental data shown in Figs. 2 and 3 could be modelled by any of the three empirical expressions given as follows (Ogawa et al., 1995; Krishnan et al., 2001; Ershov et al., 2005):

$$\Delta V_T = C_1 e^{\beta_1 V_G} t^n \exp(-E_a / kT),$$ (3)

$$\Delta V_T = C_2 e^{-\beta_2/V_G} t^n \exp(-E_a/kT), \tag{4}$$

$$\Delta V_T = C_3 E^m t^n \exp(-E_a/kT), \tag{5}$$

where V_G, E, t, and T denote stress voltage, corresponding oxide electric field, time, and temperature, respectively, E_a is activation energy, k is the Boltzmann constant, whereas C_1, C_2, C_3, β_1, β_2, m, and n are the fitting parameters. Activation energy and fitting parameters assume different values for each of three stress phases mentioned above, and these expressions can be used to calculate the NBT stress-induced threshold voltage shift in IRF9520 devices for a random combination of stress voltages and temperatures within the voltage and temperature ranges investigated (Stojadinović et al., 2005; Stojadinović et al., 2007).

3. NBTI effects on device lifetime

As already mentioned, degradation associated with NBTI could put serious limit to a lifetime of devices operated at elevated temperatures under the increased gate oxide field. Our goal is to estimate the lifetime of investigated power VDMOSFETs under normal operating conditions by using the results obtained from accelerated NBT stressing. However, the standard procedure requires first to extract the values of lifetime the devices would have if operated under the experimental conditions, and these experimental lifetime values are then used for extrapolation to normal operating conditions. Accordingly, this section will begin with estimation of the lifetime under experimental conditions, which will be followed by a procedure for extrapolation to normal operating bias conditions, including detailed discussion of various factors affecting the lifetime estimated in this way. The effects of intermittent annealing on lifetime projection will be additionally discussed.

3.1 Extraction of experimental lifetime data

To estimate the device lifetime it is necessary to choose one of device parameters affected by the stress, which will be used to monitor the level of stress-induced degradation, as well as to define the failure criterion (FC) as a maximum allowed change of the chosen parameter, which could be critical for device and/or circuit reliable operation. Various parameters, such as threshold voltage, transconductance, or drain current, can be used as a degradation monitor (Schlunder et al., 2005; Tan et al., 2005). We will use the threshold voltage, which has been shown in previous section to be affected by NBT stressing, and also has been widely accepted as a well-suited parameter, so the device lifetime for practical operation of power VDMOSFETs studied here will be estimated from the experimental results for the NBT stress-induced threshold voltage shifts. As shown in Fig. 4, experimental lifetime is obtained as the stress time required that ΔV_T reaches the predetermined value of FC (100 mV in this case). Experimental lifetime values, t_i ($i = 1 - 4$), are extracted from the plots shown in the figure and are used to plot the lifetime vs. gate voltage dependence as in the inset graph, which is further used for extrapolation to a normal operating voltage.

A proper choice of failure criterion, which will be discussed in more details later, is very important. For example, it can be seen in Fig. 4 that, if the value of FC was higher then about 120 mV, we could not get experimental lifetime for devices stressed with - 30 V at 125°C as in this case ΔV_T did not reach 120 mV even after 2000 hours of stressing. Thus, duration of
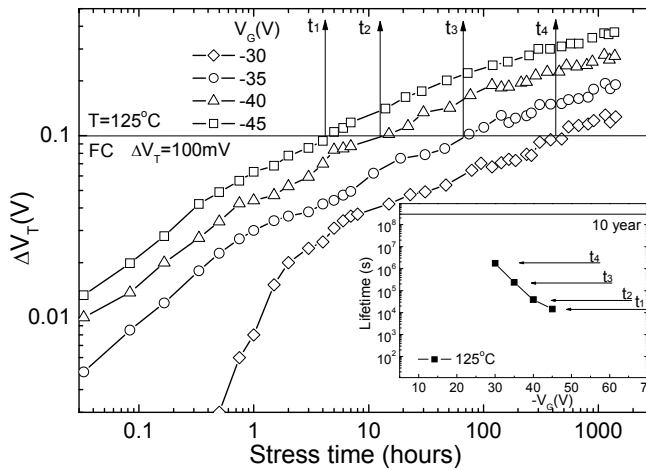
Fig. 4. Extraction of experimental lifetime data and illustration of extrapolation to a normal operating voltage (inset).

the experiment in this case was not sufficient to achieve ΔV_T as high as FC, which means the stress time should have been extended to exceed the device lifetime for a given combination of stress bias and temperature. However, this could require the stressing to be done beyond the reasonable time limit so in such cases it is convenient to use an adequate fitting model capable to provide reliable prediction of ΔV_T on the basis of available data. The simplest way is to use regular exponential model (Liu et al., 2001; Liu et al., 2002; Krishnan & Kol'dyaev, 2002):

$$\Delta V_T(t) = \Delta V_{T\max}[1 - \exp(-t/\tau)], \qquad (6)$$

where ΔV_{Tmax} is the saturation level of threshold voltage shift and τ is the characteristic time constant. However, as shown in Fig. 5, this model (dotted line) in our case yields very poor agreement with experimental data and significantly underestimates ΔV_T in saturation. Since the ΔV_T time dependences in our devices exhibit the discontinuities manifested as three distinct stress phases (Figs. 2 and 3), better agreement with experimental data is expected from the modified version of the above model, which is known as the 2-tau exponential model and is given by (Liu et al., 2001; Liu et al., 2002):

$$\Delta V_T(t) = \Delta V_1[1 - \exp(-t/\tau_1)] + \Delta V_2[1 - \exp(-t/\tau_2)], \qquad (7)$$

where the fitting parameters τ_1 and τ_2 represent the time constants closely related to transitions from early to the second stress phase and from the second phase to saturation, respectively, whereas ΔV_1 and ΔV_2 are associated with corresponding V_T shifts. This model is suitable for processes which have two distinct mechanisms operating one after the other with $100\tau_1 < \tau_2$. As can be seen in Fig. 5, the 2-tau model (dashed line) yields fairly good agreement with experimental data. However, this model appears very sensitive to small fluctuations in experimental data (the fit looks "wavy" in the second stress phase) and tends to exaggerate saturation in the case of devices stressed with - 30 V at 125°C, which does not seem justified by the experimental data.
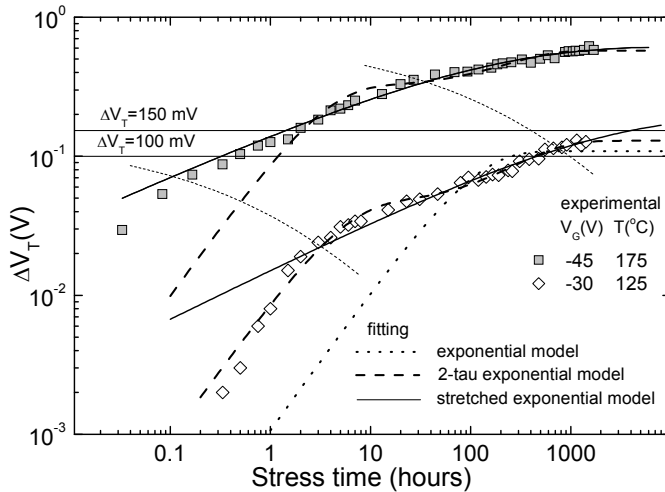
Fig. 5. Fitting of the NBT stress-induced V_T shifts by means of various models. Symbols denote the measurement data and lines are the fits.

Trying to resolve the problem, we have also considered the so-called stretched exponential model, given by (Van de Walle, 1996; Zafar et al., 2003; Zafar et al., 2004):

$$\Delta V_T(t) = \Delta V_{T\max}[1 - \exp(-(t / \tau_o)^\beta)], \tag{8}$$

where ΔV_{Tmax}, τ_o, and β are the fitting parameters. The stretched exponential model predicts ΔV_T would saturate and reach the maximum value only after a very prolonged stressing, so the ΔV_{Tmax} can be taken as a measure of ΔV_T at ten year device lifetime. Parameter β is defined as a measure of distribution width, and τ_o represents a characteristic time constant of the distribution. This model is suited for processes that, either have two or more distinct mechanisms each with its own τ, or have a single mechanism with statistically distributed values of τ (Van de Walle, 1996).

As can be seen in Fig. 5, the stretched exponential model (solid line) yields very poor agreement with our experimental data in the early phase of stressing, especially for the lower stress voltage. However, this disagreement tends to decrease with increase in stress voltage and/or temperature and, more importantly, the model is in excellent agreement with experimental data in the second stress phase and in saturation, which is of greatest practical importance since we need reliable prediction for ΔV_T at prolonged stress time to estimate the device lifetime. A practical consequence is that stretched exponential fit, if properly constructed, may replace the experimental data in saturation and shorten the experiment execution time, while also allowing the use of higher FC, e.g. 150 mV, as shown in Fig. 5. Figure also shows that time points associated with transitions from the early to second stress phase, which had been determined on the basis of parameter n in the t^n power law dependences, correspond to the points at which the stretched exponential fitting curves start agreeing with experimental data, which confirms the phase feature of NBT stress-induced degradation in power VDMOSFETs (Stojadinović et al., 2005; Danković et al., 2006). The values of parameter β in the stretched exponential fit of our experimental results obtained on VDMOSFETs are found to vary in the range 0.35~0.39 independently on stress

conditions, whereas τ_o decreases with increasing the stress voltage and temperature, which all is in good agreement with findings reported in (Zafar et al., 2003; Zafar et al., 2004). The calculated values of maximum threshold voltage shifts in saturation, ΔV_{Tmax}, are listed in Table 1. As expected, these values, which represent a measure of ΔV_T at ten year lifetime, are found to increase with both NBT stress voltage and temperature.

ΔV_{Tmax} (V)		V_G (V)			
		- 30	- 35	- 40	- 45
T (°C)	125	0.1858	0.2518	0.3109	0.4169
	150	0.2073	0.3241	0.4074	0.5563
	175	0.3188	0.3319	0.4584	0.5694

Table 1. Values of ΔV_{Tmax} in the stretched exponential fit of data obtained on NBT stressed p-channel power VDMOSFETs

3.2 Extrapolation to normal operating bias conditions
The extracted values of the lifetime under experimental conditions are used as the input data for extrapolation to normal operating gate bias conditions. However, the results obtained in this way, which represent the device lifetime values projected to normal conditions, may be strongly affected by several factors, such as the failure criterion, the range of gate voltages applied during the stress, and the mathematical function (i.e. model) used in extrapolation (Aono et al., 2005; Ershov et al., 2005), which are now to be addressed.

3.2.1 Failure criterion
As indicated in Fig. 5, we have defined two different failure criteria as the threshold voltage shifts of 100 mV and 150 mV, respectively, which are now used to project the device lifetime under normal gate bias conditions. Lifetime estimation at three different temperatures for both failure criteria, done by a standard linear extrapolation assuming a maximum normal operating gate voltage to be V_G = - 20 V, is illustrated in Fig. 6, whereas the resulting values of extrapolated lifetime are listed in Table 2. It is quite obvious that lifetime projection strongly depends on the choice of failure criterion.

The saturation tendency of the stress-induced threshold voltage shift gives contribution to a rise of the device lifetime, especially for higher failure criterion (150 mV). It can be seen in Fig. 5 that, in the case of the lower stress voltage, the 150 mV FC line only intersects with the stretched exponential fitting curve but not with experimental data, which means the duration of experiment was shorter than the lifetime at given temperature. However, lower stress voltages are closer to actual operating voltages and are expected to provide more realistic lifetime projection. In this case, the fact that stretched exponential fit successfully predicts threshold voltage shift in saturation enabled us to estimate the lifetime by using the results obtained by fitting instead of the missing experimental ones. If the failure criterion was too high (e.g. 700 mV), its value would fall far above both experimental and fitting curves in Fig. 5, which would result into device lifetime tending infinity. Alternatively, too low failure criterion (below 30 mV) could lead to rather significant underestimation of device lifetime as the value of failure criterion would fall in the early phase of stressing.

Following the above considerations, it appears that most reliable lifetime projections can be obtained by choosing the FC value within the range of threshold voltage shifts observed in the second and saturation phases of device stressing. The correct choice of FC could be verified by considering the values of ten year operation voltage, V_{G10Y}, which is defined as the maximum gate voltage that allows ten years of device operation with V_T shift below the given FC. As can be seen in Table 3, which shows the V_{G10Y} data taken from Fig. 6, the values of ten year operation voltage in most cases fall below the assumed value of maximum operating gate bias voltage of - 20 V, except in the case of 150 mV FC at 125°C. Alternatively, table indicates that in the case of 100 mV FC the devices at 175°C cannot
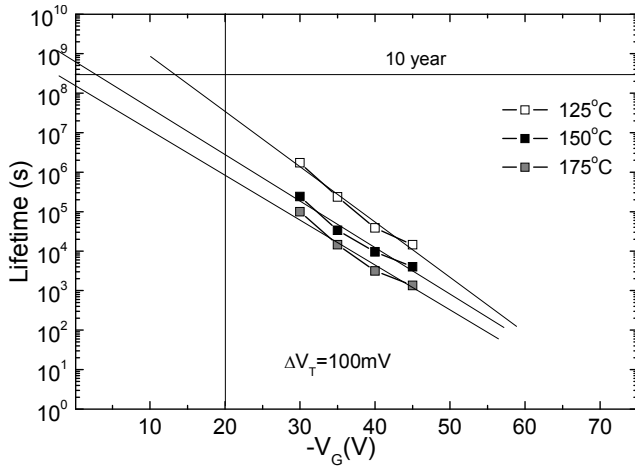
(a)
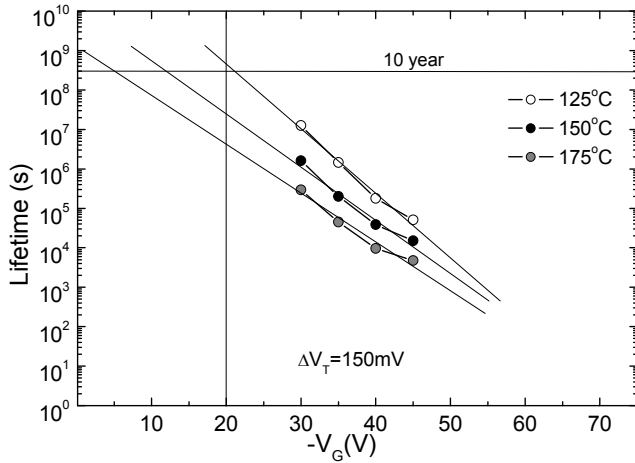
(b)

Fig. 6. Linear extrapolation of the device lifetime at different temperatures for two different values of failure criteria: *a*) ΔV_T =100 mV; *b*) ΔV_T =150 mV.

Lifetime (days)	ΔV_T = 100 mV	ΔV_T = 150 mV
125°C	391.07	5792.10
150°C	30.67	289.86
175°C	9.25	48.06

Table 2. Lifetime projections for operating voltage V_G = - 20 V under two different FC values

approach ten years of operation even without any gate bias, which does not make sense. Thus, the realistic failure criterion for the devices and experimental conditions used in our study appears to fall in the 100 - 150 mV range, which yields the ten year operation voltage within the range of normal gate operation voltages between 0 and - 20 V.

V_{G10Y} (V)	ΔV_T = 100 mV	ΔV_T = 150 mV
125°C	-13	-21
150°C	-3	-12
175°C		-5

Table 3. Estimated values of ten year operation voltage under two different failure criteria

3.2.2 Stress voltage range

Another factor that may affect the value of lifetime estimate is the range of stress voltages used in data extrapolation (Aono et al., 2005). The effects of stress voltage range on uncertainties of both lifetime projection and ten year operation voltage in the case of VDMOS devices stressed at 150°C are illustrated in Fig. 7, where the solid line shows extrapolation over the full range of four different stress voltages applied, whereas the dotted and dashed lines represent extrapolation over the higher and lower voltage ranges, respectively (in these cases the data corresponding to the lowest and highest voltages, respectively, have not been used in extrapolation). As can be noticed in the figure, the uncertainties associated with the choice of stress voltage range may cause the lifetime projection to vary for almost one order of magnitude, while the ten year operation voltage, i.e. maximum allowed V_G, varies for about 8 V.

A schematic drawing shown in Fig. 8 provides further evidence that the lifetime obtained by extrapolating the experimental data to normal operating conditions may strongly depend on the choice of both stress voltage range and failure criterion. As can be seen, experimental lifetime values determined for a given FC in the cases of devices stressed with the highest and lowest voltages fall in the early and in saturation phases, respectively. As a result, both these experimental values are higher, for Δt_2 and Δt_1, respectively, than those the lifetime would assume if found in the second phase, leading to a deviation from linearity in the extrapolation plot shown as an inset graph in Fig. 8. As a consequence, the use of higher stress voltage range for extrapolation to normal bias conditions tends to underestimate the lifetime, whereas the use of lower stress voltage range appears to overestimate it, both contributing to uncertainties shown in Fig. 7. More realistic lifetime estimates could be

expected if the lower stress voltage range, closer to normal operating voltage, was used in extrapolation, but in that case the experiment could take too much time, so the most appropriate solution is to use an expression, such as the one given by stretched exponential function, that provides good fit to experimental data in the low stress voltage range.
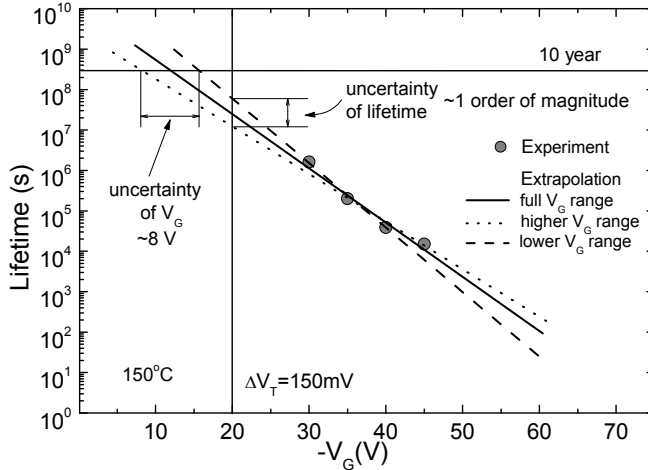
Fig. 7. Uncertainties of the lifetime and ten year operation voltage due to different ranges of stress voltages used in data extrapolation.
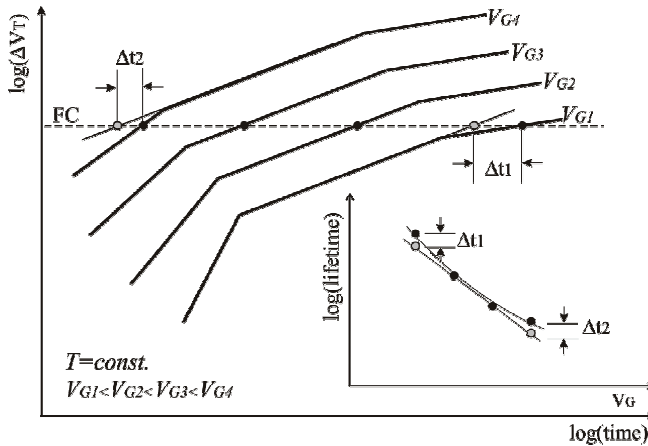
Fig. 8. Schematic illustration to explain the effects of FC and stress voltage range on lifetime projection.

3.2.3 Extrapolation models

In the above analyses of the failure criterion and stress voltage range effects on the lifetime projection we have only applied the standard model, which was based on linear extrapolation function. However, it can be seen in Figs. 6 and 7 that this model in some cases may not provide good fit to experimental data. For this reason, there are several other

commonly used models, such as the so-called V_G and $1/V_G$ models for extrapolation along the voltage axis, which are given by following expressions, respectively:

$$\tau = A \cdot \exp(-B \cdot V_G), \tag{9}$$

$$\tau = A \cdot \exp(B / V_G), \tag{10}$$

as well as the power-law model for extrapolation along the electric field axis:

$$\tau = C \cdot E^{-p}. \tag{11}$$

In the expressions above, τ is the lifetime, V_G and E are the stress voltage and corresponding electric field, respectively, whereas A, B, C, and p are the fitting parameters. These models are based on the experience of different researchers, and it is simple to show, for example, that power-law model can be derived by rearranging the empirical Eq. (5) as follows:

$$t = A^{1/n} E^{-m/n} \Delta V_T^{1/n} \exp(E_a / nkT), \tag{12}$$

and introducing the new parameters, C and p, in Eq. (12) defined as:

$$C = A^{1/n} \Delta V_T^{1/n} \exp(E_a / nkT), \tag{13}$$

$$p = m / n, \tag{14}$$

where t becomes the lifetime, τ, when ΔV_T takes the value of failure criterion.

The uncertainties of both device lifetime and ten year operation voltage in the case of p-channel VDMOSFETs stressed at 150°C, as obtained by the use of four different models for extrapolation, are shown in Fig. 9. As can be seen, each model yields different result, so the
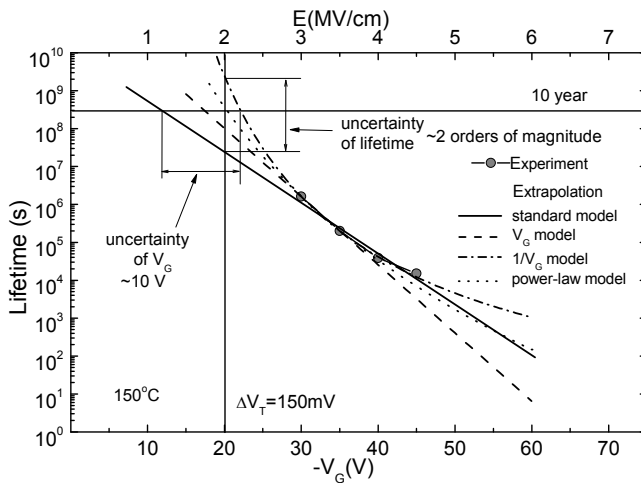
Fig. 9. Device lifetime and ten year operation voltage uncertainties due to different models used for extrapolation.

lifetime estimates obtained using different models may vary for more than two orders of magnitude, while the ten year operation voltage varies for about 10 V. The V_G and especially standard model yield lower values of both lifetime and ten year operation voltage, and both these models do not seem to provide good fit to our experimental data. The V_G model curve shows disagreement with experimental data obtained at higher stress voltages, which is in line with expectation on more realistic estimates if the lower stress voltage range, closer to actual operating voltage, was used for extrapolation. Alternatively, the power-law model provides better fit to our experimental data, but still shows certain disagreement at the higher voltage range, whereas only the $1/V_G$ model appears to fit our experimental results almost perfectly over the full range of stress voltages applied. Thus, the lifetime and ten year operation voltage estimates obtained by the $1/V_G$ model appear to be the least affected by the choice of stress voltage range.

The above considerations may have important consequences on investigations of NBTI related degradation and lifetime prediction not only in power VDMOSFETs but also in other devices exhibiting saturation in the stress-induced threshold voltage shifts over an extended period of NBT stressing. Namely, it appears that lifetime estimates obtained by the power-law, V_G and standard models would approach those obtained by $1/V_G$ model if lower voltages are used for device stressing. However, this could require very long time to perform the experiment. On the other hand, the $1/V_G$ model could provide much faster output since it appears to allow the use of higher stress voltages while still being capable to yield rather accurate lifetime estimates.

3.3 Effects of intermittent annealing

The recovery and annealing of NBTI have received an increased attention recently (Ershov et al., 2003; Tsujikawa et al., 2003; Ershov et al., 2005; Rangan et al., 2005; Huard et al., 2006), so we also have tried to get new insight into the NBTI phenomena by subjecting the devices to a sequence of NBT stress and bias annealing steps. Detailed experiments, with different stress and recovery conditions, have been performed to assess the impact of annealing phase (Danković et al., 2007; Manić et al., 2009), and here we will only present the results obtained on a set of IRF9520 devices subjected to a sequence of three interchanging NBT stress and bias annealing steps as follows: one week of NBT stressing with three different gate voltages (- 35, - 40, and - 45 V) at T = 150°C was followed by one week of positive gate bias annealing with V_G = +10 V also at 150°C, and then the devices were NBT stressed again for one week.

As can be seen in Fig. 10, the initial stress-induced increase of the threshold voltage was most significant in devices stressed with V_G = - 45 V. Initial shifts decreased in all devices during the subsequent annealing, but increased again on repeated NBT stressing. Most rapid decrease on positive bias annealing was observed in devices stressed previously with V_G = - 45 V, and highest increase during the next stressing was found in these devices again.

To assess the impact of intermittent annealing on lifetime projection more closely, Fig. 11 shows the motion of V_T shift over the full 3-step stress/anneal/stress sequence in the case of devices stressed with – 40 V. As can be seen, significant recovery of threshold voltage occurs only in an early stage of the annealing step, so the initial stress-induced ΔV_T did not fall below 100 mV. So, we can conclude that there was a non-reversible component of threshold voltage shift, which resulted from the portion of stress-induced oxide-trapped charge and interface traps that could not have been annealed.

Judging from the data shown in Figs. 10 and 11, the failure criterion that would yield reasonable lifetime projection appears to fall in the range 100-200 mV. Lifetime estimates for
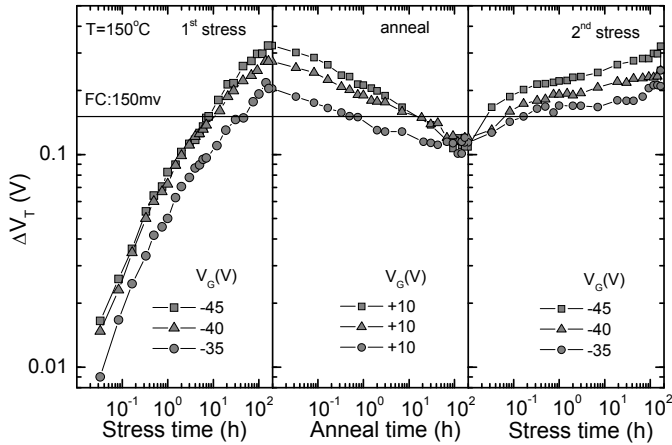
Fig. 10. Threshold voltage shifts during the full sequence of NBT stresses and positive gate bias annealing at 150°C in devices stressed with three different gate voltages.
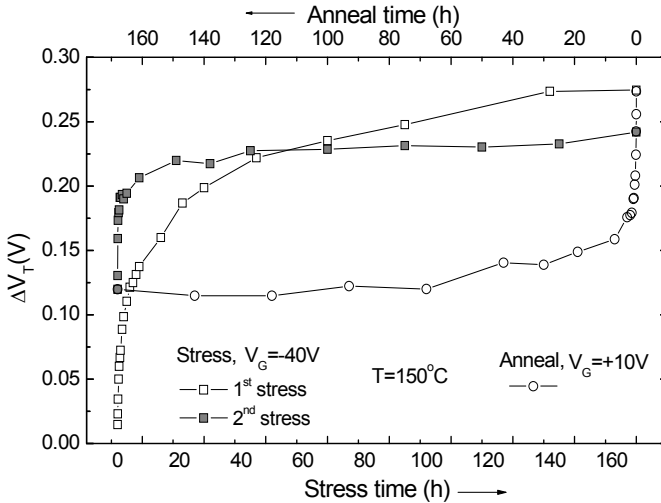
Fig. 11. Threshold voltage shift during the sequence of NBT stresses and positive gate bias annealing at 150°C in devices stressed with – 40 V.

both continuously stressed devices and those subjected to above 3-step stress/anneal/stress sequence, obtained by using the power-law model, are shown in Fig. 12. As can be seen, the two curves overlap in the semi-log plot, yielding practically the same lifetime estimates. The dashed line shows estimation of the difference between the two lifetime projections, $\Delta \tau$. It can be seen that $\Delta \tau < 10^5$ s, i.e. $\Delta \tau$ is less than 2 days, which is negligible in comparison with 10-year lifetime expectation. Therefore, intermittent annealing did not have any apparent impact on device lifetime, which could have been expected since the comparison of data shown in Figs. 3 and 10 indicated that ΔV_T at the end of the above stress/anneal/stress sequence approached ΔV_T observed in continuously stressed devices.
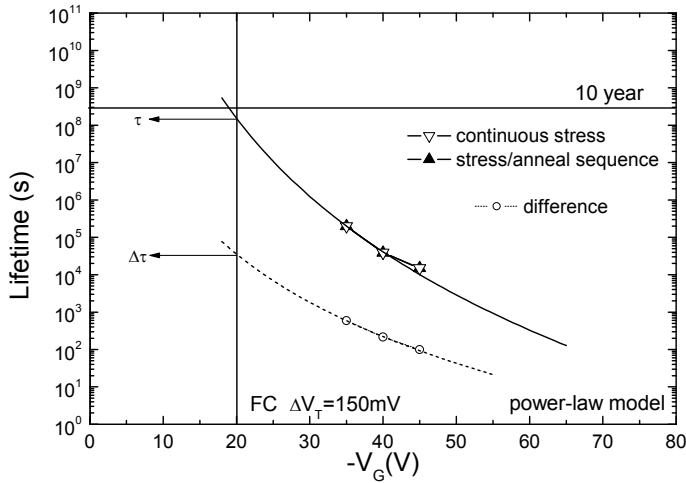
Fig. 12. Lifetime in continuously stressed and sequentially stressed & annealed devices.

4. New approach in estimating the lifetime

As already mentioned, the goal of our study was to estimate the lifetime of investigated VDMOSFETs under normal operating conditions (V_{Go}, T_o) using the results obtained by accelerated NBT stressing (V_G, T). These are the power devices, so we could assume maximum normal bias and temperature to be, for example V_{Go} = - 20 V and T_o = 100°C. In the previous section we have used the accelerated NBT stress data to estimate the lifetime our devices would have if operated under the above gate voltage by means of several models for extrapolation along the voltage axis. Further illustration is provided in Fig. 13, which shows the lifetime estimation in IRF9520 devices by power-law model for three different temperatures applied during the NBT stress. Only extrapolation to V_{Go} = -20 V is
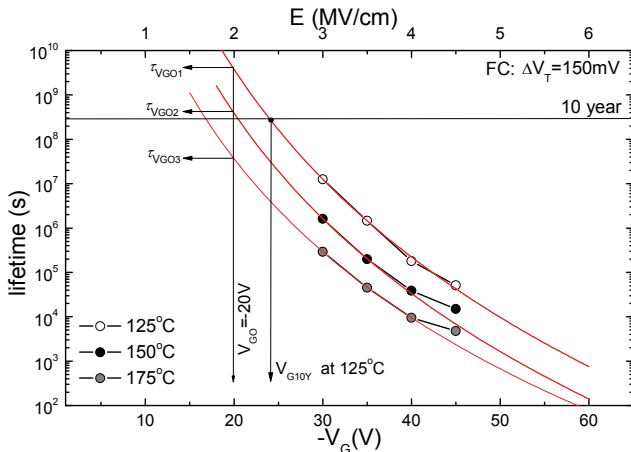
Fig. 13. Estimation of the lifetime and ten year operation voltage by power-law model.

shown, but the same procedure can be used to estimate the lifetime, τ_{VGo}, by extrapolation to any other reasonable operation voltage. In addition, the procedure allows to estimate the ten year operation voltage (V_{G10Y}), which represents the maximum gate voltage that allows 10 years of device operation with stress-induced ΔV_T below FC. However, this procedure yields the τ_{VGoj} and V_{G10Yj} values ($j = 1 - 3$) only for the temperatures applied during the accelerated stressing, which are generally higher than those actually found in device normal operation mode. So, this approach does not offer the possibility to estimate the lifetime at any temperature other than those used in the accelerated stress experiments, which generally applies to all the models discussed in previous section.

4.1 Extrapolation along the temperature axis

Trying to resolve the above issue, we note that all the empirical expressions for the threshold voltage shifts found during the NBT stressing, which were given by Eqs. (3) – (5), include the temperature dependence of stress-induced degradation by incorporating the Arrhenius temperature acceleration term. It is simple to show that any of these expressions can be used to derive a mathematical function that could be suitable as a model for extrapolation along the temperature axis. For example, Eq. (5) can be rearranged to be written as:

$$t = C_3^{-1/n} E^{-m/n} \Delta V_T^{1/n} \exp(E_a / nkT) \cdot \tag{15}$$

Now we introduce the parameters A_2 and B_2 in Eq. (15) as follows:

$$A_2 = C_3^{-1/n} \Delta V_T^{1/n} E^{-m/n}, \tag{16}$$

$$B_2 = E_a / nk, \tag{17}$$

so the Eq. (15) can be rewritten as follows:

$$\tau = A_2 \cdot \exp(B_2 / T), \tag{18}$$

where the stress time, t, has been replaced with the device lifetime, τ, which is correct if ΔV_T takes the value of failure criterion. Here we note the form of Eq. (18) to be the same as that of the $1/V_G$ model given by Eq. (10). This leads to the idea of using Eq. (18), which in analogy with the $1/V_G$ model could be called a $1/T$ model, for extrapolation along the temperature axis so we could get the lifetime for the temperatures expected in device normal operation mode. The first step again is to extract the experimental lifetime values from the ΔV_T vs. stress time plots, but in this case we cannot use the data plots for a fixed temperature shown in Fig. 4; instead, we have to use the plots for a fixed stress voltage, such as those shown in Fig. 14. The extracted values t_j ($j = 1 – 3$) are then used for extrapolation along the temperature axis by means of the model given by Eq. (18). The extrapolation to $T_o = 100°C$ in IRF9520 devices by means of the proposed model, for four different stress voltages, is illustrated in Fig. 15. As can be seen, the current procedure allows to extrapolate the lifetime to any reasonable operating temperature, as well as to estimate a new reliability parameter, which we call a ten year operation temperature, T_{10Y}, (in analogy with well-established parameter, ten year operation voltage, V_{G10Y}) and define it as a maximum temperature that allows 10 years of device operation with stress-induced ΔV_T below FC.
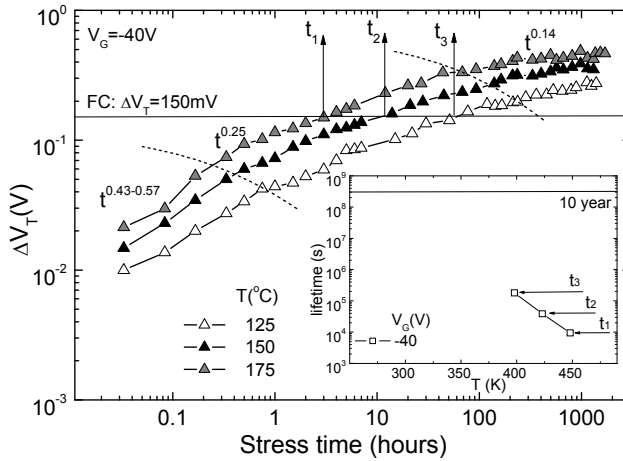
Fig. 14. Threshold voltage shifts in p-channel VDMOSFETs during the NBT stressing with V_G = - 40 V at different temperatures.
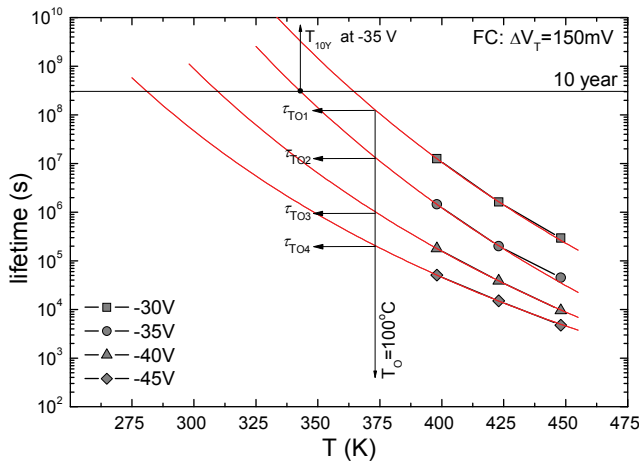
Fig. 15. Estimation of the lifetime and ten year operation temperature by 1/T model.

Thus, the proposed 1/T model offers the possibility of extrapolation to any reasonable operating temperature, but it also has a major drawback of similar nature as that of the models for extrapolation along the voltage axis. Namely, it can be seen in Fig. 15 that the above procedure yields the values of both device lifetime, τ_{Toi}, and ten year operation temperature, T_{10Yi}, (i = 1 - 4) only for the gate voltages applied during accelerated stressing, which are definitely higher than those actually found in device normal operation mode.

4.2 Double extrapolation along the voltage and temperature axes

Each of the two extrapolation procedures considered so far disregards one of the stress acceleration factors, either temperature or voltage, so both procedures may underestimate the device reliability parameters. A reasonable solution could be found by combining the

two procedures, i.e. by performing two successive extrapolations along the gate voltage (or corresponding electric field) and temperature axes, where the latter extrapolation uses the results of the former one as the input data (Danković et al, 2008a). As an example, Fig. 16 illustrates extrapolation along the temperature axis by the proposed 1/T model, where the values of τ_{VGoj} ($j = 1 - 3$) from Fig. 13, obtained by extrapolation along the voltage axis using the power-law model, have been taken as the input data. As can be seen, the two successive extrapolations yield a single lifetime projection, τ_o, which can be associated with devices operated under the normal conditions, both voltage, V_{Go}, and temperature, T_o. The two extrapolations can be done in reverse order as well: the τ_{Toi} ($i = 1 - 4$) data from Fig. 15 obtained by extrapolation to a normal operation temperature can be used for extrapolation to a normal operation voltage by power-law model, which is illustrated in Fig. 17. The two τ_o values obtained in Figs. 16 and 17 are almost identical, so the order in performing the two extrapolations does not seem to be of importance.
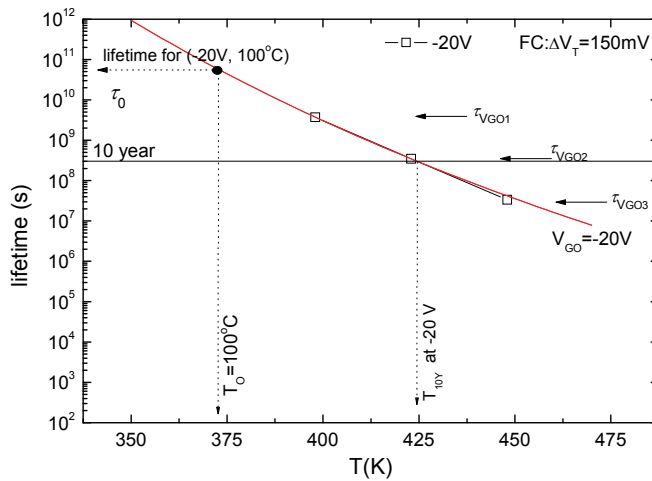
Fig. 16. Lifetime extrapolation to normal operating conditions by 1/T model with input data taken from Fig. 13.

It should be noted that power-law model has been chosen for the above consideration only to explain the proposed procedure requiring double extrapolation along both voltage and temperature axes. In practice, 1/T model for extrapolation along the temperature axis could be combined with any of the other models available for extrapolation along the voltage axis (standard linear model, V_G model, $1/V_G$ model), where the best choice should be the model that provides the best fit to the data. It is also important to note that a double extrapolation approach can be used to estimate the device lifetime for any realistic combination of operating voltages and temperatures, which is schematically illustrated by a drawing shown in Fig. 18. The drawing illustrates both extrapolation routes explained above (extrapolation along the voltage axis followed by the one along the temperature axis and vice versa) and shows the uncertainties in the estimated value of the device lifetime, which have earlier been discussed to originate from the varieties in stress conditions applied, failure criteria chosen, and models used for extrapolation to normal operating conditions. The uncertainties could be reduced if both extrapolation routes were applied in parallel, so the average of the
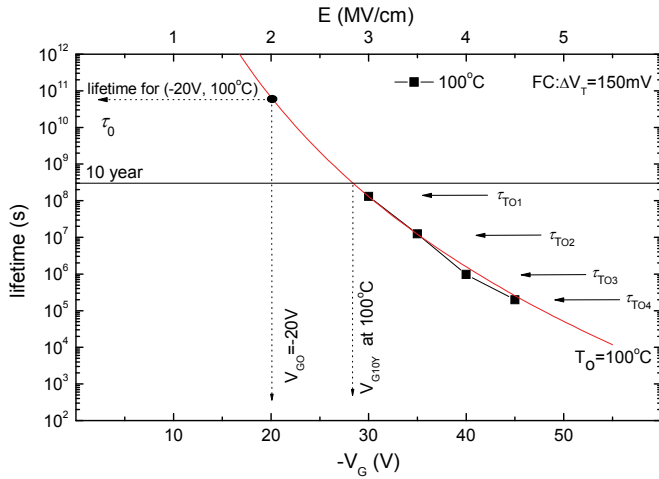
Fig. 17. Lifetime extrapolation to normal operating conditions by power-law model with input data taken from Fig. 15.
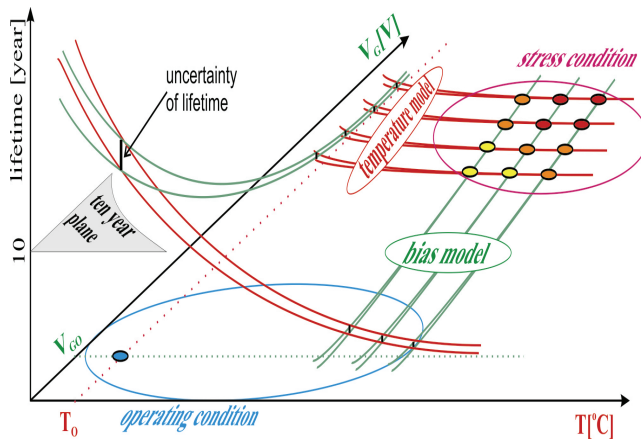
Fig. 18. Schematic drawing illustrating lifetime estimation by double extrapolation to a random combination of operating voltage, V_{GO}, and operating temperature, T_O.

two estimated lifetime values could be used. The procedure of double extrapolation, similar to that illustrated in Fig. 18, can be applied to estimate the other reliability parameters, such as the ten year operation voltage at normal operation temperature, $V_{G10Y}(T_o)$, and the ten year operation temperature under normal operation voltage, $T_{10Y}(V_{Go})$.

Finally, we will demonstrate a possibility of constructing the surface representing the lifetime values projected to a full range of operating voltages and temperatures. As already mentioned, the above double extrapolation approach can be applied to estimate the lifetime for any reasonable combination of operating voltages and temperatures (V_{Go}, T_o), which means the procedure can be re-done for each combination falling within the entire range of operating voltages and temperatures. The set of results obtained in this way can be used to construct the surface representing the lifetime values corresponding to a full range of device

operating conditions. The approach has been applied to our experimental results, and Fig. 19 shows such a surface representing lifetime projections to a full range of operation in the case of NBT stressed p-channel power VDMOS devices IRF9520, where the threshold voltage shift of 150 mV has been taken as the failure criterion. Similar surfaces can be created for different failure criteria, and can be of help in estimating either the lifetime or maximum allowed voltage and temperature for every single device in the operation environment.
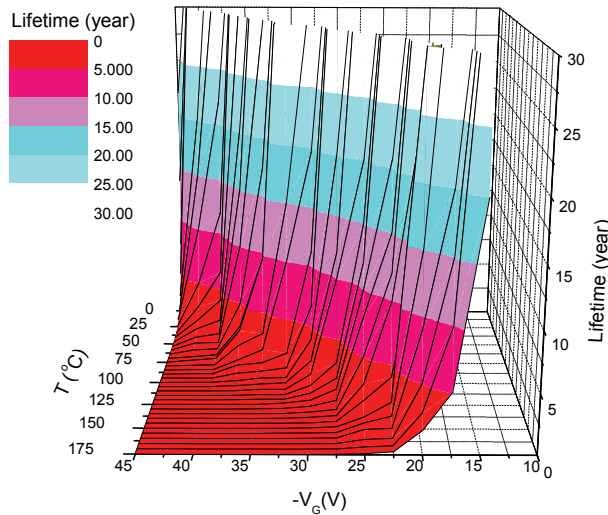
Fig. 19. Surface representing the lifetime estimates in NBT stressed devices for a full range of operating voltages and temperatures with $\Delta V_T = 150$ mV taken as a failure criterion.

5. Conclusion

The NBT stress-induced threshold voltage instabilities in commercial p-channel power VDMOSFETS, as well as the implications of related degradation on device lifetime have been reviewed. The stress-induced threshold voltage shifts have been fitted using different models to estimate the device lifetime and to discuss the impacts of stress conditions, failure criteria, extrapolation models, and intermittent annealing on lifetime projection. Excellent agreement between the stretched exponential fit and experimental data found in later stress phases allowed for an accurate estimation of device lifetime for the lowest stress voltage applied, justifying the need for using the stretched exponential or some other suitable fitting function. The realistic failure criterion for devices and experimental conditions used in our study was found to fall in the 100 - 150 mV range. Estimated values of device lifetime were found to strongly depend on the model used for extrapolation to normal operating conditions, whereas intermittent annealing did not have any apparent impact on device lifetime. The $1/V_G$ model appeared most suited to our experimental results and allowed the use of higher stress voltages while still being capable to yield rather reliable lifetime estimates. However, $1/V_G$ and other models available in the literature offer only extrapolation along the voltage axis, so they are able to provide lifetime estimates only for

the temperatures applied during the accelerated stressing, which are generally above the temperature range observed by normally operated devices. To alleviate this issue, a new approach in estimating the device lifetime, which assumes double extrapolation along both voltage and temperature axes, was proposed. The proposed approach was shown to yield the device lifetime for any reasonable combination of operating voltages and temperatures, including those falling within the ranges normally found in usual device applications.

6. References

Alam, M.A. & Mahapatra, S.A. (2005). A comprehensive model of PMOS NBTI degradation, *Microelectron. Reliab.*, Vol. 45, No. 1, (January 2005) pp. 71-81, ISSN 0026-2714.

Aono, H., Murakami, E., Okuyama, K., Nishida, A., Minami, M., Ooji, Y. & Kubota, K. (2005). Modelling of NBTI saturation effect and its impact on electric field dependence of the lifetime, *Microelectron. Reliab.*, Vol. 45, No. 7-8, (July-August 2005) pp. 1109-1114, ISSN 0026-2714.

Baliga, B.J. (1987). *Modern power devices*, John Wiley & Sons, ISBN 0-471-81986-7, New York.

Benda, V., Gowar, J. & Grant, D.A. (1999). *Power semiconductor devices*, John Wiley & Sons, ISBN 0-471-97644-X, Chichester (UK).

Danković, D., Manić, I., Djorić-Veljković, S., Davidović, V., Golubović, S. & Stojadinović, N. (2006). NBT stress-induced degradation and lifetime estimation in p-channel power VDMOSFETs, *Microelectron. Reliab.*, Vol. 46, No. 9-11, (September-November 2006) pp. 1828-1833, ISSN 0026-2714.

Danković, D., Manić, I., Djorić-Veljković, S., Davidović, V., Golubović, S. & Stojadinović, N. (2006a). Lifetime estimation in NBT stressed p-channel power VDMOSFETs, *Proc. 25th Int. Conference on Microelectronics (MIEL)*, pp. 645-648, ISBN 1-4244-0116-X, Belgrade (Serbia), May 2006, IEEE EDS, Niš (Serbia).

Danković, D., Manić, I., Davidović, V., Djorić-Veljković, S., Golubović, S. & Stojadinović, N. (2007). Negative bias temperature instabilities in sequentially stressed and annealed p-channel power VDMOSFETs, *Microelectron. Reliab.*, Vol. 47, No. 9-11, (September-November 2007) pp. 1400-1405, ISSN 0026-2714.

Danković, D., Manić, I., Davidović, V., Djorić-Veljković, S., Golubović, S. & Stojadinović, N. (2008). Negative bias temperature instability in n-channel power VDMOSFETs, *Microelectron. Reliab.*, Vol. 48, No. 8-9 (August 2008) pp. 1313-1317, ISSN 0026-2714.

Danković, D., Manić, I., Davidović, V., Djorić-Veljković, S., Golubović, S. & Stojadinović, N. (2008a). New approach in estimating the lifetime in NBT stressed p-channel power VDMOSFETs, *Proc. 26th Int. Conference on Microelectronics (MIEL)*, pp. 599-602, ISBN 987-1-4244-1881-7, Ni š (Serbia), May 2008, IEEE EDS, Niš (Serbia).

Demesmaeker, A., Pergoot, A. & De Pauw, P. (1997). Bias temperature reliability of p-channel high-voltage devices, *Microelectron. Reliab.*, Vol. 37, No. 10-11, (October-November 1997) pp. 1767-1770, ISSN 0026-2714.

Ershov, M., Saxena, S., Karbasi, H., Winters, S., Minehane, S., Babcock, J., Lindley, R., Clifton, P., Redford M. & Shibkov, A. (2003). Dynamic recovery of negative bias temperature instability in p-type metal-oxide-semiconductor field-effect transistors, *Appl. Phys. Lett.*, Vol. 83, No. 8 (August 2003) pp. 1647-1649, ISSN 0003-6951.

Ershov, M., Saxena, S., Minehane, S., Clifton, P., Redford, M., Lindley, R., Karbasi, H., Graves S. & Winters, S. (2005). Degradation dynamics, recovery, and

characterization of negative bias temperature instability, *Microelectron. Reliab.*, Vol. 45, No. 1, (January 2005) pp. 99-105, ISSN 0026-2714.

Gamerith, S. & Polzl, M. (2002). Negative bias temperature stress in low voltage p-channel DMOS transistors and role of nitrogen, *Microelectron. Reliab.*, Vol. 42, No. 9-11, (September-November 2002) pp. 1439-1443, ISSN 0026-2714.

Huard, V., Denais M. & Parthasarathy, C. (2006) NBTI degradation: From physical mechanisms to modelling, *Microelectron. Reliab.*, Vol. 46, No. 1, (January 2006) pp. 1-23, ISSN 0026-2714.

Jeppson, K.O. & Svensson, C.M. (1977) Negative bias stress of MOS devices at high electric fields and degradation of MNOS devices, *J. Appl. Phys.*, Vol. 48, No. 5, (1977) pp. 2004-2014, ISSN 0021-8979.

Kimizuka, N., Yamamoto, T., Mogami, T., Yamaguchi, K., Imai, K. & Horiuchi, T. (1999). The impact of bias temperature instability for direct–tunneling ultra–thin gate oxide on MOSFET scaling, *Dig. of Tech. Papers 1999 Symp. on VLSI Tech.*, pp. 73 -74, ISBN 4-930813-93-X, Kyoto (Japan), June 1999.

Krishnan, A.T., Reddy, V. & Krishnan, S. (2001). Impact of charging damage on negative bias temperature instability, *Techn. Dig. 2001 Int. Electron Dev. Meeting* (*IEDM*), pp. 865-868, ISBN 0-7803-7050-3, Washington DC (USA), December 2001.

Krishnan, M.S. & Kol′dyaev, V. (2002). Modelling kinetics of gate oxide reliability using stretched exponents, *Proc. 40th Ann. Int. Reliab. Phys Symp. (IRPS)*, pp. 421-422, ISBN: 0-7803-7649-8, Dallas, Texas (USA), April 2002.

Liu, C.H., Lee, M.T., Lin, C.Y., Chen, J., Schruefer, K., Brighten, J., Rovedo, N., Hook, T.B., Khare, M.V., Huang, S.F., Wann, C., Chen, T.C. & Ning, T.H. (2001). Mechanism and process dependence of negative bias temperature instability (NBTI) for pMOSFETs with ultrathin gate dielectrics, *Techn. Dig. 2001 Int. Electron Dev. Meeting* (*IEDM*), pp. 861-864, ISBN 0-7803-7050-3, Washington DC (USA), December 2001.

Liu, C.H., Lee, M.T., Lin, C.Y., Chen, J., Loh, Y.T., Liou, F.T., Schruefer, K., Katsetos, A.A., Yang, Z., Rovedo, N., Hook, T.B., Wann, C. & Chen, T.C. (2002). Mechanism of threshold voltage shifts (ΔVth) caused by negative bias temperature instability (NBTI) in deep submicron pMOSFETs, *Jpn. J. Appl. Phys.*, Vol. 41, No. 4B, (2002) pp. 2423-2425, ISSN 0021-4922.

Ma, T.P. & Dressendorfer, P.V. (1989), *Ionizing Radiation Effects in MOS Devices and Circuits*, John Wiley & Sons, ISBN-10 047184893X, New York.

Manić, I., Danković, D., Djorić-Veljković, S., Davidović, V., Golubović, S. & Stojadinović, N. (2009). Effects of low gate bias annealing in NBT stressed p-channel power VDMOSFETs, accepted for *20th European Symp. Reliab. Electron Dev., Failure Phys. and Analysis (ESREF)*, Bordeaux (France), October 2009.

Ogawa, S., Shimaya, M. & Shiono, N. (1995). Interface-trap generation at ultrathin SiO_2 (4-6 nm)-Si interfaces during negative-bias temperature aging, *J. Appl. Phys.*, Vol. 77, No. 3, (February 1995) pp. 1137-1148, ISSN 0021-8979.

Rangan, S., Mielke, N. & Yeh, E.C.C. (2003). Universal recovery behaviour of negative bias temperature instability", *Techn. Dig. 2003 Int. Electron Dev. Meeting* (*IEDM*), pp. 341-344, ISBN 0-7803-7872-5, Washington DC (USA), December 2003.

Schlunder, C., Brederlow, R., Ankele, B., Gustin, W., Goser, K. & Thewes, R. (2005). Effects of inhomogeneous negative bias temperature stress on p-channel MOSFETs of

analog and RF circuits, *Microelectron. Reliab.*, Vol. 45, No. 1, (January 2005) pp. 39-46, ISSN 0026-2714.

Schroder, D.K. & Babcock, J.A. (2003). Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing, *J. Appl. Phys.*, Vol. 94, No. 1, (July 2003), pp. 1-18, ISSN 0021-8979.

Schroder, D.K. (2005). Negative bias temperature instability: What do we understand, *Microelectron. Reliab.*, Vol. 47, No. 6, (June 2007) pp. 841-852, ISSN 0026-2714.

Stathis, J.H. & Zafar, S. (2006). The negative bias temperature instability in MOS devices: A review", *Microelectron. Reliab.* Vol. 46, No. 2-4, (February-April 2006), pp. 270-286, ISSN 0026-2714.

Stojadinović, N., Danković, D., Djorić-Veljković, S., Davidović, V., Manić, I. & Golubović, S. (2005). Negative bias temperature instability mechanisms in p-channel power VDMOSFETs, *Microelectron. Reliab.*, Vol. 45, No. 9-11, (September-October 2005) pp. 1343-1348, ISSN 0026-2714.

Stojadinović, N, Manić, I., Davidović, V., Danković, D., Djorić-Veljković, S., Golubović, S. & Dimitrijev, S. (2006) Electrical stressing effects in commercial power VDMOSFETs, *IEE Proc. Circuits, Devices & Systems*, Vol. 153, No. 3, (June 2006) pp. 281-288, ISSN 1350-2409.

Stojadinović, N., Danković, D., Manić, I., Davidović, V., Djorić-Veljković, S. & Golubović, S. (2007). Impact of negative bias temperature instabilities on lifetime in p-channel power VDMOSFETs, *Proc. 8th Int. Conf. Telecomm. in Modern Satellite, Cable and Broadcasting Services*, pp. 275-282, ISBN 1-4244-1467-9, Niš (Serbia), September 2007.

Tan, S.S., Chen, T.P., Ang, C.H. & Chan, L. (2005). Mechanism of nitrogen-enhanced negative bias temperature instability in pMOSFET, *Microelectron. Reliab.*, Vol. 45, No. 1, (January 2005) pp. 19-30, ISSN 0026-2714.

Tsujikawa, S., Mine, T., Watanabe, K., Shimamoto, Y., Tsuchiya, R., Ohnishi, K., Onai, T., Yugami, J. & Kimura, S. (2003). Negative bias temperature instability of pMOSFET with ultra-thin SiON gate dielectrics", *Proc. 41st Ann. Int. Reliab. Phys Symp. (IRPS)*, pp. 183-188, ISBN: 0-7803-7649-8, Dallas, Texas (USA), March-April 2003.

Van De Walle, C.G. (1996). Stretched-exponential relaxation modelled without invoking statistical distributions, *Physical Review B*, Vol. 63, No. 17, (May 1996) pp. 11292-11295, ISSN 0163-1829.

Zafar, S., Callegari, A., Gusev E. & Fischetti, M.V. (2003). Charge trapping related voltage instabilities in high permittivity gate dielectric stacks", *J. Appl. Phys.*, Vol. 93, No. 11, (June 2003) pp. 9298-9303, ISSN 0021-8979.

Zafar, S., Lee, B.H. & Stathis, J. (2004). Evaluation of NBTI in HfO_2 gate-dielectric stacks with tungsten gates, *IEEE Electron. Dev. Lett.*, Vol. 25, No. 3, (March 2004) pp. 153-155, ISSN: 0741-3106.

Radiation Hardness of Semiconductor Programmable Memories and Over-voltage Protection Components

Boris Lončar[1], Miloš Vujisić[2], Koviljka Stanković[2] and Predrag Osmokrović[2]
[1]Faculty of Technology and Metallurgy,
University of Belgrade,
[2]Faculty of Electrical Engineering,
University of Belgrade,
Serbia

1. Introduction

The development of electronics and computer engineering, reflected particularly in the increasing degree of miniaturization and integration of electronic components, leads to their growing usage. Stability of their characteristics is very important in specific working conditions, determined by their specific implementation in different environments. These working conditions include exposure to ionizing radiation. Due to all this, a considerable amount of study and investigation is currently focused on radiation hardness of electronic components. The best indicator of this is the fact that the operation strategy of the United States Ministry of Defense states "nuclear hardness and resistive characteristics should be a part of the design, acquisition and operation of the main and auxiliary systems which perform critical missions in nuclear conflicts". From the scientific point of view, the best proof for stated claims is that each year, since 1964, USA or Canada hosts the IEEE Nuclear and Space Radiation Effects Conference (NSREC), while every two years, starting with 1991, the Radiation and Its Effect on Components and Systems Conference (RADECS) is held in Europe, at which numerous papers and workshops present results of the investigation of radiation hardness and reliability of electronic components in different working conditions and environments. Papers dealing with topics in this field are frequently cited, which is a reliable evidence of the importance of this subject.

Considering the facts stated so far, the topic of this chapter is undoubtedly in the focus of scientific interests and investigations. The chapter may be viewed as a part of a broader study of electronic components reliability, aiming to increase their radiation hardness and produce a construction design which best suits the application.

Reliability of two types of electronic components has been investigated: 1) Commercial Off the Shelf (COTS) semiconductor programmable memories (EPROM and EEPROM); and 2) over-voltage protection elements (transient voltage suppression diodes (TVS diodes), metal-oxide varistors (MOVs), gas-filled surge arresters (GFSA) and polycarbonate capacitors).

2. Radiation hardness of programmable memories

2.1 Theory

Reliability of programmable memories is of great importance due to their widespread application in electronic devices. When hardness design of these components is not efficient enough, radiation effects can cause their partial damage or complete destruction.

Major advantages of Electrically Erasable Programmable Read Only Memory (EEPROM) over Erasable Programmable Read Only Memory (EPROM) components are the elimination of UV erase equipment and the much faster in-the-system erasing process (in milliseconds compared with minutes for high-density EPROM). On the other hand, a major drawback of EEPROMs is the large size of their two transistor memory cells compared to single transistor cells of EPROMs (Prince, 1991). Following the shift from NMOS to CMOS transistor technology, presentday programmable non-volatile memories are mostly CMOS-based, as is the case with both memory models investigated in this paper. Since EPROM and EEPROM cells utilize a similar floating gate structure, their radiation responses are similar as well.

The influence of neutron displacement damage, reflected in the change of minority carrier lifetime, is negligible in all MOS (Metal-Oxide Semiconductor) structures, since they are majority carrier devices. Other types of neutron damage, including secondary ionization and carrier removal, are minimal and indirect (Ma & Dressendorfer, 1989).

CMOS is naturally immune to alpha radiation, due to the shallow well. The formation of electron-hole pairs by an alpha particle will primarily take place in the substrate below the well. The well forms an electrical barrier to the carriers, preventing them from reaching the gate and influencing transistor operation. Any carriers generated in the well itself recombine quickly or get lost in the flow of majority carriers (Srour, 1982).

Gamma radiation may cause significant damage to programmable memories, deteriorating properties of the oxide layer, and was therefore considered in this chapter.

2.2 Experimental procedure

Examination of EPROM and EEPROM radiation hardness was carried out in cobalt-60 (^{60}Co) gamma radiation field. Absorbed dose dependence of the changes in memory samples caused by irradiation was monitored.

The ^{60}Co source was manufactured at Harwell Laboratory. Air kerma rate was measured at various distances from the source with a Baldwin-Farmer ionization chamber. Absorbed dose was specified by changing the duration of irradiation and the distance between the source and the examined memory samples. Absorbed dose in Si was calculated from the absorbed dose in air, by using the appropriate mass energy-absorption coefficients for an average energy of ^{60}Co gamma quanta equal to 1.25 MeV. Mass energy-absorption coefficients for silicon (μ_{enSi}(1.25 MeV) = 0.02652 cm^2/g) and air (μ_{enAIR}(1.25 MeV) = 0.02666 cm^2/g) were obtained from the NIST tables (Hubell & Seltzer, 2004).

Testing was performed on samples of COTS EPROMs and EEPROMs. EPROMs used for the investigations were NM27C512 8F85 components, with 64 KB storage capacity, packaged in a DIP 28 chip carrier. EEPROM samples used were M24128 - B W BN 5 T P, with 16 KB storage capacity, packaged in a PDIP 8 chip carrier. Fifty samples were used for both EPROM and EEPROM testing, from which the average results presented in the paper were obtained. All tests were performed at room temperature (25ºC). Irradiation of a 50-sample batch was conducted in consecutive steps, corresponding to the increase of total absorbed dose. Dose increment was 20 Gy per irradiation step for EPROMs and 60 Gy for EEPROMs.

All memory locations (cells) were initially written into a logic '1' state, corresponding to an excess amount of electron charge stored on the floating gate. This state has been shown to be more radiation sensitive than the '0' state, responding with a greater threshold voltage shift for the same absorbed dose. (Snyder et al., 1989). Effects of gamma radiation were examined in terms of the number of "faults" in memory samples following irradiation. A fault is defined by the change of a memory cell logic state as a consequence of irradiation. The content of all memory locations was examined after each irradiation step, whereby the number of read logic '0' states equaled the number of faults.

Although ionizing radiation effects in MOS structures are generally dose-rate dependent, effects in EPROM and EEPROM cells don't depend on dose rate. Radiation induced charge changes on the floating gate occur extremely fast, and so are in phase with any incident radiation pulse (Lončar et al., 2001).

2.3 Experimental results and discussion

Both the differential and aggregate relative change of the number of faults with the absorbed dose in EPROM samples is shown in Fig. 1 a) and b). These plots were obtained as an average over the results for all 50 samples, and the corresponding statistical dispersion is presented in Fig. 1 c). First faults, of the order of 0.1%, appeared at 1220 Gy. The number of faults increased with the rise of the absorbed dose. At dose values above 1300 Gy, significant changes in memory content were observed.

Changes in EPROMs proved to be reversible, i.e. after UV erasure and reprogramming all EPROM components became functional again - consecutive erasing, writing and reading of the previously irradiated samples was efficiently performed several times.

A repeated irradiation procedure of EPROM samples, following erasure and reprogramming to '1' state, produced faults already at 80 Gy, with significant failures in memory content occurring above 130 Gy, as shown in Fig. 2 a) and b). The lower threshold of fault occurrence upon repeated irradiation shows evidence of the cumulative nature of radiation effects. Fig. 2 c) presents the corresponding statistical dispersion.

The differential and aggregate relative change of the number of faults with the absorbed dose in irradiated EEPROM samples is shown in Fig. 3 a) and b). The plots were obtained as an average over all 50 samples used, and the corresponding statistical dispersion is presented in Fig. 3 c). First faults appeared at 1000 Gy, proving EEPROMs to be more sensitive to gamma radiation than EPROM components. With further dose increase, the number of faults also increased. Moreover, the changes in EEPROMs appeared to be irreversible. Irreversibility of radiation damage in EEPROMs was established based on the fact that the standard erasure procedure was unable to erase the contents of any of the irradiated memory samples. In CMOS EPROMs and EEPROMs, utilizing either N-well or P-well technology, the dual polysilicon gate, consisting of the control and the floating gate, resides over an N-channel transistor. Polysilicon layer floating gate, insulated from the control gate above it and the silicon channel below it by the gate oxide, is used to store charge and thus maintain a logical state. Charge is stored on the floating gate through hot electron injection from the channel in EPROMs, and through cold electron tunneling from the drain in EEPROMs. The stored charge determines the value of transistor threshold voltage, making the memory cell either 'on' or 'off' at readout (Vujisić a et al., 2007).
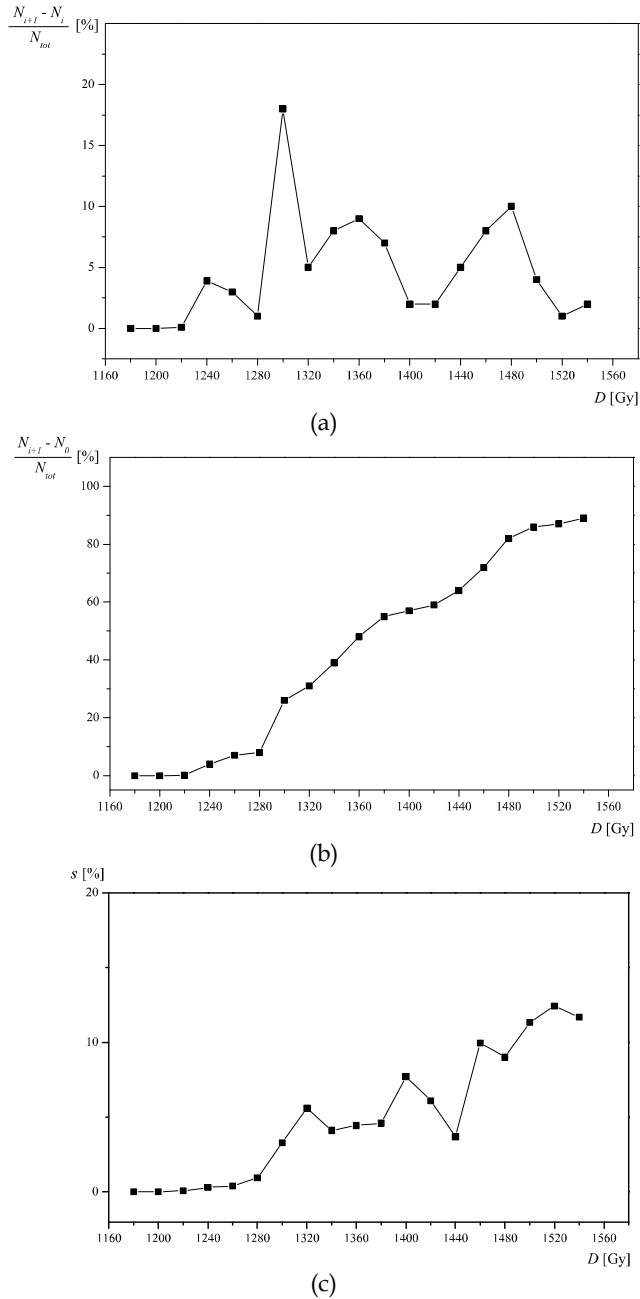
Fig. 1. Average relative change of the number of faults with absorbed dose
in irradiated EPROM samples: a) differential, b) aggregate (N_{tot} = 512 bits, N_0 = 0),
c) corresponding statistical dispersion.
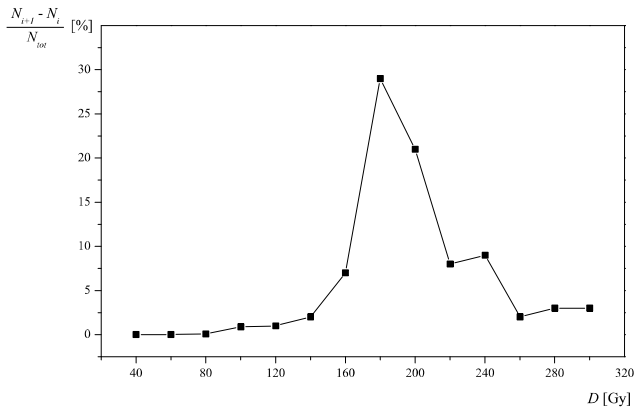
(a)
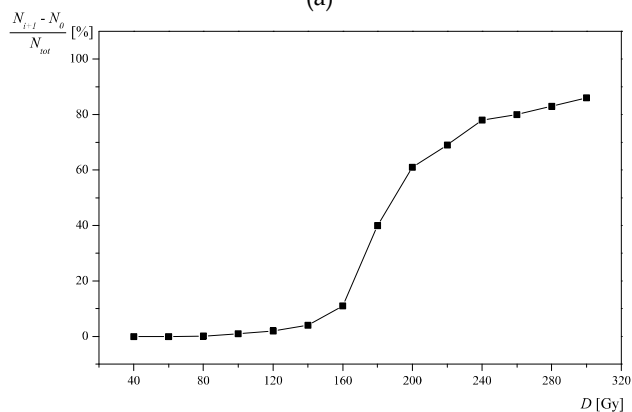
(b)
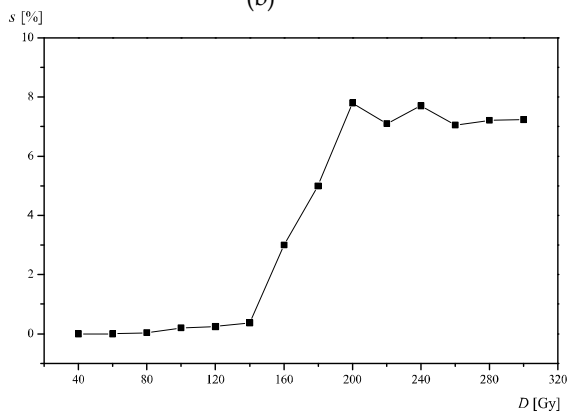
(c)

Fig. 2. Average relative change of the number of faults with absorbed dose in reprogrammed and repeatedly irradiated EPROM samples: a) differential, b) aggregate (N_{tot} = 512 bits, $N_0 = 0$) ,c) corresponding statistical dispersion.
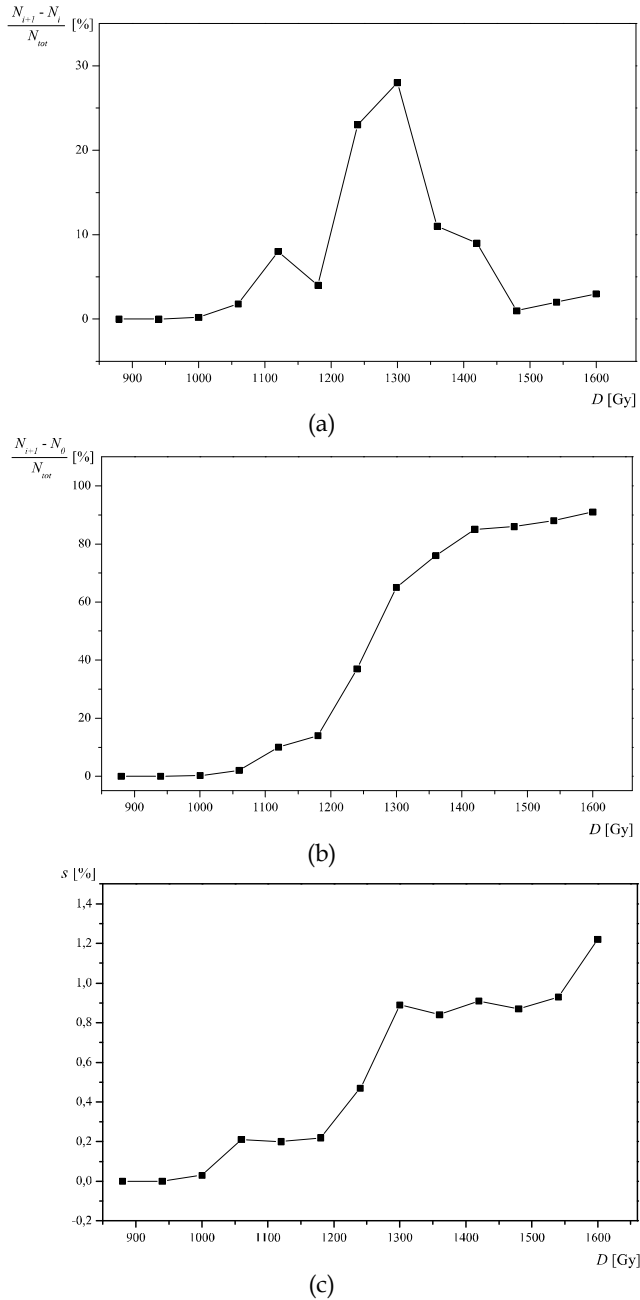
Fig. 3. Average relative change of the number of faults with absorbed dose
in irradiated EEPROM samples: a) differential, b) aggregate (N_{tot} = 128 bits, N_0 = 0),
c) corresponding statistical dispersion.

Passing through the gate oxide (SiO_2), gamma radiation breaks Si−O and Si−Si covalent bonds, creating electron/hole pairs. The number of generated electron/hole pairs depends on gate oxide thickness. Recombination rate of these secondary electrons and holes depends on the intensity of electric field in the irradiated oxide, created by charge trapped at the floating gate, and modulated by the change in charge carrier concentration and their separation within the oxide. The greater the electric field, the larger the number of carriers evading recombination. Incident gamma photons generate relatively isolated charge pairs, and recombination is a much weaker process than in the case of highly ionizing particles.

Secondary electrons which escape recombination are highly mobile at room temperature. In the '1' state of the memory cell, excess amount of electrons stored on the floating gate maintains an electric field in both oxide layers, that swiftly drives the secondary electrons away from the oxide to the silicon substrate and the control gate. While traversing the oxide, radiation-generated secondary electrons themselves create additional electron/hole pairs. Some of the secondary electrons may be trapped within the oxide, but this is a low-probability event, due to their high mobility and the low concentration of electron trapping sites in SiO_2 (Ristić et al., 1998).

In addition to electron/hole pair creation, secondary electrons may produce defects in the oxide by way of impact ionization. Colliding with a bonded electron in either an unstrained silicon-oxygen bond (≡Si−O−Si≡), a strained silicon-oxygen bond, or a strained oxygen vacancy bond (≡Si−Si≡), a secondary electron may give rise to one of the hole trapping complexes. Interaction with a unstrained silicon-oxygen bond gives rise to one of the energetically shallow complexes (≡Si−O•.⁺Si≡ or ≡Si−O⁺.•Si≡, where • denotes the remaining electron from the bond). Strained silicon-oxygen bonds, distributed mainly near the oxide/substrate and oxide/floating gate interfaces, are easily broken by the passing electrons, giving rise to the amphoteric non-bridging oxygen (NBO) center (≡Si−O•) and the positively charged ≡Si⁺ center (known as the E'_s center). Collision of the secondary electron with one of the strained oxygen vacancy bonds, also concentrated near the interfaces, leads to the creation of the ≡Si⁺ •Si≡ center (known as the E'_γ center). Hole traps generated in the bulk of the oxide are shallow, while the centers distributed in the vicinity of the interfaces (NBO, E'_s, E'_γ) act as deep hole traps (Ristić et al., 2000).

Holes generated in the oxide by incident gamma radiation and through secondary ionization are far less mobile than the electrons. They are either trapped in the oxide, or move toward the floating gate under the influence of the electric field in the logic '1' state. Hole transport through the oxide occurs by means of two mechanisms: hopping transport via direct hole tunneling between localized trap sites, and trap-mediated valence band hole conduction. The holes not trapped in the oxide are injected into the floating gate, reduce the net amount of electron charge stored on it, and thereby decrease the threshold voltage of the cell's NMOS transistor. The trapping of holes occurs mostly at the oxide/floating gate interface, where the concentration of deep hole traps is high. The positive charge of these trapped holes, will tend to mask the negative electron charge on the floating gate, again reducing the transistor threshold voltage. Thus, the trapped and the injected holes both produce a negative threshold voltage shift.

Holes moving through the oxide cannot break unstrained silicon-oxygen bonds, but may create defects by interacting with either ≡Si−H or ≡Si−OH bonds, whereby hydrogen atoms and ions (H^o and H^+) are released. Once reaching the oxide/floating gate interface, holes can break both strained silicon-oxygen bonds and strained oxygen vacancy bonds, producing

NBO, E'_s and E'_γ centers. Holes trapped at the oxide/substrate interface which recombine with electrons injected from the substrate may produce another kind of amphoteric defect ($Si_3{\equiv}Si\bullet$, a silicon atom at the interface back bonded to three silicon atoms from the floating gate) (Ristić et al., 2007).

Interface traps may also be generated through direct interaction of incident gamma photons. Another mechanism of interface trap buildup includes hydrogen atoms and ions released by the holes in the oxide. Hydrogen atoms and ions diffuse and drift toward the oxide/floating gate interface. When a H^+ ion arrives at the interface, it picks up an electron from the floating gate, becoming a highly reactive hydrogen atom $H°$, which is able to produce interface traps (Fleetwood, 1992).

Small oxide thickness gives rise to considerable fluctuation of absorbed energy, directly influencing the number of faults in the examined samples. Moreover, the amount of radiation-induced defects acting as electron and hole traps is a complex function of the gate oxide material, as well as of the doping and processing methods used in securing the oxide onto the silicon surface. These are the reasons for the observed variation in the number of faults among the examined memory samples.

Another effect caused by gamma radiation is electron emission from the floating gate. This kind of emission is the basis for standard EPROM erasure by UV radiation. During irradiation, gamma photons cause electrons to be emitted over the floating gate/oxide potential barrier. Once in the oxide, electrons are swept to the substrate or control gate by the electric field. The loss of electrons from the floating gate causes additional decrease of the threshold voltage.

The net effect of charge trapping in oxide and at oxide/floating gate interfaces, as well as of floating gate hole injection and electron emission, is the change of the NMOS transistor threshold voltage. Radiation induced change of the threshold voltage may affect memory cell logic state at readout. Threshold voltage V_T is, hence, the key parameter of memory cell state. Modeling charge stored at the NMOS floating gate as charge on a parallel plate capacitor, threshold voltage can be expressed as:

$$V_T = V_{T0} + \frac{q_s d}{\varepsilon} \tag{1}$$

where V_{T0} is the initial transistor threshold voltage, q_s is the gate surface charge density, d is the oxide thickness between the control and floating gate, ε is the oxide dielectric constant. This model disregards the dependence of threshold voltage on the actual position of the trapped charge sheet within the oxide. The influence of gamma irradiation on programmable memories is manifested through the change of the net gate surface charge density. Threshold voltage as a function of the absorbed dose can be represented by the empirical relation:

$$V_T(D) = V_T^{eq} + (V_{T0} - V_T^{eq})e^{-\alpha D} \tag{2}$$

where a is a constant dependent on the type and energy level density of the traps in the oxide, V_T^{eq} is the threshold voltage at extremely high doses, when an equilibrium of the three dominant processes contributing to the change of gate surface charge density - hole trapping, hole injection and electron emission - is achieved (Messenger & Ash, 1992).

The cumulative nature of gamma radiation effects observed in EPROM components can be attributed to the fact that no annealing of radiation induced interface states occurs at ambient temperature. Higher sensitivity of the tested EEPROM components to gamma radiation is a consequence of a more pronounced radiation induced electron emission from the floating gate over the thin oxide region (≈ 10 nm) between the floating gate and the drain (Wrobel, 1989).

UV photons with an energy lower than the bandgap of silicon dioxide (≈ 9 eV) are incapable of creating electron-hole pairs in the oxide, but are capable of exciting electrons from the silicon substrate into the oxide, where they recombine with the trapped holes. Irradiation of EPROMs by UV light during erasure partially reduces the radiation-induced trapped charge from previous exposure to gamma photons. This light-induced annealing of trapped holes can account for the observed reversibility of changes in EPROMs. Since EEPROM erasing process involves no UV irradiation, this effect is absent for these components. Thermal annealing of holes trapped at deep interface traps is not evident at ambient temperatures. Current-induced annealing of trapped holes, due to recombination of holes with electrons being driven from the floating gate to the drain, could be expected to occur during EEPROM electrical erasure. However, this kind of annealing is known to require much longer times compared to the duration of a standard EEPROM erasing procedure. On the whole, no significant annealing of trapped holes occurs in EEPROMs, and hence radiation-induced changes in these components appeared irreversible on the time scale of experiments performed in this paper (~ 10 hours) (Raymond, 1985).

3. Radiation hardness of over-voltage protection components

3.1 Theory

The power or signal lines over-voltage transients arise directly from the commutation process, electrostatic discharge, lightening stroke, or indirectly as a consequence of interaction between wire structures of the system and an electromagnetic field. An efficient over-voltage protection is highly important for reliable functioning of the protection equipment. Continous or temporary malfunctioning of the equipment caused by surge, could be a result of an improperly designed circuitry protection. Resistance to the occurrence of over-voltage is significantly reduced through miniaturization of electronic components. Therefore, more attention is paid to the protection of over-voltage components. This problem is particularly interesting when a fast electromagnetic pulse and radiation are simultaneously applied to the electronic components.

The over-voltage components could be generally divided into non-linear and linear components. Non-linear group includes the following: Transient Suppressers Diodes (TSD), Metal-oxide Varistors (MOV) and Gas Filled Surge Arresters (GFSA). Linear group of components includes capacitors, inductors, resistors or their combinations-filters.

The main disadvantage of linear over-voltage protection components is the frequency dependent protection efficiency. For all non-linear protective devices this is only a marginal problem - only at high frequencies may the protection efficiency of these devices suffer certain degradation (Lončar et al., 2005). Reliable operation of all electronic components (including over-voltage protection) at high temperatures is extremely important - particularly in a "mission critical" application. If the over-voltage protection is stabile and effective over a wide temperature range, it will significantly contribute to overall system reliability (Markov, 1987).

3.2 Experiment
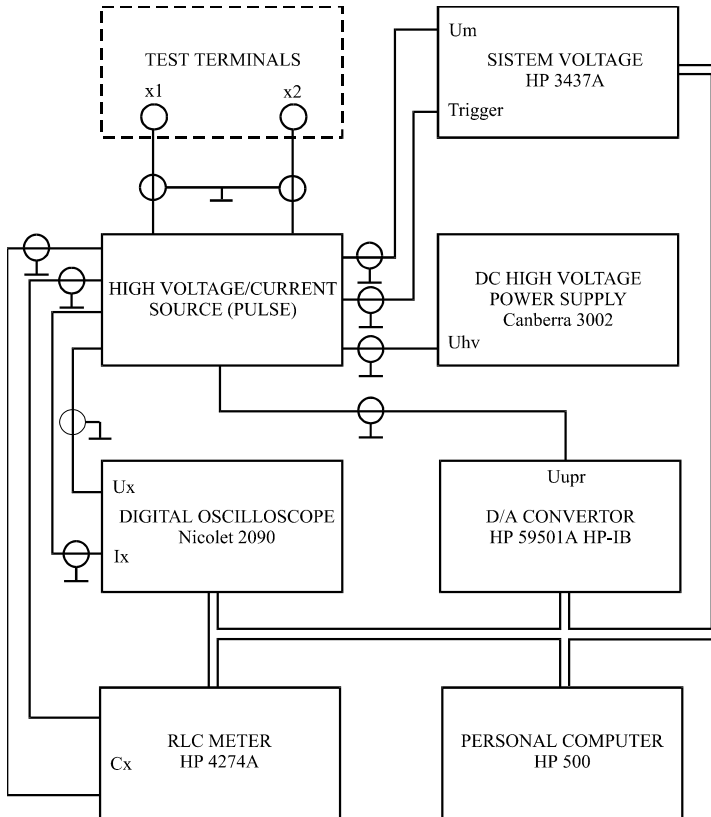
Fig. 4. Block scheme of measuring system

The examination of over-voltage protection components was carried out on the following commercial components:

1. Transient Suppresser Diodes (TSD) nominal voltage 250 V, maximum pulse current 1A;
2. Metaloxide Varistors (MOV) nominal voltage 230V, maximum impulse electric current 1200 A/ 2500 A;
3. Polycarbon Capacitors (nominal voltage 160 V, capacity 1μF);
4. Gas Filled Surge Arresters (GFSA) d.c. breakdown voltage 470V.

We examined the $n_+\gamma$ radiation from californium ^{252}Cf isotope influence on following characteristics of TSD and MOV:

1. volt-ampere characteristic;
2. volt-ohm characteristic;
3. nonlinear coefficient $\alpha = \log (I_2/I_1)/\log (U_2/U_1)$;
4. breakdown voltage.

The effects of $n_+\gamma$ radiation on following Polycarbon Capacitors characteristics were examined:

1. the dielectric loss factor, tg δ (measuring frequency f = 100 kHz);
2. capacitance, C (measuring frequency f = 100 kHz).

The effects of $n_+\gamma$ radiation on following GFSA characteristics were examined:
1. the random variable "pulse breakdown voltage";
2. the random variable "d.c. breakdown voltage";
3. the pulse shape (volt-second) characteristic.

Experiments were conducted using high voltage and measuring equipment consisting of current source, having the maximum voltage between terminals x_1 and x_2 of 3000 V, digital oscilloscope, D.C. high voltage power supply and personal computer, (Fig. 4.).

For TSD and MOV testing, the double exponential (8 x 20μs) current pulse method was applied. Double exponential voltage pulses (1,2 x 50μs) were applied in the GFSA and Polycarbon Capacitors testing. A random variable "D.C. breakdown voltage" test conducted on a GFSA consisted of 20 series each one having 50 measurements (1000 activation). Prior to experiments in each of the series, the element to be tested was conditioned with 25 breakdowns. A 30-second pause between two successive measurements was introduced. Polycarbon Capacitors were examined using the same measuring procedure used for impulse tests of GFSA. A RLC meter was used for this experiment.

All measuring instruments were protected from electromagnetic interference by electromagnetic shielding. The experimental procedure was fully automated. This approach assures very high accuracy of measurement and good repeatability of results. Specialized PC based control software (HP-IB or IEEE488 protocol) was developed to provide overall experiment sequencing, measurement and data acquisition (digital oscilloscope) and easy change of parameters (pulse modification - voltage and current source was generated using a D/A converter) (Osmokrović et al., 2006).

Table I and II present F_n and F_γ (neutron fluence and gamma flux respectively) dependencies versus N_F (the exposure number) for TSD & MOV and Polycarbon Capacitors, respectively.

N_F	$Fn(n/cm^2) \cdot 10^{10}$	$F\gamma(\gamma/cm^2) \cdot 10^{13}$
0	0	0
1	3.55	8.66
2	7.10	17.3
3	10.66	26

Table 1. Values for neutron fluence (F_N) and gamma flux (F_γ) versus N_F (the exposure number) for TSD and MOV

N_F	$Fn(n/cm^2) \cdot 10^{10}$	$F\gamma(\gamma/cm^2) \cdot 10^{13}$
0	0	0
1	2.79	6.8
2	5.59	13.6
3	8.37	20.4

Table 2. Values for neutron fluence (F_N) and gamma flux (F_γ) versus N_F (the exposure number) for Polycarbon Capacitors

3.3 Results and discussion
3.3.1 Radiation hardness of TSD

In Figures 5 (a)-(d) volt-ampere, volt-ohm, breakdown voltage and the nonlinear coefficient versus n+γ fluence/flux characteristics are shown respectively (Vujisić b et al., 2007).
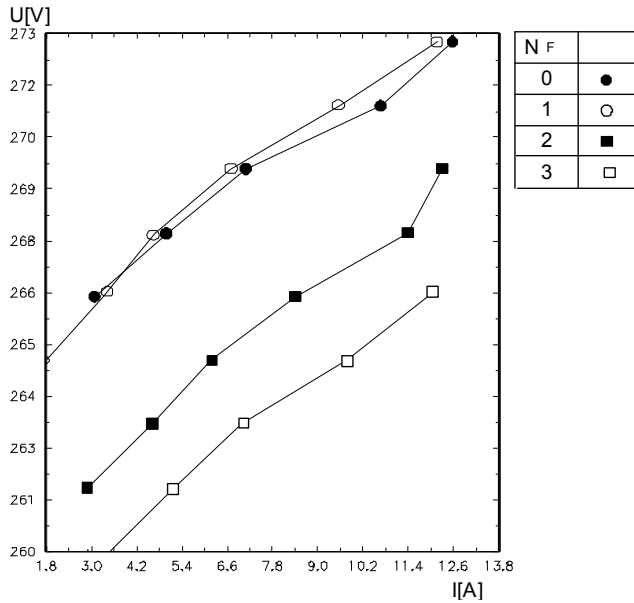
Fig. 5(a). The TSD volt-ampere versus n+γ fluence / flux characteristic
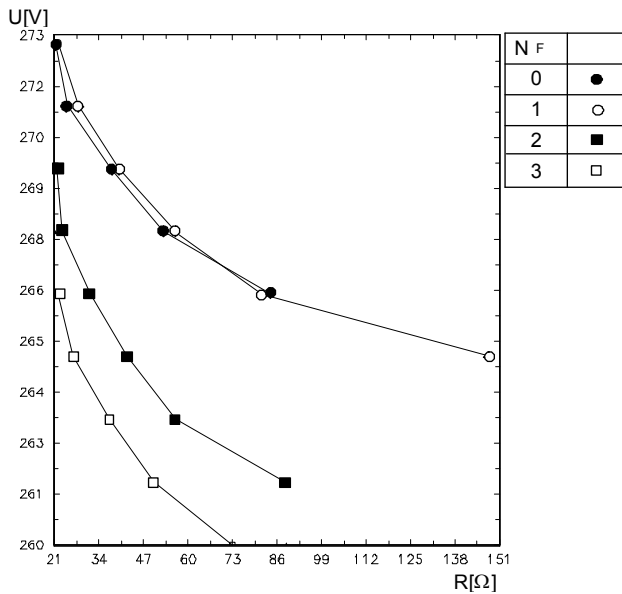
Fig. 5(b). The TSD volt-ohm versus n+γ fluence / flux characteristics

Upr [V]
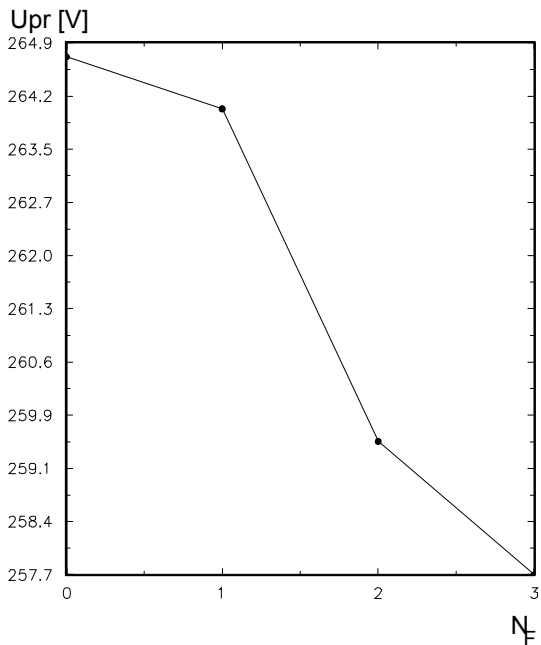
Fig. 5(c). The TSD breakdown voltage versus n+γ fluence / flux characteristics

α
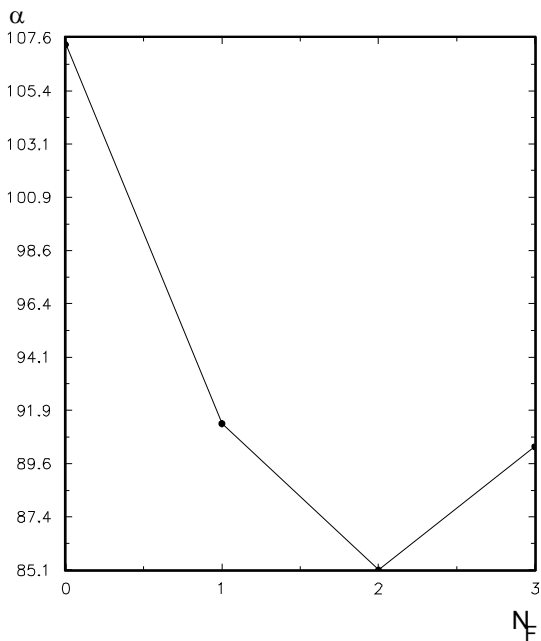
Fig. 5(d). The TSD nonlinear coefficient versus n+γ fluence / flux characteristics

According to these diagrams TSD exhibits breakdown voltage drop, due to increase of the volt-ampere slope, and consequently decrease of non-linear coefficient.

The changes which were noticed could be explained by irreversible TSD material changes caused by $n_+\gamma$ radiation. This radiation influences generation different radiation defects which directly causes degradation of electric characteristics of over - voltage protection components (life time, minority charge carrier concentrations, mobility, specific resistance). The most typical radiation defect for applied doses is Frenkel (Van Lint et al., 1980). The energy levels of these defects are located inside the energy gap. Although minority charge carriers life time is determined by bulk recombination velocity at traps and local centers, recombination efficiency, e.g. minority charge carriers life time depends on traps concentration and probability of minority charge carriers trapping by recombination centers. For this reason these defects represent very efficient recombination centers. Consequently, recombination likelihood becomes higher, which decreases minority charge carrier's life time. Thus, TSD Reverse Current becomes larger which causes breakdown voltage drop. Tested TSD has relatively high voltage protection element (250V), therefore relatively small radiation doses caused significant breakdown voltage drop (Fig. 5c) (approximately 50 %).

Second radiation effect is related to minority charge carriers mobility and concentration decrease which causes increase of specific resistance of the initial semiconductor material. Thus, TSD material exhibits specific resistance increase. Consequently, nonlinear coefficient decreases worse and the TSD protection characteristics degrades. Influence of radiation neutron components on degradation of TSD protection characteristics is much larger then influence of γ component. It is noticed that one displacement of atoms from silicon crystal lattice under influence of γ photons with energy of 1,5 MeV causes 140 displacements under influence of same energy neutrons. On other hand that ratio is 2,7 by γ photon, e.g. 150 by neutrons at energy of 120 MeV. This shows that contribution of γ radiation effects to displacement is negligible to that of neutron radiation effects (Holmes- Siedle & Adams, 2002).

3.3.2 Radiation hardness of MOV

In Fig. 6(a)-(d) the MOV volt-ampere, volt-ohm, breakdown voltage and nonlinear coefficient versus $n_+\gamma$ fluence/flux characteristics are represented, respectively.

According to diagrams, MOV exhibits breakdown voltage drop due to increase of the volt-ampere slope and consequently decrease of nonlinear coefficient.

Generally, metaloxide varistors exhibits more pronounced protection characteristics degradation compared to TSD.

The obtained results are not in full agreement with the expected ones (Malinarić, 1985). Therefore, the experiments are performed a number of times in strictly controlled conditions with high rate of reliability and repeatability and every time with same results.

Therefore, appearance of dislocation in metaloxide varistor structure does not change carrier life time, so larger influence of $n_+\gamma$ radiation on metaloxide varistor characteristics (volt - ampere, volt -ohm, breakdown voltage and nonlinear coefficient) is not expected. But, experimental results show significant influence of $n_+\gamma$ radiation on degradation of metaloxide varistor characteristics. These results could be explained by dislocation capturing charge carriers, which causes decrease of free charge carriers. These further leads to decrease of metaloxide varistors specific conductivity and to increase of local electric

field nearby dislocations (Mahan et al., 1979). Decrease of specific conductivity explains volt - ampere and volt - ohm curves. On the other hand decrease of breakdown voltage is explained by increase of local electric fields.
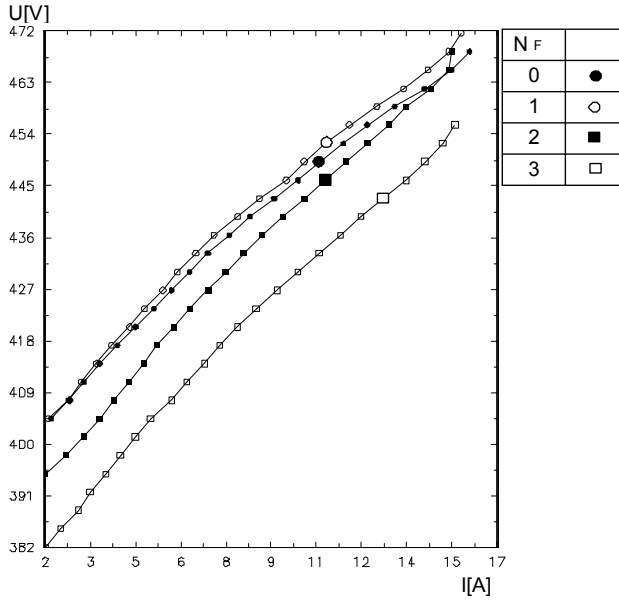
Fig. 6(a). MOV volt-ampere versus n+γ fluence / flux characteristics
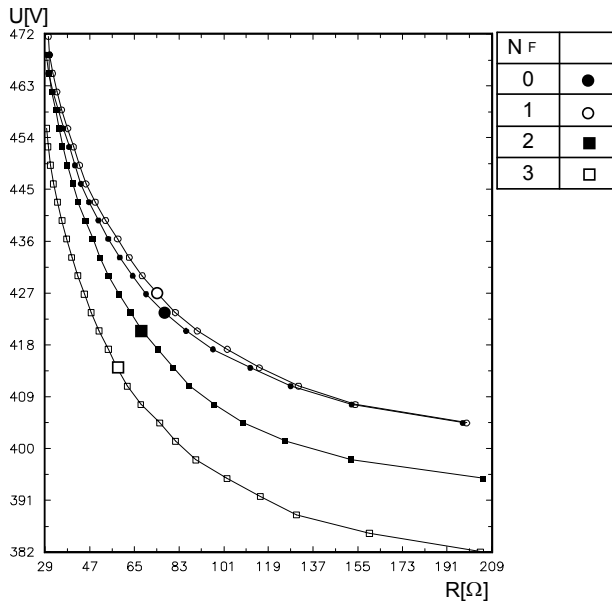
Fig. 6(b). MOV volt-ohm versus n+γ fluence / flux characteristics
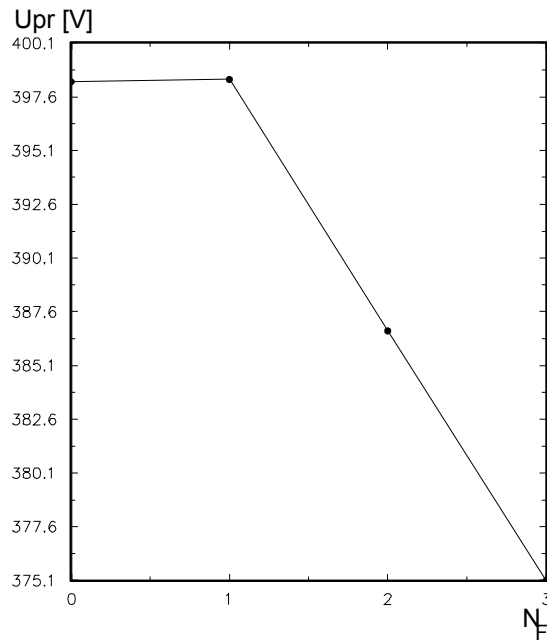
Fig. 6(c). MOV breakdown voltage versus n+γ fluence/ flux characteristics
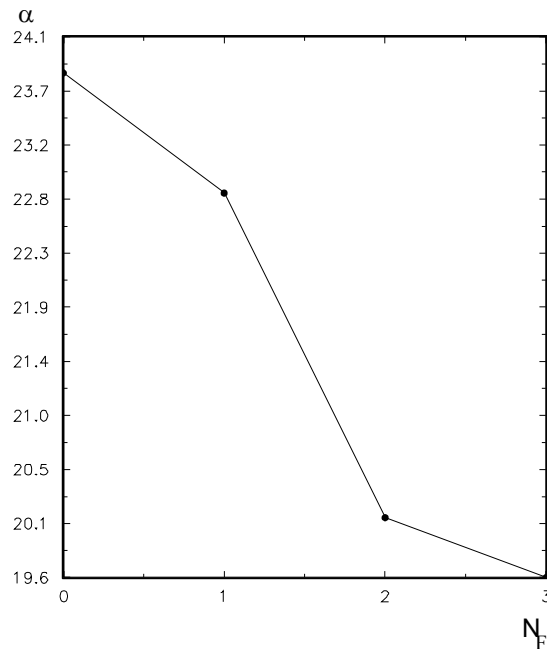
Fig. 6(d). MOV nonlinear coefficient versus n+γ fluence / flux characteristics

Irreversible changes of electric characteristics of metaloxide varistors are caused by inelastic interactions with atom material nucleus. Since the cross section of inelastic interaction of neutron component is larger than corresponding γ component noticed effect is mostly result of neutron component radiation (Messenger & Ash, 1992).

3.3.3 Radiation hardness of polycarbonate capacitors

In Fig. 7 the change in the capacity of polycarbon capacitors versus $n+\gamma$ fluence/flux radiation is depicted. Examined capacitor has a very high insulated degree (10 -100 TΩ) and small losses (tgδ = 0,0015). During examination a measurable influence of $n+\gamma$ radiation on the loss factor tgδ was not found. According to this diagram we can conclude that radiation causes a decrease in the capacity of capacitors. After repeating this experiment in 120 hours, effects of reversible nature were noticed (Lončar et al., 2007).

Influence of neutron and γ radiation on polycarbonat can be observed in two phases. As a result of first phase ionizing radiation free electrons, positive ions and excess molecules are formed. In second phase free radicals are formed. Huge amount of free radicals in polymer leads to irreversible changes which is consist of destruction processes and structure changes. Destruction is process of breaking main chains of connected macro molecules which causes molecule mass drop and gas occurring. Destruction mechanism is determined by individual characteristics of radiated material. Nowadays, for larger number of polymers destruction process is not well known to the end. It should be noticed that small radiation doses can change physical characteristics of polymer materials (Clegg & Collyer, 1991).
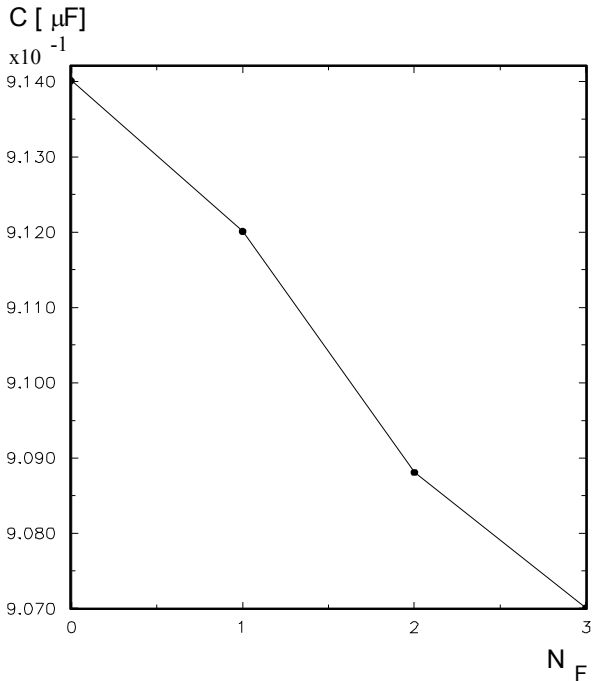
Fig. 7. Polycarbon capacitor value versus $n+\gamma$ fluens/ flux characteristics

The decrease of capacitors capacity when under the radiation of n+γ flux can be explained by forming the ionized structure inside the dielectric volume. Those structures influence the partial screening of the electric field in the capacitors. Greater number of ions appearing in dielectric initiate higher influence of ion type polarization on dielectric constant of material. This leads to decrease of capacity. Also, ion pairs by there local field partially screening external electric field (Gromov & Evdokinov, 1984). Still the change in the capacitance of the capacitors caused by it is relatively small. The existence of these structures in the dielectric leads to its aging and to the decrease of the breakdown voltage . The reversibility of these phenomena is a result of recombination processes. This shows that the doses of radiation were not high enough to cause significant changes in the molecular structure of the dielectric.

3.3.4 Radiation hardness of GFSA

3.3.4.1 The effect of induced radiation

A GFSA dc breakdown voltage versus neutron fluence dependence is presented in Fig. 8. Fig. 9 present the GFSA pulse shape characteristics before and after the exposure to the radioactive source (Lončar et al., 2003).
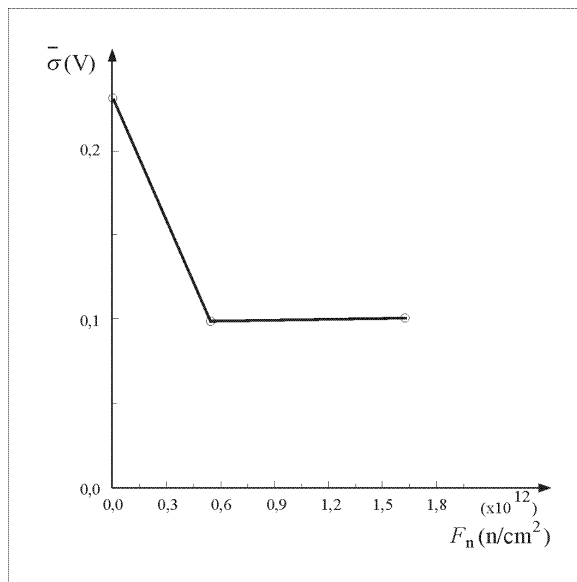
Fig. 8. The GFSA dc breakdown voltage versus neutron fluence characteristic

Since cross section for capturing neutrons has large enough value only for thermal and slow neutrons and due to structure of californium source fission spectra relatively small part of neutrons takes part in activation of GFSA material (Messenger & Ash, 1992). Besides this restriction as a result of GFSA radiation a change of its electrical characteristics was noticed. GFSA was influenced by neutron fluence of $5,4148 \times 10^9$ n/cm² same as neutron fluence of $16,244 \times 10^{11}$ n/cm². Although, except the neutron component part of radiation was also γ
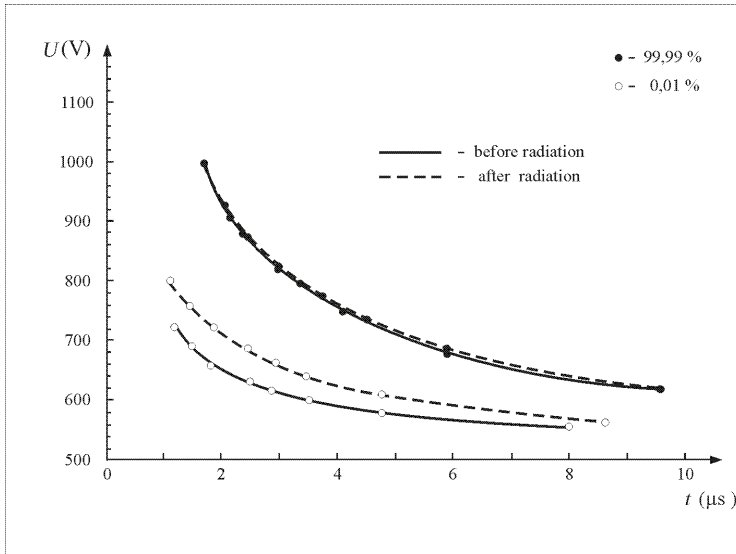
Fig. 9. The GFSA pulse shape characteristic

component, which can influence only changes of electric characteristics of GFSA until its displace from field of radiation. It means that we can observe effects of radiation of GFSA only as influence of neutron fluence.

The obtained results show that after irradiating GFSA the standard deviation of the static breakdown voltage is significantly decreased. The pulse voltage tested GFSA shows that the irradiated GFSA acts more readily and has somewhat narrower voltage-time characteristic than unirradiated GFSA. In other words it had a smaller discharge value of dc breakdown voltage (Fig.10) (1,84% and before radiation 3,11%). This means that its protective characteristics are improved.
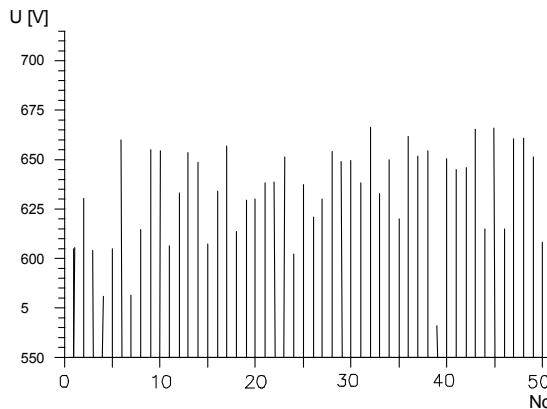
Fig. 10(a). Chronological series of measured values of GFSA d.c. Breakdown Voltage before radiation
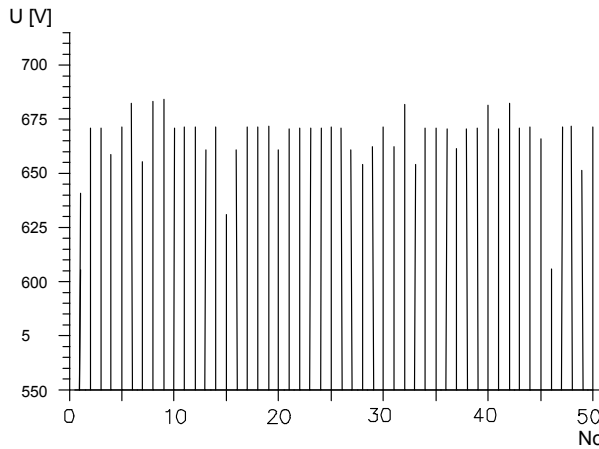
Fig. 10(b). Chronological series of measured values of GFSA d.c. Breakdown Voltage immediately after radiation

A faster response of an irradiated GFSA was the consequence of the higher concentration of free electrons in the inter-electrode gap stimulated by a gas ionization. This ionization was induced by neutron activation of a GFSA material. Faster response time of GFSA can be explained as a consequence of the increased availability of free electrons that caused a statistic time decrease.

Figures 11(a) and 11(b) present the GFSA activation analysis diagrams immediately after the exposure to the radioactive source and six hours after radiation effects obtained by γ-ray
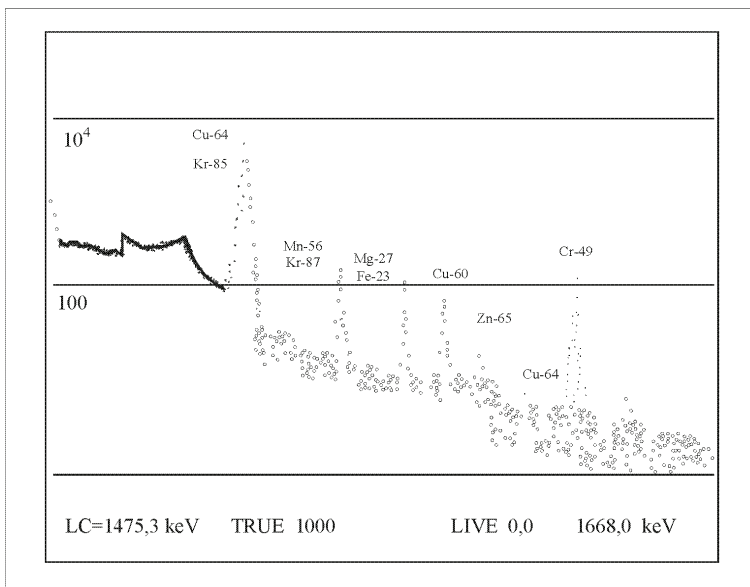
Fig. 11(a). Diagram of the GFSA activation analysis immediately after radiation effects
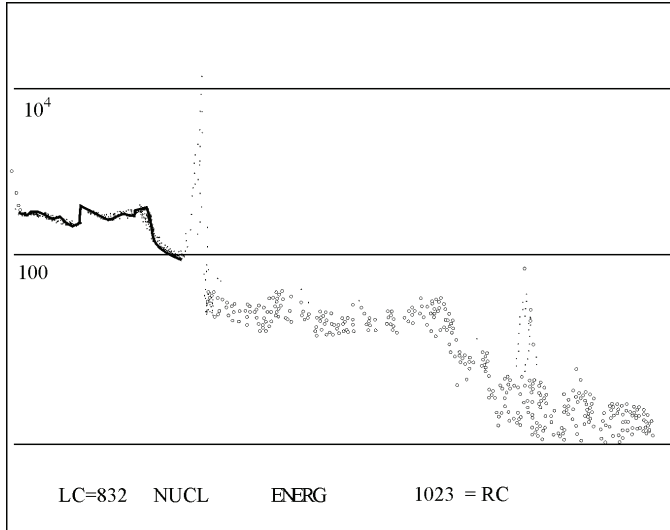
Fig. 11(b). Diagram of the GFSA activation analysis six hours after radiation effects

spectrometer, respectively. Identified radioactive isotopes are recorded close to the expected energy peaks. The activity of these isotopes consists of both γ and β component. Due to the induced radioactivity, the gas ionization is intensified and the statistical time of a pulse breakdown voltage is reduced. Neutron radiation effects improve the pulse shape characteristic for a short period of time. This effect of neutron radiation disappears quickly, as the half-life time of induced activity varies from several minutes to several hours. This fact is confirmed clearly in the diagram for the activation analysis of radiated GFSA taken after six hours (Fig. 11(b)). From these diagrams one notices, that neutron activation products have completely decayed and the device GFSA had recovered back to the unirradiated state.

3.3.4.2 The effect of built-in radiation

In GFSA model α and β sources were built-in. [241]Am used as an α-emitter in experiment has the activity of 1735 Bq with the energy of 5485.6 keV and the activity of 267 Bq with the energy of 5442.9 keV. As a β-emitter [90]Sr -^{90}Y source was used (first continuous spectrum from ^{90}Sr has a total average of β- energy of 196 keV and the maximum energy of β electrons as highest energy in energy range from endpoint of internal Bremsstrahlung 546 keV, second continuous spectra from ^{90}Y has total average β- energy 934 keV and maximum energy of β electrons as the highest energy in the energy range from the endpoint of internal Bremsstrahlung 2282 keV). In order to get a detailed insight into how the radiation influences the GFSA characteristics, the types of gas and its pressure were varied.

In Fig. 12, the GFSA dc breakdown voltage versus a variable 'pd' (p is the gas pressure in the chamber, d is the interelectrode gap) is presented before the radiation and with a built-in α and β radioactive sources. In Fig. 13, GFSA voltage-time characteristics with built-in α and with built-in β radioactive sources are given (Osmokrović a et al., 2002).

On the basis of the experimental results we can conclude that built-in radioactive sources significantly improve the protective characteristics of GFSA. This is especially obvious in the
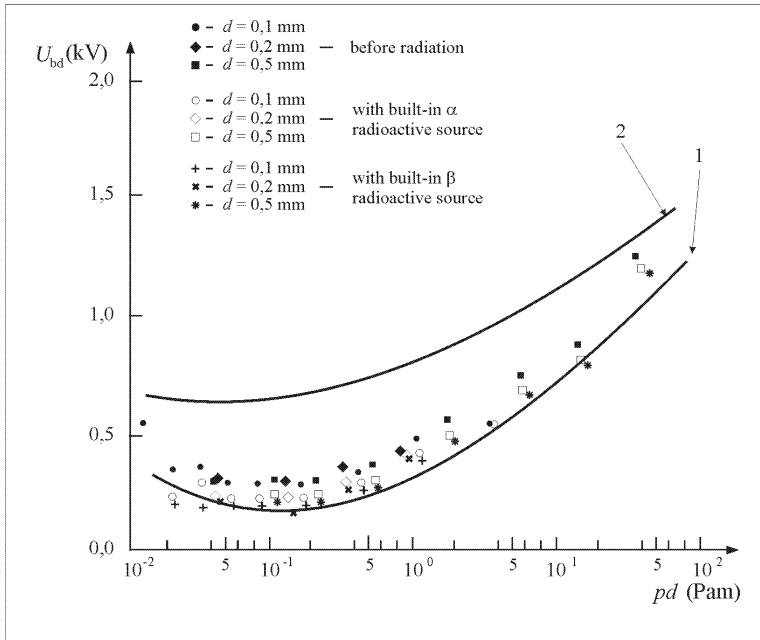
Fig. 12. The GFSA dc breakdown voltage versus variable pd before radiation, with built-in α and with built-in β radioactive source (p is Argon gas pressure in the chamber, d is interelectrode gap) Curve 1- Townsend breakdown mechanism, Curve 2- Streamer breakdown mechanism
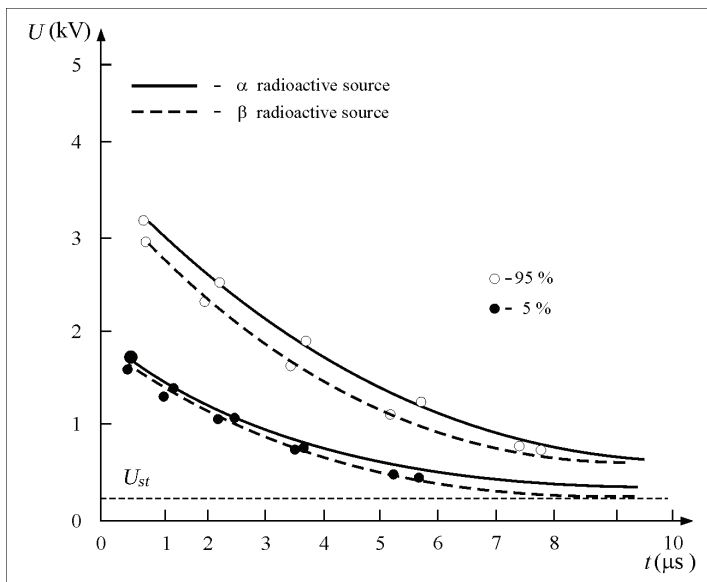
Fig. 13. The GFSA pulse shape characteristic with built-in α and β radioactive source

decrease of the response time by more than three times compared to that without radioactive sources (e.g. from 100 ns to 30 ns). At lower pressures (10^2 Pa – 10^3 Pa) it is more effective to use the α-emitters while at the higher pressures (10^5 Pa – 10^6 Pa) the β-emitters has proved to be more effective because of the greater cross section for ionization by β-particles (Osmokrović et al., 2005).

4. Conclusion

This text presents the results of the examination of programmable memories' radiation reliability. Influence of cobalt-60 gamma radiation was tested on NM27C512 8F85 EPROM and M24128 - B W BN 5 T P EEPROM components. EPROM components proved to have better radiation reliability than EEPROMs. Significant faults in EPROM and EEPROM components appeared at 1300 Gy and 1000 Gy, respectively. Changes in EPROMs are reversible, and after erasing and reprogramming, all EPROM components are functional. Reversibility of changes in EPROMs is attributed to partial light-induced annealing of trapped holes during UV erasure. Due to the cumulative radiation effects, first failures of the previously irradiated EPROMs appear at significantly lower doses. On the other hand, EEPROM changes are irreversible. All observed phenomena have a plausible theoretical explanation, based on the interaction of gamma radiation with the oxide layer of memory cell MOS transistors. The influence of gamma radiation is basically manifested through the change of the net gate surface charge density, and consequently of transistor threshold voltage.

For future work we planned the following:
1. To investigate these results for other EPROM and EEPROM components;
2. To include the gate insulator material and its thickness in research;
3. To consider the processing and doping methods used in securing the gate insulator onto the silicon surface;
4. To investigate the radiation-induced mobility changes;
5. To examine the dose rate effects;
6. To investigate the temperature influence of the memory during irradiation.

Also, in this chapter the examination of radiation resistance results of the components of over-voltage protections were presented. The influence of $n_+\gamma$ radiation was tested on the TSD, MOV, GFSA and Polycarbon Capacitors (the most sensitive elements of the electric filter).

It was concluded that under $n_+\gamma$ radiation the protection characteristics of TSD degrades considerably worse. A similar effect, only more pronounced was noticed in the case of the MOV. In the case of GFSA, the protection characteristics were improved under $n_+\gamma$ radiation. Polycarbon Capacitors exposed to $n_+\gamma$ radiation in a given fluence range, under goes a change in its capacitance towards smaller values.

These results show that TSD, MOV and Polycarbon Capacitors are radiation sensitive components. The GFSA shows a considerable improvement of the protection characteristics under $n_+\gamma$ radiation. Therefore by investigating the obtained results, it can be noticed that GFSA shows the best characteristics stability under radiation influence, among examined over - voltage protection components. The results obtained when GFSA is in a radiation environment, show that there is no need to use GFSA with built-in radioactive sources,

because the induced radiation causes the same or a very similar effect. This conclusion is very important from both the environment protection and the GFSA production point of view. Moreover, using short living radioactive isotopes, it is possible to produce faster GFSA. After a certain period of time, the GFSA would lose their characteristics, so their storage and disposal would not cause any more risk to the environment. We believe that future experiments should be directed towards the investigation of the possibility of a combined use of short lived radioactive isotopes and the hollow cathode effect.

It means that in conditions of increased hazardous radiation GFSA gives the best protection. Of course it doesn't mean that other over - voltage protection elements are not more reliable in some other hales (temperature (Lončar et al., 2002), electromagnetic contamination, aging (Osmokrović b et al., 2002), dimension, price). However, if several factors influence the devices simultaneously, hybrid protective scheme should be employed, which includes different types of over - voltage protection elements.

5. References

Clegg, D. W. & Collyer, A. A. (1991). *Irradiation Effects on Polymers*, Kluwer Academic Publishers, New York

Fleetwood, D. M. (1992). Border Traps in MOS Devices. *IEEE Trans. Nucl. Sci.*, Vol. 39, No. 2, 269–271

Gromov, V.V. & Evdokimov, O.B. (1984). *Trapping Charge Phenomena In Irradiated Dielectrics*, Pergamon Press, New York

Holmes – Siedle, A. & Adams, L. (2002). *Handbook of Radiation Effects*, Oxford University Press, Oxford

Hubbell, J. H. & Seltzer, S.M. (2004). *Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients* (version 1.4), National Institute of Standards and Technology, Gaithersburg, MD. Available Online: http://physics.nist.gov/xaamdi

Lončar B., Osmokrović P., Stojanović M., & Stanković S. (2001). Radioactive Reliability of Programmable Memories, *Jpn. J.Appl. Phys.*, Vol. 40 Pt. 1, No. 2B, 1126 –1129.

Lončar B., Osmokrović P. & Stanković S. (2002). Temperature Stability of Components for Over-voltage Protection of Low-voltage Systems, *IEEE Trans. Plasma Sci.*, Vol. 30, No. 5, 1881-1885.

Lončar B., Osmokrović P. & Stanković S. (2003). Radioactive Reliability of Gas Filled Surge Arresters, *IEEE Trans. Nucl. Sci.*, Vol. 50, No. 5, 1725–1731

Lončar B., Stanković S., Vasić A. & Osmokrović P. (2005). The Influence of Gamma and X-radiation on Pre-breakdown Currents and Resistance of Commercial Gas Filled Surge Arresters, *Nucl. Tech. & Rad. Prot.*, Vol. 20, No. 1, 59-63.

Lončar B., Osmokrović P.; Vujisić M. & Vasić A. (2007). Temperature and Radiation Hardness of Polycarbonate Capacitors, *J. Opt. & Adv. Mat.*, Vol. 9, No. 9, 2863-2866

Ma, T. P. & Dressendorfer, P.V. (1989). *Ionizing Radiation Effects in MOS Devices and Circuits*, John Wiley & Sons, New York

Mahan, C. D.; Levinson, L. M. & Philipp, H. R. (1989). Theory of Conduction in ZnO Varistors. *Journal of Applied Physics*, Vol. 50, No. 4, 424-428

Malinaric, P. (1985). Transient Suppressor Design with Varistor Composite Materials. *IEEE Trans. Electromagnetic Compatability*, Vol. 23, No. 4, 338-343

Markov, Z. (1987). *Over-voltage Protection in Electronics and Telecommunications*, Technical book, Belgrade, Yugoslavia

Messenger, G.C. & Ash, M.S. (1992). *The Effects of Radiation on Electronic Systems*, Van Nostrand Reinhold, New York

Osmokrović P.; Lončar B. & Stanković S. (2002). Investigation the Optimal Method for Improvement the Protective Characteristics of Gas Filled Surge Arresters-with/without the Built in Radioactive Sources, *IEEE Trans. Plasma Sci.*, Vol. 30, No. 5, 1876-1880.

Osmokrović P.; Lončar B., Stanković S. & Vasić A. (2002). Aging of the Over-voltage Protection Elements Caused by Over-Voltages, *Microel. Rel.*, Vol. 42, No. 12, 1959 1966.

Osmokrović P.; Lončar B. & Šašić R. (2005). Influence of the Electrode Parameters on Pulse Shape Characteristic of Gas-filled Surge Arresters at small Pressure and Inter-electrode Gap Values, *IEEE Trans. Plasma Sci.*, Vol. 33, No. 5, 1729-1735.

Osmokrović P.; Lončar B. & Stanković S. (2006). The New Method of Determining Characteristics of Elements for Overvoltage Protection of Low-voltage System, *IEEE Trans. Instr. & Meas.*, Vol. 55, No. 1, 257-265.

Prince, B. (1991). *Semiconductor Memories*, *A Handbook of Design, Manufacture and Applications*, John Wiley & Sons, New York

Raymond, J. P. (1985). IEEE Nuclear and Space Radiation Effects Short Course Notes, *Proceedings of Nuclear and Space Radiation Effects Conference,* pp. 5-15, New York, July 1985, New York

Ristić, G. S.; Pejović M. M. & Jaksić A. B., (1998). Modelling of Kinetics of Creation and Passivation of Interface Traps in Metal-oxide-semiconductor Transistors during Postirradiation Annealing, *J. Appl. Phys.*, Vol. 83, No. 6, 2994–3000

Ristić, G. S.; Pejović M. M. & Jakšić A. B., (2000). Analysis of Postirradiation Annealing of *n*-channel

Power Vertical Double-diffused Metal-oxide-semiconductor Transistors. *J. Appl. Phys.*, Vol. 87, No. 7, 3468–3477

Ristić, G. S.; Pejović M. M. & Jakšić A. B., (2007). Physico-chemical Processes in Metal-oxide-semiconductor Transistors with Thick Gate Oxide during High Electric Field Stress, *J. Non-Cryst. Sol.*, Vol. 353, 170–179

Snyder, E. S.; McWhorter P. J.; Dellin T.A. & Sweetman J. D. (1989). Radiation Response of Floating Gate EEPROM Memory Cells, *IEEE Trans. Nucl. Sci.*, Vol. 36, No. 6, 2131–2139

Srour, J.R. (1982). *Basic Mechanisms of Radiation Effects on Electronic Materials, Devices and Integrated Circuits,* DNA-TR-82-20, New York.

Van Lint, V. A. J.; Flanagan, T. M.; Leadon, R.E. & Rogers, C. V. (1980) *Mechanisms of Radiation Effects in Electronic Materials*, John Wiley & Sons, New York

Vujisić M.; Osmokrović P.; & Lončar B. (2007). Gamma irradiation effects in programmable read only memories, *J. Phys. D: Appl. Phys.*, Vol. 40, 5785–5789

Vujisić M.; Osmokrović P.; Stanković K.; & Lončar B. (2007). Influence of Working
Conditions on Over-voltage Diode Operations, *J. Opt. & Adv. Mat.*, Vol. 9, No. 12,
3881-3884

Wrobel, T. F. (1989). Radiation Characterization of a 28C256 EEPROM. *IEEE Trans. Nucl.
Sci.*, Vol. 36, No. 6, 2247–2251

ANN Application to Modelling of the D/A and A/D Interface for Mixed-mode Behavioural Simulation

Miona Andrejević Stošović and Vančo Litovski
University of Niš, Faculty of Electronic Engineering
Serbia

1. Introduction

The design of electronic and telecommunication integrated circuits is unavoidably faced with the simulation of analog subsystems of ever rising complexity as a result of building more and more complex mixed-signal systems containing both analog and digital parts. The design of such systems requires simulation tools that are both fast and accurate at the same time.

The main obstacle to this requirement is related to the difficulties in high-level modelling of the analog part, and sufficiently accurate modelling of the digital-analog (D/A) and analog-digital (A/D) interfaces which are frequently encountered in such systems (Trihy & Kundert, 1995; Kundert, 1999). We will consider here analog mixed-level behavioural modelling as a special case of the behavioural modelling of mixed-signal systems.

According to McAndrew: "It is difficult to develop complete, accurate, numerically robust models that work properly with numerical algorithms in simulators and with the CAD tools in an IC design system" (McAndrew et al., 1998). Modelling dynamic non-linear networks is practically the most difficult. One needs to solve not only the problem of the approximations to be used (modelling) and the method of evaluation of the coefficients within them (characterisation) but also to determine appropriate test signals able to activate the complete non-linear and dynamic properties of the circuit to be modelled.

Considering Fig. 1., we may have two distinct situations. First, for the A/D interface, we need the voltage and the current at the interface so that the input impedance of the digital part can be modelled. Modelling of the input impedance of a CMOS circuit is, to our knowledge, usually considered simple: a capacitor is used as a model. However, this capacitance is highly non-linear so that there is need for more realistic modelling as we propose in the next. Second, at the D/A interface, we need an electrical circuit to model the digital output driving the analog load. That model should perform D/A signal conversion and maintain electrical compatibility. Among the solutions available in the literature are an approach where time-varying resistors have been used (Petković & Litovski, 1989; Petković & Litovski, 1991), and a macromodelling procedure which has difficulties in characterising the reactive part of the non-linear output impedance of the model (Nichols et al., 1992; Brown et al., 1994). Recently, while attempting to enable high-level behavioural simulation of complex combinational CMOS circuits, successful techniques have been developed for

equivalent inverter or equivalent transistor synthesis represented as a non-linear voltage controlled current source (Chatzigeorgiou et al., 1999). Here an alternative, algorithmically simpler procedure, but robust and with broader scope of application, is proposed based on use of artificial neural networks (ANNs).

After the first application of ANNs for modelling electronic components, it became clear that this concept could be successfully used for modelling dynamics too (Litovski et al., 1992). The dynamics in a micro-electro-magneto-mechanical actuator were modelled first, taking advantage of the fact that the dynamic and the resistive properties may be expressed separately (Litovski et al., 1997a).
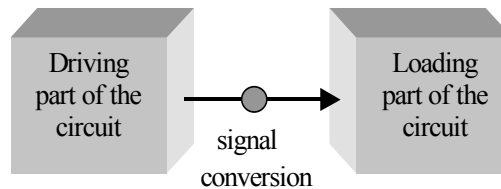
Fig. 1. Schematic depiction of the interface to be modelled between analog and digital parts of the circuit.

Full non-linear dynamic modelling using ANNs was described later – the approximation being performed in the frequency domain using direct and inverse Fourier transforms repeatedly (Citterio et al., 1999). Here we propose, for the first time, ANNs to be used for behavioural modelling of the digital output with modelling performed in the time domain.

The proceedings in the next will go in two directions. First we will describe the solution to the problem of A/D interface modelling. Since we assume that the digital (loading) part of the interface has no own feedback to its input, we will use its input impedance as the load to the analog part. In that way the problem of modelling of the A/D interface will be reduced to the problem of modelling the input impedance of the digital part. Second, the solution to the problem of the D/A interface will be proposed. In this case a complete new analog circuit will be proposed that will model the output behaviour of the digital driver at the interface. It will be shown by inspection that despite the fact that the model was developed for unloaded digital part, it behaves perfectly for every load used as a control.

2. Modelling the A/D interface

The simplest circuit representing input impedance of a MOS logic circuit is shown in Fig. 2. Here and on, for simplicity, the very conversion (analog to digital value) of the signal is not shown. The circuit is linear, with constant element values, what is not good enough approximation of the real MOS input impedance.

Input resistance and capacitance of bipolar circuits is nonlinear, so the simplest input TTL circuit is diode. Fig. 3. shows one solution with diodes (Corman & Wimborow, 1988).

The problem of an analog circuit output can be generally solved by connecting nonlinear resistance and nonlinear capacitance in parallel controlled by the voltage of the analog part of A/D interface, Fig. 4., (Petković & Stojanović, 1992). A special procedure is proposed in order to get the characteristics of the nonlinear elements depicted.
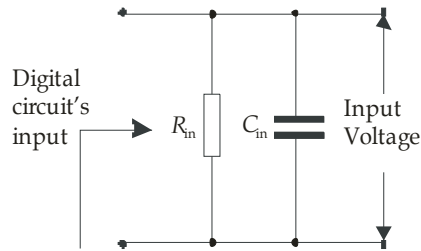
Fig. 2. RC input circuit
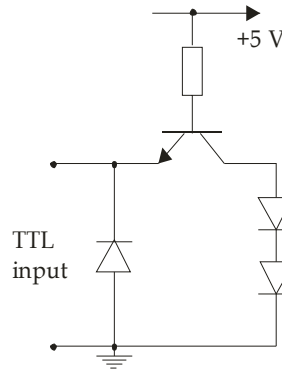
Fig. 3. Input circuit-one possible solution
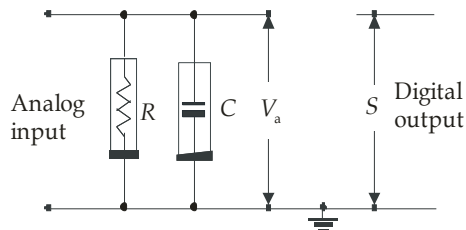
Fig. 4. A/D interface model

In the following (Litovski & Andrejević, 2002) we will show that MOS circuit input impedance is highly nonlinear, so more precise method of modelling is required. In order to get more information concerning our problem, as an example of two-terminal non-linear dynamic circuit, the input of an inverter is considered, as shown in Fig. 5. It is capacitive but highly non-linear, the non-linearity coming mostly from the Miller effect. As shown in Fig. 6., the inverter's gain is strongly non-linear, so that the Miller capacitances have to be too.

Inverter has only one input, so we are modelling only one input impedance. When there are more inputs, we have to model as many impedances as there are circuit inputs.

We attacked the problem of implementation of ANNs to analog modelling by first solving the simulation of an electro-magneto-mechanical actuator, as mentioned above. Thereafter, the problem of modelling linear dynamic networks was solved (Ilić et al., 2000). Two main ideas were implemented. First, to enable application of ANNs for implementation of the

model the so-called recurrent time-delay structure was used. It is already proven (Chow & Li, 2000) "that any finite time trajectory of a given *n*-dimensional dynamical continuous system with input can be approximated by the internal state of the output units of a continuous time recurrent neural network".
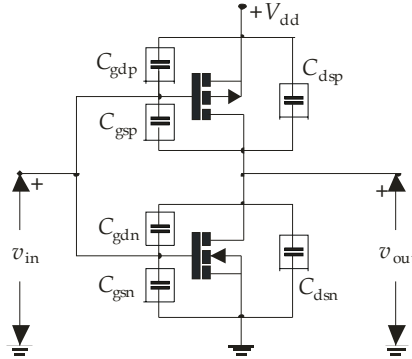
Fig. 5. CMOS inverter with non-linear capacitances.
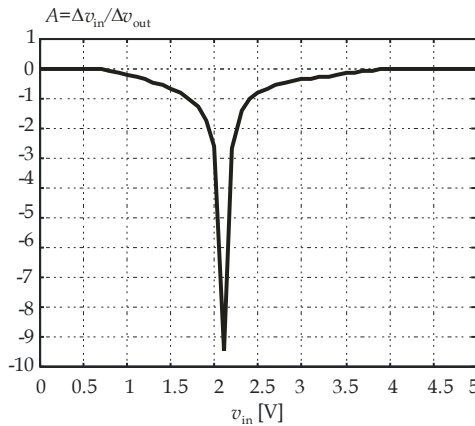
Fig. 6. The dependence of the inverter's gain on the input signal.

The structure of such a network is depicted in Fig. 7. The need for both time-delay and recurrence comes from the theory of dynamic system modelling (Bernieri et al., 1994). In our experience, however, time delay of the input signal facilitates modelling of the nonlinear properties while recurrence better accounts for the memory property of the system to be modelled. The core of the structure is a feed-forward ANN that, apart of the input layer, has one hidden layer having neurons with sigmoidal activation function, and output layer with a single linear output neuron. In the special case of modelling a two-terminal device, x in Fig. 7. stands for the exciting current while y denotes the response (voltage). The discretization time step is Δt .

Second, to capture the dynamic behaviour of the circuit to be modelled, a chirp signal (i.e., a frequency-modulated sinusoidal signal) was used as excitation. It has evenly-distributed spectral components across the "passband" of the circuit. If the change in frequency is

sufficiently slow, one gets the proper time-domain response of the dynamic system that is being modelled. This response is then used for training the ANN described above. The generalisation property of the ANN was verified by testing the models obtained as above with new signals, such as square pulses (Ilić et al., 2000). In order to model non-linear circuits with the concept described above, one needs amplitude large enough of the chirp signal to encompass all non-linearities encountered during excitation.
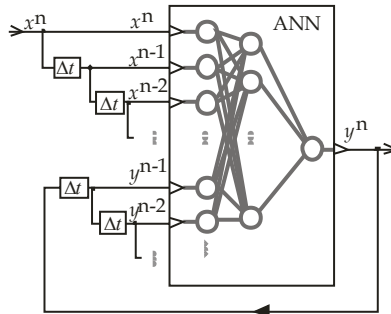
Fig. 7. The topology of the proposed ANN.
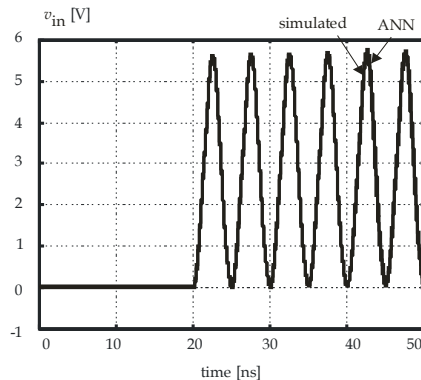
Fig. 8. Response (voltage) at the inverter's input obtained for sinusoidal excitation. Here "simulated" denotes the circuit simulation of the inverter of Fig. 5., while "ANN" stands for the response of the model.

The response observed was the input voltage of the inverter. The modelling results are shown in Fig. 8. where the difference between the simulated circuit and the model response is difficult to discern.

For the example used here, five inputs (three direct input connections and two feed-backs from the output as shown in Fig. 7.), four hidden units, and one output neuron are incorporated in the ANN. Note that the input units are simply distributing signals to the hidden layer. The hidden units have sigmoidal activation functions, while the output neuron is linear. Table 1. shows the weights and the thresholds of the neurons. A modified steepest-descent learning algorithm was used to train the ANN (Zografski, 1991). Other details of the training procedure are practically the same as described elsewhere (Bernieri et al., 1994).

No.	Hidden-layer neurons (First figure stands for the input neuron)	Output neuron (First figure stands for the hidden neuron)
1	$w^1(1,1) = -1.15618$ $w^1(2,1) = 0.303841$ $w^1(3,1) = 0.801999$ $w^1(4,1) = -0.655468$ $w^1(5,1) = -0.112933$ $\theta_1^1 = 1.69522$	$w^2(1,1) = -1.24339$ $w^2(2,1) = 1.41557$ $w^2(3,1) = -1.98177$ $w^2(4,1) = -2.18655$ $\theta_1^2 = 2.91066$
2	$w^1(1,2) = 0.138613$ $w^1(2,2) = -1.16098$ $w^1(3,2) = 1.01521$ $w^1(4,2) = -0.128065$ $w^1(5,2) = 0.915597$ $\theta_2^1 = -0.890606$	
3	$w^1(1,3) = -0.537022$ $w^1(2,3) = 1.64986$ $w^1(3,3) = -1.1164$ $w^1(4,3) = -2.64625$ $w^1(5,3) = 1.49825$ $\theta_3^1 = -1.15733$	
4	$w^1(1,4) = -1.13048$ $w^1(2,4) = 1.87005$ $w^1(3,4) = -0.716303$ $w^1(4,4) = -3.13684$ $w^1(5,4) = 2.4126$ $\theta_4^1 = 1.53233$	

Table 1. Weights and thresholds of ANN used to model the inverter in Fig. 5.

3. Generalisation capabilities

To show the quality of the approximation procedure and the generalisation capabilities of the ANN, a new example circuit is now considered. It consists of two inverters. The first (the driver) is considered to be analog as in Fig. 5. The second (the load) is considered to be digital. The whole is depicted in Fig. 9a. The voltage at the A/D interface (v_1) is of interest here. The excitation used is depicted in Fig. 10.

To get reference data for modelling, the logic part was substituted by its circuit equivalent as shown in Fig. 9b. SPICE-like simulation was performed and the results are depicted in Fig. 11 denoted as "simulated". Here the *Alecsis* simulator was used for all simulations (Litovski et al., 2001). Now, the logic branch presents effectively infinite impedance so that the second analog inverter is responsible for all loading effects.

Implementation of the model for mixed-signal simulation is depicted in Fig. 9c. Here the loading effects are modelled by the ANN. The simulation results are given in Fig. 11. denoted by "ANN". Note that for multi-input circuits separate models (tables similar to Table 1.) have to be created for every distinct input terminal. One needs a behavioural simulator to exercise such a model (Litovski et al., 2001). This concludes the discussion related to modelling of the A/D interface.
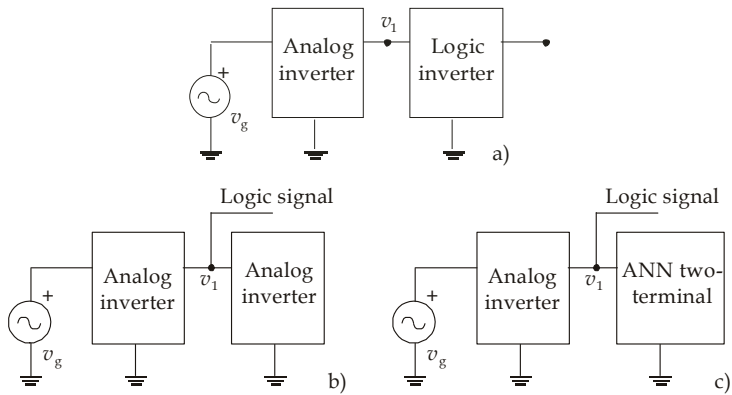
Fig. 9. Circuits used for verification. a) Analog inverter loaded by logic inverter, b) analog inverter loaded by analog inverter that generates reference signal v_1, and c) analog inverter loaded by ANN model.
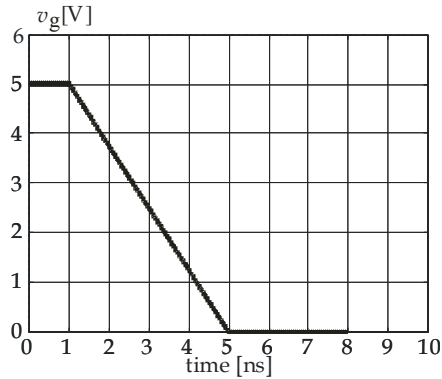
Fig. 10. Signal exciting the circuits in Fig. 9a.
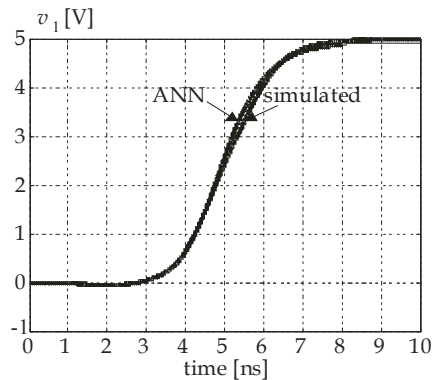
Fig. 11. Response (v_1) obtained by simulation of the circuit of Fig. 9b. (simulated) and the circuit of Fig. 9c. (ANN).

4. Modelling the D/A interface

For modelling of the D/A interface, the output circuit of the digital part is to be represented by a circuit that is supposed to drive an analog load. Note that mixed-mode simulation is considered. This means that an event scheduler is active, marking the controlling input of the digital circuit (Litovski & Zwolinski, 1997b). The event scheduler does not allow for two inputs to be active simultaneously because that is considered as a hazard. Hence, modelling the output of an inverter is general enough for verification of the modelling procedure.
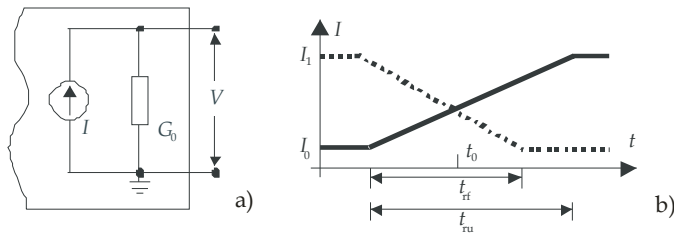
Fig. 12. a) Simple D/A conversion circuit, b) Current generator waveform t_{ru} stands for the rising edge while t_{rf} for the falling edge duration of the transition

Modelling of the D/A interface is more complex problem than modelling of the A/D interface, because we need to generate voltage waveform that excites the analog part of the circuit out of a set of logic states. Conversion algorithms are mostly based on synthesis of an electronic circuit that replaces the logic element's output, and is connected as an excitation to the particular node. Logic gate's delays also need to be considered and extracted by the event scheduler.

The simplest solution of the D/A conversion is illustrated in Fig. 12. (Zwolinski et al., 1989). There is a branch consisting of a constant conductance G_0 and current generator I, and it is applied to D/A node. The delay time is denoted by t_0.

Ratios I_1/G_0 and I_0/G_0 correspond to levels of logic 1 and logic 0, respectively, and different transition times from logic 1 to 0 and vice-versa, are permitted. Current waveforms for transitions from logic 1 to 0 and vice-versa are given in Fig. 12b.

A more complex output circuit is shown in Fig. 13. (Arnout & De Man, 1978). There are two voltage generators (E_0, R_0) and (E_1, R_1) applied to the analog node depending on logic element's output state. This function is realized by a switch controlled by Boolean function. R_0 and R_1 are logic gate's output resistances, when there are logic 0 or 1 at the output, respectively, meaning that there are two different resistance values, in contrast to previous case, when G_0 was used in both cases. The logic gate's delay is included in the switching time instant.
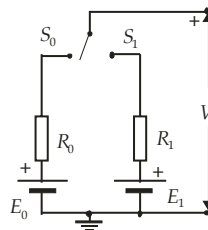
Fig. 13. D/A conversion with voltage levels

To further improve accuracy one may use the meliorated version depicted in Fig. 14. (Acuna et al., 1990). Sequence of pairs (E_i, R_i) and voltage controlled switch are used.
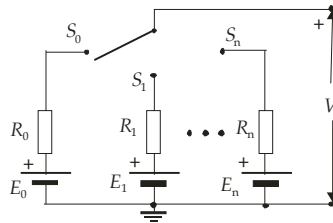
Fig. 14. Conversion when using several signal states on the logic gate output
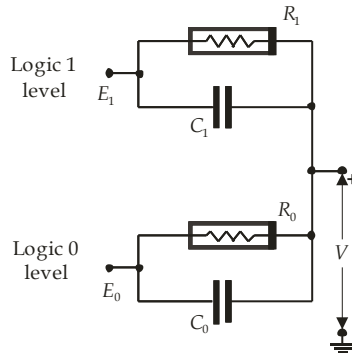
Fig. 15. D/A conversion using pair of voltage controlled resistors

In the circuit in Fig. 15. (Corman & Wimborow, 1988), the logic gate's output is observed as a voltage divider output. The capacitance values are constant and determined by the user if needed, and resistances are nonlinear and determined by user. In the circuits depicted in Figs. 14. and 15., the logic signal is firstly converted into electrical one and then values of this analog signal are discretized by comparing to sequence of thresholds. On that basis switches (Fig. 14.) or resistors values (Fig. 15.) are controlled. The circuits in previous examples approximate analog signal by discontinuous functions, what is inappropriate for most nonlinear circuit analysis methods.

Example of an output circuit approximated by analytical function is given in (Petković & Litovski, 1989; Petković & Litovski, 1991). Only nonlinear resistance is included, and using an approximation procedure, analytical expressions were produced expressing the output resistance dependence on the output voltage. Fig. 16. represents the output resistance of a CMOS inverter as a function of the output voltage. The dashed line is an approximation that was expressed in closed form.

The circuit depicted in Fig. 17. (Petković & Stojanović, 1992) includes output capacitances also. It consists of a nonlinear controlled ideal voltage generator E, a nonlinear resistor R and two output nonlinear capacitors C_0 and C_1. The transfer function, delays, output resistance and capacitances of digital gates are precisely modelled. While in (Petković & Litovski, 1989; Petković & Litovski, 1991) two constant values representing the logic level were used only, here the transfer characteristics and the delay are expressed in a more sophisticated way. Namely, a ramp signal, obtained by conversion of the logic output signal (similarly to Fig. 12b), is first delayed and as such, it represents a controlling signal for the

nonlinear generator E, whose dependence on the controlling voltage is actually the static transfer characteristic of an equivalent inverter. C_0 and C_1 are space-charge capacitances of the complementary transistors in the equivalent inverter.
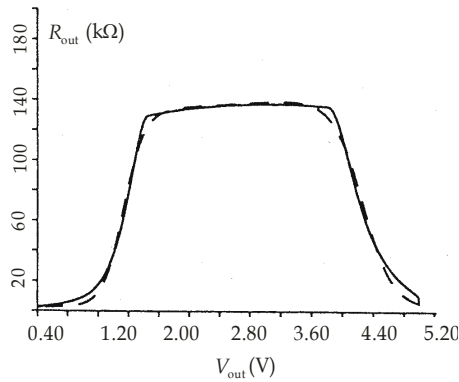
Fig. 16. Output resistance approximation
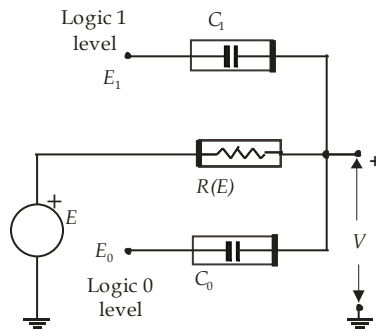
Fig. 17. D/A conversion using nonlinear reactive part
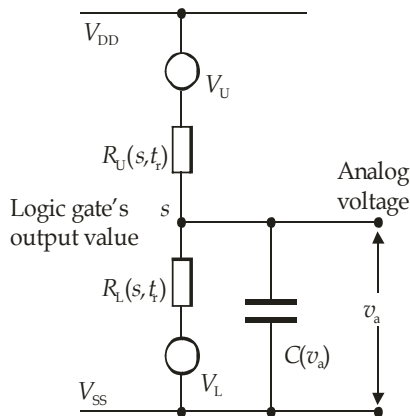
Fig. 18. Output impedance model

Time-dependent resistors are used in (Nichols et al., 1992). The model of the output impedance of a logic gate is shown in Fig. 18. The values of the resistances R_U and R_L depend on value s, as well as on transition time t_r, but not on the analog output voltage v_a. On the other side the capacitance C depends on this voltage. The voltages V_{DD} and V_{SS} are logic gate supply voltages. V_U and V_L are fixed offset values for the given type of logic gate. The resistances R_U and R_L linearly change their values from minimum to maximum and reversely, depending on time t_r. Linear change does not cause problems in analog simulation, because the analog voltage value is continuous. The parameter t_r is chosen large enough in order to hinder too fast analog voltage change, even if the capacitance value reaches zero. The capacitance change is given in (Nichols et al., 1992).

Similar solution where resistors change their values is given in (Brown et al., 1994).

In the next, solution based on artificial neural networks is given. Main property of this solution is its topological generality. Namely, we have no need to look for the topology of the model depending on the approximation procedure or on the topology of the digital original. Simply, the topology is always the same. In addition, the approximating function is general in the sense that only the parameters within the approximating function are mapping the properties of the instantiated digital circuit.

The topology of the new model is depicted in Fig. 19. In the figure, v_{in} stands for a controlling ramp-shaped voltage-waveform:

$$i(v_{in}) = I_{max}\left[1 - \tanh(v_{in} - v_T)\right],$$

(1)

and Z is a recurrent time-delay neural network approximating the function:

$$v_{out} = Z(i).$$

(2)
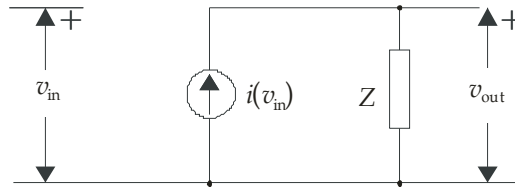
Fig. 19. Circuit representation of the model expressed by (1) and (2)

Here, I_{max} is the maximum supply current during the transition in the inverter, and v_T is (usually) equal to $V_{DD}/2$, V_{DD} being the supply voltage. Obviously, the ANN model of Z has one input (current) and one output (voltage) terminal. The network is trained using input-output pairs $[i(t), v_{out}(t)]$, where $i(t)$ is calculated from (1) while $v_{out}(t)$ is obtained by simulation using the *Alecsis* simulator of the circuit to be modelled (here an inverter). Note that we need the electrical schematic of the digital part during the modelling phase.

First results are shown in Fig. 20. Here the output waveforms of the original inverter and the model are shown to illustrate the quality of the approximation procedure. Unloaded circuits are simulated. The ANN has five input units (their role being the same as in Table 1.), three hidden units, and one output unit. Weights and thresholds are given in Table 2.

No.	Hidden-layer neurons (First figure stands for the input neuron)	Output neuron (First figure stands for the hidden neuron)
1	$w^1(1,1) = 2.28185$ $w^1(2,1) = -3.51137$ $w^1(3,1) = 1.36815$ $w^1(4,1) = 3.54312$ $w^1(5,1) = -1.37367$ $\theta_1^1 = -1.3177$	$w^2(1,1) = 0.644039$ $w^2(2,1) = 0.644042$ $w^2(3,1) = 0.644043$ $\theta_1^2 = -0.408248$
2	$w^1(1,2) = 2.28187$ $w^1(2,2) = -3.51135$ $w^1(3,2) = 1.36816$ $w^1(4,2) = 3.54312$ $w^1(5,2) = -1.37366$ $\theta_2^1 = -1.31769$	
3	$w^1(1,3) = 2.28187$ $w^1(2,3) = -3.51135$ $w^1(3,3) = 1.36816$ $w^1(4,3) = 3.54313$ $w^1(5,3) = -1.37366$ $\theta_3^1 = -1.31769$	

Table 2. Weights and thresholds of ANN used to model the inverter circuit.
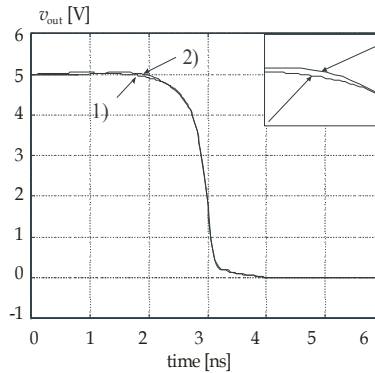
Fig. 20. Responses: 1) of unloaded CMOS inverter (considered as digital output) and 2) of the new model.

5. Further examples

The following three examples are intended to check the modelling procedure based on situations not present during training. The first trace (marked 1)) in Fig. 21. is the output voltage of an inverter being loaded by an inverter, all modelled by regular transistor models, i.e., obtained by regular circuit simulation. The second one (marked 2)) represents the

response of the same circuit with the ANN model used for the driving part and circuit model for the loading. This situation was not encountered in the training process. Excellent agreement was obtained, especially in the steepest part of the response that defines both the gain and the delay of the loaded inverter.

Further, Fig. 22. gives a similar comparison the loading element here being a transmission line modelled by a π-RC network (Chatzigeorgiou et al., 2001). Finally, a TTL load (diode) was used to demonstrate the success of the ANN model in the case of a 'large' non-linear dynamic load, Fig 23. Note the average value of the output voltage is less than 0.5 V while the difference is still smaller than 10 mV. Once again, the ANN model was developed using an unloaded inverter.
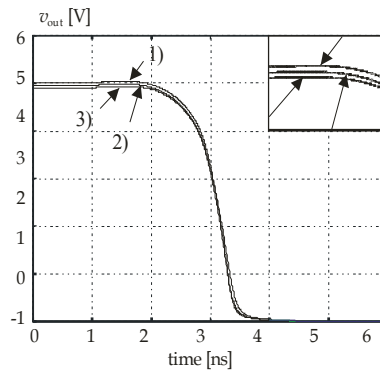
Fig. 21. Responses of 1) inverter loaded by inverter, 2) a model loaded by inverter, and 3) an ANN (modelling the output) loaded by an ANN modelling the input of an inverter.
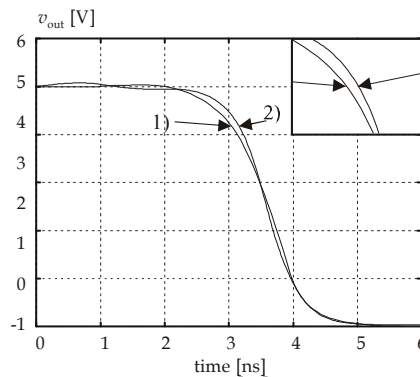
Fig. 22. Responses of 1) an inverter loaded by RC π-network and 2) a model loaded by RC π-network.

6. Application to high level analog simulation

Mixed-level analog behavioural modelling may need application of both concepts. In some situations, one will need to model the output circuit of the driver but in other cases, one will need to model the input circuit of the load; at very high levels of presentation, one will need

both. Such an example for the D/A interface is given in Fig. 21. Here trace 3) represents a response obtained by behavioural simulation using ANN models for both the driver and the load. In this way, the type of modelling we propose offers the opportunity to be implemented in analog behavioural simulation at any level.
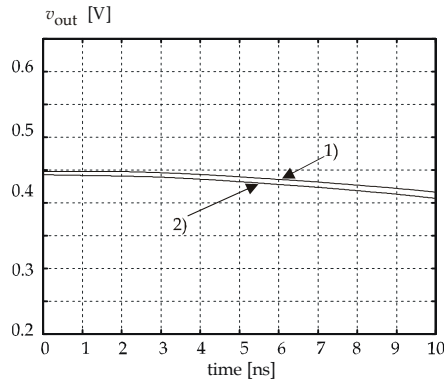
Fig. 23. Responses of a) inverter loaded by a diode and b) ANN model loaded by a diode.

7. Conclusion

An approach to the modelling of the A/D and D/A interface in mixed-mode circuit using ANNs has been described. The main difference in these two is the type of the input signal used for capturing the dynamic properties of the circuit to be modelled. For the D/A interface we use a ramp being, simply, the natural signal, while a sinusoidal signal was used for the input impedance modelling at the A/D interface in conjunction with general two-terminal non-linear dynamic modelling.

To summarise, a new method for modelling non-linear dynamic electronic circuits is described and applied to the modelling of A/D and D/A interfaces for mixed-signal simulation. It is *general and robust*. From the point of view of speed of simulation, one should bear in mind that ANNs are complex structures with exponential non-linearities requiring additional evaluation time compared to linear models. However, having in mind the complexity of modern models of MOS transistors (the BSIM3v3 model that is used in most modern electronic simulation capabilities needs more than a hundred parameters); we claim that the ANN approach is both efficient and convenient.

8. References

Acuna, E. L., Dervenis, J. P., Pagones, A. J., Yang, F. L., Saleh, R. A. (1990). Simulation techniques for mixed analog/digital circuits, *IEEE Journal of Solid-State Circuits*, Vol. 25, No. 2, pp. 353-363, ISSN 0018-9200.

Arnout, G., De Man, H. J. (1978). The use of threshold functions and Boolean-controlled network elements for macromodelling of LSI circuits, *IEEE Journal of Solid-State Circuits*, Vol. SC-13, No. 3, pp. 326-332, ISSN 0018-9200.

Bernieri, A., D'Apuzzo, M., Sansone, L. and Savastano, M. (1994). A Neural Network Approach for Identification and Fault Diagnosis on Dynamic Systems, *IEEE Trans. on Instrumentation and Measurements*, Vol. 43, pp. 867-873, ISSN 0018-9456.

Brown, A. D., Nichols, K. G., Zwolinski, M. and Kazmierski, T. J. (1994). CLASS Simulator Comparable Mixed-Mode Interfacing, *1994 Research Journal, Department of ECS, University of Southampton*, pp. 99-101, England.

Chatzigeorgiou, A., Nikolaidis, S. and Tsukalas, I. (1999). A Modelling Technique for CMOS Gates, *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 18, pp. 557-575, ISSN 0278-0070.

Chatzigeorgiou, A., Nikolaidis, S. and Tsukalas, I. (2001). Modelling CMOS Gates Driving RC Interconnect Loads, *IEEE Transactions on Circuits and Systems–II: Analog and Digital Signal Processing*, Vol. 48, pp. 413-418, ISSN 1057-7130.

Chow, T. S. W., and Li, X.-D. (2000). Modelling of Continuous Time Dynamical Systems with Input by recurrent Neural Networks, *IEEE Transactions on CAS–I: Fundamental Theory and Applications*, Vol. 47, pp. 575-578, ISSN 1057-7122.

Citterio, C., Pelagotti, A., Piuri, V. and Rocca, L. (1999). Function Approximation – A Fast-Convergence Neural Approach Based on Spectral Analysis, *IEEE Transactions on Neural Networks*, Vol. 10, pp. 725-740, ISSN 1045-9227.

Corman, T., Wimborow, M. U. (1988). Coupling a digital logic simulator and an analog circuit simulator, *VLSI System Design*, pp. 40-47.

Ilić, T., Zarković, K., Litovski, V. B., and Mrčarica, Ž. (2000). ANN Application in Modelling of Dynamic Linear Circuits, *Proceedings of the Small Systems Simulation Symposium, SSSS'2000*, pp. 43-47, Niš, Yugoslavia, September 2000.

Kundert, K. S. (1999). Introduction to RF Simulation and Its Application. *IEEE Journal of Solid-State Circuits*, Vol. 34, pp. 1298-1318, ISSN 0018-9200.

Litovski, V. B., Radjenović, J., Mrčarica, Ž. and Milenković, S. (1992). MOS Transistor Modelling Using Neural Network, *Electronics Letters*, Vol. 28, pp. 1766-1768.

Litovski, V. B., Mrčarica, Ž., and Ilić, T. (1997a). Simulation of Non-linear Magnetic Circuits Modelled Using Artificial Neural Network, *Simulation Practice and Theory*, Vol. 4, pp. 553-570, ISSN 0928-4869.

Litovski, V., and Zwolinski, M. (1997b). *VLSI Circuit Simulation and Optimization*, Chapman and Hall, ISBN 0412638606.

Litovski, V., Maksimović, D., and Mrčarica, Ž. (2001). Mixed-Signal Modelling With AleC++: Specific Features of the HDL, *Simulation Practice and Theory*, Vol. 8, pp. 433-449, ISSN 0928-4869.

Litovski, V., Andrejević, M. (2002). ANN application in Modelling of A/D interfaces for mixed-mode behavioural simulation, *Proceedings of XLVI Conference of ETRAN*, pp. I51-I54, Banja Vrućica, Bosnia & Herzegovina, June 2002.

Litovski, V., Andrejević, M., Damper, R. (2003). Modelling the D/A Interface for Mixed-Mode Behavioural Simulation, *Proceedings of EUROCON 2003*, pp. A.130-A.133, September 2003, Ljubljana, Slovenia.

Litovski, V., Andrejević, M., Petković, P., Damper, R. (2004). ANN Application to Modelling of the D/A and A/D Interface for Mixed-Mode Behavioural Simulation, *Journal of Circuits, Systems and Computers*, Vol. 13, No. 1, pp. 181-192, ISSN 0218-1266.

McAndrew, C. C. (1998). Practical Modelling for Circuit Simulation, *IEEE Journal of Solid State Circuits*, Vol. 33, pp. 439-448, ISSN 0018-9200.

Nichols, K. G., Brown, A. D., Zwolinski, M., and Kazmierski, T. J. (1992). A Logic-Analog Interface Model, *1992 Research Journal, Department of ECS, University of Southampton*, pp. 106-109, England.

Petković, P., and Litovski, V. (1989). Time Domain Black-box Modelling of CMOS Structures and Analog Timing Simulation, *Proceedings of the Third Annual European Computer Conference, COMPEURO'89*, pp. 5.142-5.143, Hamburg, Germany.

Petković, P., and Litovski, V. (1991). Output Resistance of CMOS Logic Cells, *Proceedings of the 3rd Mid-European Conference on Custom/ASICS, CCC1991*, pp. 237-244, Sopron, Hungary.

Petković, P., Stojanović, Z. (1992). Primena analognih makromodela logičkih ćelija u modeliranju D/A sprege kod hibridnog simulatora, *Proceedings of XXXVI Yugoslav Conference of ETAN*, pp. 51-57, Kopaonik, Yugoslavia.

Trihy, R., and Kundert, K. (1995). Top Down Design with VHDL-A, *Proceedings EURO-SIM'95-Session Software Tools and Products*, pp. 53-56, ISBN 0444822410, Vienna, Austria, September 1995, IEEE Computer Society Press.

Zografski, Z. (1991). A Novel Machine Learning Algorithm and its Use in Modelling and Simulation of Dynamical Systems, *Proceedings of Fifth Annual European Computer Conference, COMPEURO'91,* pp. 860-864, Bologna, Italy.

Zwolinski, M. et al. (1989). The "HOMICIDES" mixed-mode circuit simulator, *Proceedings of the Silicon Design Conference,* Heathrow, England.

Electronic Circuits Diagnosis using Artificial Neural Networks

Miona Andrejević Stošović and Vančo Litovski
University of Niš, Faculty of Electronic Engineering
Serbia

1. Introduction

Whenever we think about why something does not behave as it should, we are starting the process of diagnosis. Diagnosis is therefore a common activity in our everyday lives (Benjamins & Jansweijer, 1990). Every complex system is liable to faults or failures. In the most general terms, a fault is every change in a system that prevents it from operating in the proper manner. We define diagnosis as the task of identifying the cause and location of a fault manifested by some observed behaviour. This is often considered to be a two-stage process: first the fact that fault has occurred must be recognized – this is referred to as *fault detection*. That is, in general, achieved by testing. Secondly, the nature and location should be determined such that appropriate remedial action may be initiated.

The explosion of integrated circuit technology has brought with it some difficult testing problems. The recent growth of mixed analogue and digital circuits complicates the testing problem even further. It becomes more complicated to determine a set of input test signals and output measurements that will provide a high degree of fault coverage. There is also a timing problem of testing the circuits even on the fastest automated equipment.

The general structure of a diagnostic system is shown in Fig. 1. Signals $u(t)$ and $y(t)$ are input and output to the system, respectively. Faults and disturbances (here measurement errors) also influence the system under test, here denoted as the "Process", but there is no information about the values of these errors. The task of the diagnostic system is to generate a diagnostic statement S, which contains information about fault modes that can explain the behaviour of the Process. Note that the diagnostic system is assumed to be passive i.e. it cannot affect the Process itself.

The whole diagnostic system can be divided into smaller parts referred here to as tests. These tests are also diagnostic systems, DS_i. It is assumed that each of them generates diagnostic statement S_i. The purpose of the decision logic (voting system) is then to combine this information in order to form the final diagnostic statement S.

The number of possible faults in an electronic system may be large and can be located everywhere in the system. To diagnose in such conditions one frequently uses hierarchical approach where successive diagnostic statements are generated as the level of description of the system is lowered going down towards the fault itself (Ho et al., 2001; Sheu & Chang, 1997). This allows for smaller sets of faults to be considered at a time for the given hierarchical level. Modern automatic test pattern generator may support such concepts (Soma et al., 2001).

2. Concepts of diagnosis

Besides the human expert that is performing the diagnosis, one needs tools that will help, and ideally, perform the diagnosis automatically. Such tools are a great challenge to design engineers because, usually, the diagnostic problem is underspecified. In addition, it is a deductive process with one set of data creating, in general, unlimited number of hypotheses among which we try to find a solution. This is why the research community continues to be attracted by this problem (Bandler & Salama, 1985).

During the life-cycle of a product, testing is implemented in both the production phase and the implementation phase. We claim, however, that the sustainability of a product is strongly influenced by the design phase. So, to make a sustainable product, one should design the test procedure and synthesize test signals early in the design phase.

It is frequently possible to perform functional verification of the system. That, most frequently, happens when a small number of input/output terminals is present. In the majority of cases however, full functional testing becomes time consuming and is not acceptable. So, one applies defect-oriented (structural) testing, as will be discussed in more detail in what follows.

We consider testing to be: the selection of a set of defects regarded as the most probable, the description of a set of measurements, the selection of a set of testing points (or output signals) and most importantly, the synthesis of optimal testing signals that will be applied at the system inputs allowing for detectability and observability of the listed fault effects. Here, optimality means that one test signal covers as many faults as possible.

Selection of the type of measurements and testing points is specific to the circuit. One should stick to those measurements that are prescribed for functional verification. Specific measurements such as supply current monitoring are frequently adopted, too. Separate test points may be added in order to improve detectability or observability. Specific design for testability concepts can be applied.

Thanks to the advances in computational intelligence in the last decades new diagnostic paradigms have been applied based on: model-based concepts (Benjamins & Jansweijer, 1990); production rule based artificial intelligence (Pipitone et al., 1991); ANNs (Hayashi et al., 2002); genetic algorithms (Golonek & Rutkowski, 2002); and fuzzy-reasoning (Pous, et al., 2002); all trying to create an approach that contains properties that we might consider to be "intelligent behaviour".

In order to get an idea of why and how ANNs are applied to analogue electronic circuit diagnosis, the diagnostic concept (Fig. 1) will be elaborated in some detail first. It involves collaboration of design, test, and field engineers and the mutual distribution of responsibilities throughout the life cycle of an electronic product. We assume that field engineers are expected to react after a functional failure of the system. In order to diagnose such a system they need to be supplied with: testing equipment, a list of specific measurements to be done (including a set of signals and test points), and diagnostic software to process the measurement data. A similar set of data and tools would be given to a test engineer in a production-plant environment in order to evaluate the production yield and create feedback to process engineers when prototyping the circuit. We believe, however, design engineers are the most familiar with the product and the most qualified and capable to synthesize test and diagnostic signals, and procedures. This means the SBT (simulation before test) has to be applied to create fault dictionaries containing exhaustive lists of faults and corresponding responses. The fault dictionary is in fact a table
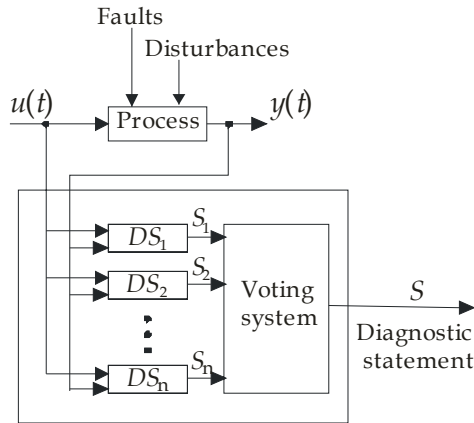
Fig. 1. A general diagnostic system.

representing the mapping from the fault list into a list of faulty (or possibly, fault-free) responses. In that way the diagnostic process becomes a search through the fault dictionary. Alternatively, modern diagnostic techniques using traditional artificial intelligence and reasoning methods typically fall into the simulation after test (SAT) category. This will increase the time spent on diagnosing systems at production time (Spina & Upadhyaya, 1997). SBT systems typically require more initial computational costs, but provide faster diagnosis at production time being the second reason why this concept was accepted here.

We claim here that ANNs, being universal approximators (Scarselli & Tsoi, 1997), are the best way both to capture the mapping, and to search through the dictionary, thereby to perform diagnosis.

3. Diagnosis of nonlinear dynamic analogue circuits

Analogue electronic circuits are known to be difficult to test and diagnose. Apart from the huge number of possible faults, this difficulty is a consequence of the inherent nonlinearity of these circuits. Even linear circuits (having linear input-output signal interdependence) exhibit nonlinear relations between circuit parameters and the output response. There are no linear active networks. Active networks are nonlinear with nonlinear reactive elements. They may be linearized and thought of as such in situations where signal and parameter changes are small in comparison to nominal values. When large parameter changes or even catastrophic faults occur (affecting the DC state), however, one must distinguish between linear and analogue circuits. This is not the case in most research reports bringing confusion into the subject.

Several concepts were applied to diagnosis of analogue networks. Among them we will first mention the ones relying on reasoning based on measured data and some measure of distance between the response of the good circuit and the faulty one. Starting with the basic research reported in (Bandler & Salama, 1985) and (Milor & Visvanathan, 1989) several ideas were reported. In (Luchetta et al., 2002) the fault location phase is considered as an optimization problem where the parameter value is searched for in order to minimize the difference among the actual and simulated response. Linear circuits in the frequency domain are considered being characterized by symbolic functions. Similarly, in (Catelani &

Giraldi, 1998) applying SAT multiple faults may be diagnosed in linear circuits described by symbolic functions what is characterized as model based method. SBT based method for soft faults diagnosis in linear circuit was proposed in (Alippi et al., 2002) where harmonic analysis was used for selecting the most suitable test input stimuli and nodes by means of global sensitivity approach. In (Huang & Cheng, 2000) and (Yoon et al., 1998) passive circuits were diagnosed based on graph theoretical approach, and on pass and fail regions for the circuit poles and zeroes in the real-imaginary plane, respectively, while in (Chang, 2002) a Boolean decision scheme was proposed for the diagnosis of linear circuits described in the frequency domain. In order to diagnose multiple soft faults in the same type of circuits the Woodbury formula was applied to the modified nodal equation to construct the so called fault equation in (Liu & Starzyk, 2002). A decomposition method was proposed in (Starzyk & Liu, 2002) aiming to cope with circuit complexity. In one approach, small parameter changes were allowed in nonlinear circuits (Tadeusuewicz et al., 2002). Soft faults were considered only when linear programming method was used for diagnostic decisions. Large parametric fault diagnosis was described in (Worsman & Wong, 2002) using piecewise linear models for DC analysis, and separate considerations were given for diagnosis of faults in the dynamic part of the network (considered linear) based on large change sensitivity computations. Further, in (Cota et al., 1999) the diagnostic method applied consists of injecting probable faults in a mathematical model of the linear circuit, and later comparing its output with the output of the real faulty circuit. Transfer functions transformed into the Z domain were created and fault injection was performed. In (Cherubal & Chatterjee, 1999) methodology based on linear regression model using prior circuit simulation which relates a set of measurements to the circuit's internal parameters was applied in order to solve for the circuit parameter values using iterative numerical techniques. Linear circuits in the frequency domain were diagnosed in (El-Yazeed & Mohsen, 2003) where the AC response to a set of sinusoidal input frequencies was calculated at selected test nodes. Prony's method was then utilized as a preprocessor to extract an optimal set of features representing nodal voltage waveforms. In (Dai & Xu, 1999) a solution to the same problem was proposed based on noise measurements.

Soft and hard faults (shorts and opens) in nonlinear dynamic circuit were diagnosed in (Pinjala et al., 2003). The procedure employs a statistical method of computing Mahalanobis distance to find defects in load board traces and components. Short list of defects was reported. A low-noise amplifier was diagnosed in (Liobe & Margala, 2004) by using digital signatures suitable for built-in self test design concepts. Hard and soft faults were diagnosed the former modelled as resistors having convenient values.

A specific aspect of diagnosis is the number and location of the test points. Simply, we can say that internal test points should be avoided and measurements on the primary inputs and outputs are preferred. This is not only related to their automatic accessibility but also to the nature of the diagnostic reasoning. Namely, one looks for functionality to diagnose something, and the function is seen at the primary terminals. Of course, in order to compensate for the small number of test points more measurements with different types of applied signals are, generally, needed to extract complete information about the system behaviour. For complex analogue systems, however, hierarchical approaches based on decomposition (Ho et al., 2001; Sheu & Chang, 1997; Bandler & Salama, 1985; Starzyk & Liu, 2002) are inevitable provided that no propagation of the fault effect arises between partitions what is not easy to achieve. Of course, there are circuits that may be partitioned based on functionality known *a priori* from the design process as mentioned in the introduction.

Another aspect of fundamental importance is related to the choice of the output quantities that are to be measured. In most cases these are voltages at the output of the circuit under test (CUT) or at selected test points. It is shown, however, that measurement of the supply current (Iddq) may be successfully used for testing of both analogue and digital circuit (Dragic & Margala, 2002; Margala et al., 2002; Papakostas & Hatzopoulos, 1991; Bell et al., 1991; Zwolinski et al., 1996). This idea was used for diagnosis of analogue circuits using ANN that will be discussed later.

Several results were reported where the so called artificial intelligence concepts were applied to diagnosis of analogue circuits or at least linear ones. In (Savioli et al., 2005) method based on fault trajectory concept for fault diagnosis of analog linear continuous time networks, which relies on evolutionary techniques, where a genetic algorithm (GA) was coded to optimize test vector generation, was reported. GA was applied into (Golonek & Rutkowski, 2002) creation "transfer functions" enabling creation of a new type of fault dictionary. The classical signature dictionary has been replaced by fault decoder based on transfer functions. In order to obtain a sharp diagnosis about the possible wrong component of the circuit, a tool based on qualitative reasoning was used in (Pous et al., 2002). In particular, the results were refined by means of fuzzy techniques. This means that inputs, outputs, rules and the corresponding operators to combine them were defined. A production rule based concept was reported in (Pipitone et al., 1991).

ANNs have previously been applied to diagnosis (Spina & Upadhyaya, 1997; Materka, 1994; Rodrigez et al., 1994; Aminian & Aminian, 2000; He et al., 2002; Andrejević & Litovski, 2004; Aminian et al., 2002; Stopjakova et al., 2004; Yu et al., 1994; Collins et al., 1994; Catelani & Gori, 1996; Maidon et al., 1997; Yang et al., 2000). As in the case with the classical concepts, however, ANNs were predominantly applied to linear analogue circuits. In (Materka, 1994) feed-forward ANNs were used for parameter identification (soft fault diagnosis) of linear circuits. In (Rodrigez et al., 1994) linear power networks were diagnosed by feed-forward ANNs. In order to enhance the performance of the ANN applied for diagnosing of soft faults in linear active networks, in (Spina & Upadhyaya, 1997), new "criteria" - a discriminating measure based on discrepancy of the autocorrelation function of the fault-free and the correlation function of the faulty and fault-free circuit, were introduced. The same problem was attacked in (Aminian & Aminian, 2000; Aminian et al., 2002) where the impulse response was analyzed by wavelet decomposition, principal component analysis, and data normalization preprocessors before introduced to the ANN. Soft faults were considered only. In (He et al., 2002) a method based on extraction of a "feature vector" from the differences between vectors of node voltages of faulty and fault-free linear circuit for every fault was described. This feature vector is then presented to the ANN as a teaching session. Network tearing is applied in order to manage the circuit complexity in an 11 transistor bipolar circuit. Every partition was considered linear although catastrophic faults were present (e.g. transistor base disconnected). Two faults were diagnosed only. In (Andrejević & Litovski, 2004) a linear resistive circuit was diagnosed using feed-forward neural nets. Soft and hard faults (shorts and opens) were considered. Comparably large set of faults was taken into account. In the scheme presented in (Catelani & Gori, 1996) (one opamp/one capacitor, three resistors and two diodes) programmable function generator was used to generate the set of stimuli sequentially injected into the input of the CUT. Six test frequencies were chosen. For each stimulus the frequency response of the CUT has been considered and five Fourier components were measured at the output test point with the spectrum analyzer. For the purpose of diagnosis, four neural networks were used. Euclidian

distance was to be learned by the ANN in order conclusions to be created on the origin of the fault. Bipolar analogue integrated circuits (Maidon et al., 1997) were diagnosed and their resistances determined from the magnitudes of the Fourier harmonics in the spectrum responses to a sinusoidal input test signal using multilayer perceptron ANN. The input vector to the ANN consists of the magnitudes of the Fourier harmonics of the response waveforms owing to the input stimulus, and the class represents the type of circuit faults, while the outputs map to resistance values of the faults. Probabilistic neural network was applied in (Yang et al., 2000). It is a four layer feedforward neural network that realizes the Bayes classifier. The ANN creates the probability that a circuit is faulty and points to the type of fault. In (Stopjakova et al., 2004) a large number of circuit versions was created by introducing sets of models for every separate fault. In fact, hard faults were considered while the opens and the shorts were modelled by resistors of variable resistivity. Then statistical properties of the time domain response (in this case the supply current) to a pulse excitation were extracted in order to create knowledge of the fault to fault-effect mapping. The supply current was successfully used for diagnosing gate oxide shorts in CMOS circuits by the help of ANNs in (Yu et al., 1994; Collins et al., 1994). After introducing a fault model of the MOS transistor built as a series connection of two MOS transistors with a common gate (i.e. considering this as a soft fault), several faults per transistor (for all transistors in an 11 transistor operational amplifier) were created by changing the possible position of the gate short relative to the source-to-drain ends of the channel. Sinusoidal and ramp signals were used for creation of a fault dictionary in an SAT method. The response i.e. the supply current was sampled to give a series of values used to train the feed-forward (Yu et al., 1994) and a Cohonen (Collins et al., 1994) neural network.

In this chapter we will give two examples of fault diagnosis in non-linear dynamic circuit. The first one refers to an analogue circuit, and the second to the mixed-mode circuit.

We describe the results of applying feed-forward ANNs to the diagnosis of non-linear dynamic electronic circuits with no restriction on the number and type of faults. This method is based on fault dictionary creation and using an ANN for data compression by memorizing the table representing the fault dictionary. Only DC and small signal sinusoidal excitations will be applied, so preserving the usual measurement procedure for generating the data given in a component's and/or a circuit's data-sheets. The ANN so created is, consequently, used for diagnosis by applying to it the signals obtained by measuring the faulty network. This process may be considered as looking-up a fault in the fault dictionary. The ANN finds the most probable *fault code* that corresponds to the measured signals.

Putting this in the general context of diagnosis we first note that the fault dictionary contains all the knowledge we need. In other words by applying the SBT concept all hypotheses are memorized (within the ANN) and no further hypothesis needs to be created after the dictionary is known. This is equivalent to the structural concept of testing. The fault not conceived in advance can't be tested nor diagnosed. Now we look among the hypotheses (by searching the dictionary i.e. by running the ANN) to find the one most similar to the actual (faulty) circuit response. The difficulties here are the complexity of the search and the decision algorithm that finds the "most similar" entry in the dictionary. As will be shown with an example this can be an extremely difficult task. It has been successfully solved using ANNs.

The network used for the first diagnostic example is a feed-forward neural network structured in three layers. It has only one hidden layer, which has been proved sufficient for

this kind of problem (Masters et al., 1993). The neurons in the hidden layer are activated by a sigmoidal function, while the neurons in the output layer are activated by a linear function. The learning algorithm used for training this network is a version of the steepest-descent minimization algorithm (Zografski et al., 1991).

4. Fault dictionary creation and application example

In order to describe the way in which the fault dictionary was created, the circuit in Fig. 2 is used as an application example. This is a CMOS operational amplifier consisting of seven transistors. To our knowledge this example belongs to the category of the most complex ones reported, both from the number of circuit elements point of view and the number of faults inserted. Note that three (nonlinear) capacitors are associated with every transistor totalling the number of nonlinear circuit elements to 28 but, for the sake of simplicity, are not shown in the figure. In order to emphasize the method as such, while not offering a full solution of the diagnostic problem for this circuit, having in mind abundance of possible faults, a reduced set of faults was considered. To this end only single transistor faults are sought. That, of course will not affect the generality of the ideas implemented in the next. We do not intend to diagnose simultaneous presence of several faults.
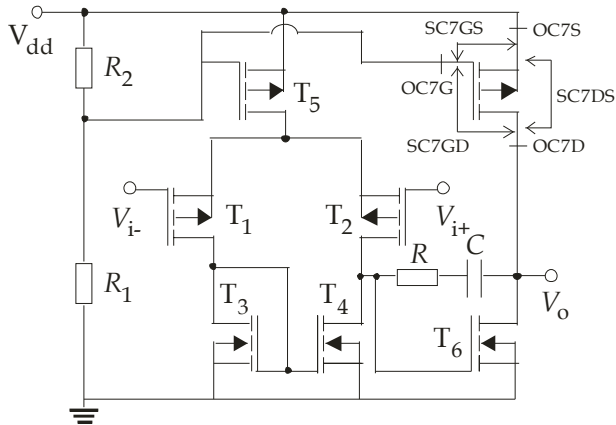
Fig. 2. The operational amplifier circuit. SC=short circuit, OC=open circuit

Ten faults per transistor, six catastrophic and four parametric were added to the dictionary. As shown in the figure (using T_7 as an example) there exist three open-circuit faults (OC) and three short-circuit faults (SC) per transistor (for example, OC3G stands for open gate of transistor T_3, and SC1DG stands for drain and gate shorted in transistor T_1, Table 1.). As opposed to (Stopjakova et al., 2004) and some others, the shorts (some of them behaving as bridging fault) and opens were really implemented instead of resistors modelling them. To effectively simulate perfect short and opens we used our model of the ideal switch (Mrčarica et al., 1999) what is not possible in the SPICE simulator. Of course, there was no obstacle for us to use resistors to model shorts and opens. Simply, what we did, we considered satisfactory. In addition, two faulty values for every channel length (±20%) (denoted as L+ and L- in Table 1.), and two for every channel width (±20%) (denoted as W+ and W- in Table 1.) were introduced, totalling 10 faults per transistor. The soft faults considered here are

expected to model design errors and, in a specific way, gate oxide short having in mind the fault model reported in (Yu et al., 1994). For the whole circuit this gives a set of 70 faults observed.

The DC output values ($V_{\text{oDC }m}$) were first obtained by simulation. Here $m=0,1,2,\ldots,69$ stands for the fault code. In addition, the frequency response of the circuit (the non-inverting input terminal was excited by a signal of amplitude 1mV) was obtained by simulation over a fixed frequency range in order to extract two response parameters: the nominal gain (A_m) and the 3-dB cut-off frequency ($f_{3\text{dB}m}$). For the example given, we considered this signature to be satisfactory complex. If additional fault need to be used one might think on additional measurements such as supply current. Note that, for the DC supply current point of view, the fault effects of most open faults at sources and drains in series connected transistors, may have equivalent signatures.

Type	A_m	$f_{3\text{dB}m}$ [MHz]	$V_{\text{oDC}m}$ [V]	Code (m)
FF	419	0.01527	0.127	0
1L+	0.0053	6.791	0.0497	37
OC1G	0.047	501.187	0.127	49
OC3G	0.049	544.042	0.093	47
SC1DG	0.042	320.440	0.0458	6
SC2DS	0.071	312.071	3.3	27
SC5DS	0.656	0.57	0.0186	55
6W-	5770	0.0018	0.2146	13
OC5D	0.056	507.298	3.3	25
SC5GS	0.109	0.036	0	2

Table 1. Part of the fault dictionary for the circuit of Fig. 2. The faults are chosen at random.

Because of the nonlinearity of the circuit, every fault is expected to change the transistor's quiescent points. Consequently, new linear transistor-models are created by SPICE-like program and used for frequency domain performance extraction for each fault. In order to find the new quiescent point for every fault, we have to insert the fault i.e. to create a faulty model of the circuit for DC analysis. This procedure is described elsewhere (Milovanović & Litovski, 1991; Milovanović & Litovski, 1994) and will not be discussed here.

Fault SC3DG is untestable because of the existing connection between the gate and drain of T_3. This reduces the fault dictionary to 69 elements. Therefore, the fault dictionary created here has four columns containing the set of circuit performances i.e. the signatures and the fault code: {$V_{\text{oDC}m}$, A_m, $f_{3\text{dB}m}$, m}. First three items in a row are considered inputs to the neural network, while the fault code is learnt as an output.

The fault coding is an important issue. In fact, some defects exhibit very similar effects. So, input data (signatures) can have very close numerical values, and if the output values (defect codes) were also similar, the network could not always be trained successfully. Such an example is given in Fig. 3. Here the signatures of three faults are compared. By careful inspection we can see that only the $f_{3\text{dB}}$ values suggest a difference between the fault effects. Faults are coded randomly, so that faults with similar effects are unlikely to have similar codes. This approach is proven to be good, because the way of coding influenced the training time, and also, the training error. Part of the fault dictionary for the circuit in Fig. 2, is given in Table 1., where $m=0$ denotes the fault-free circuit.
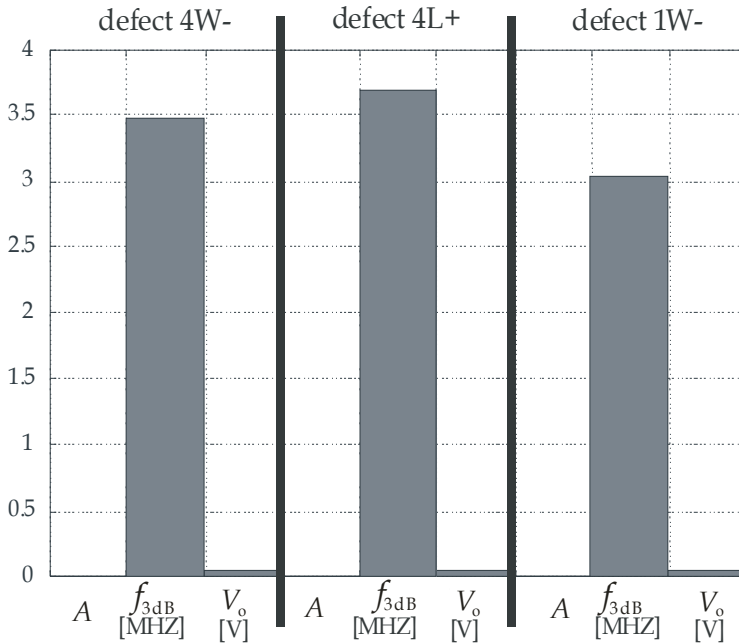
Fig. 3. Fault effects of faults 4W-, 4L+ and 1W-

Here we come to an additional issue concerning the applicability of rule-based approaches to the diagnosis of systems of this kind. Because of the similarity of the circuit performance in the presence of different faults, no set of rules can be established to distinguish between these three faults. Furthermore, by inspection of Table 1., we can see that the performance values cover a broad range. For example, for the voltage gain, the smallest value in the Table is 0.0053, and the largest is 5770. So, we cannot establish a rule defining the difference that occurs as a consequence of the presence of a certain fault. Thus rule-based approaches are impractical for systems exhibiting responses as continuous functions. Note, in addition, that we expect noisy data to be obtained when field measurements are performed for diagnosis which further complicates the creation of any rules. We may claim that similar is related to use fuzzyfication in order to boost the difference among signatures. To go further, this puts in a similar prospective the simulation-after-test i.e. the model based concept. Namely, in this concept one is supposed to create a set of hypotheses that will be checked against the measurement data by successive simulations of the circuit under test at the repairing site. Having in mind Fig. 3., however, we do not believe the creation of a qualified set of hypotheses is an achievable task.

The fault dictionary can be further reduced in size by processing the *ambiguity groups* or the groups of equivalent faults. According to (Manetti & Piccirilli, 2003) "an ambiguity group is, essentially, a group of components where, in case of fault, it is not possible to uniquely identify the faulty one". Here, we can say that an ambiguity group consists of a set of *faults* that propagate identical signatures to the output, making the faults detectable and the circuit testable, but no distinction between the individual faults is possible making them undiagnosable. Table 2. shows all ten ambiguity groups for this example, systematically

collected after simulation. The faults italicized in Table 2. represent the same topological connection in the circuit, so the effect would be expected to be the same.

Ambiguity group	Faults included	A	f_{3dB} [MHz]	V_{oDC} [V]
1	OC1D	0.31	20000	0.0179
	OC1S			
2	OC3D	0.041	365.8	3.3
	OC3S			
3	OC4D	0.303	20000	0.0458
	OC4S			
	SC4GS			
	SC3GS			
	SC3DS			
4	OC5D	0.056	507.298	3.3
	OC5S			
5	OC6D	0.063	0.039	3.3
	OC6S			
6	OC7D	$A \to \infty$	Indeterminate	0
	OC7S			
7	SC1GS	0.055	515.993	3.3
	SC2GS			
8	SC5GS	0.109	0.036	0
	SC7GS			
9	SC4DS	$A=0$	Indeterminate	3.3
	SC6GS			
	SC7DS			
10	3L+	0.05	2.37	3.3
	4W+			

Table 2. Ambiguity groups and fault effects.

A specific ambiguity group is the case when the gain (*A*) although small ever rises within the given frequency range. This is denoted as $A \to \infty$. The 3-dB cut-off frequency is in these cases indeterminate. Some of defects exhibiting this property, however, have different V_{oDC} values, so they can be distinguished. In these cases, to avoid use of infinite numbers during the training of the ANN we assigned a value of 1000 to the gain. This is the case when simulating defects OC7S and OC7D, but since these defects produce completely the same effect, they form ambiguity group number 6.

A similar situation occurs when the gain is almost zero, *A*=0. The 3-dB cut-off frequency is then again indeterminate. Ambiguity group number 9 covers three such defects, with the same V_{oDC}.

Only one representative of each ambiguity group was included in the fault dictionary. From Table 2., we find that the complete fault dictionary in this case has 70-1-24+10=55 elements.

With three pieces of data for each fault, the neural network input structure was restricted to three input terminals. The ANN diagnoses the fault by outputting the fault-code (*m*) as a signal level, so we needed only one output neuron. The number of hidden neurons, *n*, was

found by trial and error after several iterations starting with an estimation based on that in (Baum & Haussler, 1989). The goal was to find the optimum n that leads to a satisfactory classification even with noisy excitations. Using too many neurons would increase the training time, but using too few would starve the network of the resources needed to solve the problem. Also, an excessive number of hidden neurons may cause the *overfitting* problem (Masters et al., 1993), when a network has so much information capability that it learns insignificant aspects of the training sets, irrelevant to the general population. In practice, 30 hidden neurons were used. After successful training, no mistakes were observed for all 55 faults.

Code	A_j	$f_{3\mathrm{dB}j}$ [MHz]	$V_{\mathrm{oDC}j}$ [V]	ANN response
0	419	**0.0145**	0.127	-0.02128
1	129.6	**0.0248**	0.079	1.09057
2	0.109	0.036	**-0.05**	2.01405
3	6028	**0.001575**	0.1712	2.93868
5	4453	**0.002415**	1.0255	5.03203
6	**0.0441**	320.44	0.0458	6.03224
9	1000	1000	-0.05	9.0707
10	**0.043**	365.8	3.3	10.0278
12	1000	1000	**3.39**	12.1771
13	5770	**0.00171**	0.2146	13.2376
16	8220	**0.00197**	0.4876	16.031
18	0.32	1000	**0.133**	17.8458
20	0	1000	**3.46**	20.4409
21	0.83	1000	**3.46**	20.6497
25	**0.0588**	507.298	3.3	25.0605
26	11.739	**0.114**	0.127	26.0098
27	0.071	312.071	**3.46**	27.0091
34	5809	**0.00169**	0.1811	33.7541
35	209	0.0237	**0.115**	35.47
36	**0.05**	1000	0.8824	36.3514
37	**0.00556**	6.791	0.0497	37.2652
43	**0.004**	17.191	0.0509	43.0008
46	**0.0523**	515.993	3.3	45.99
47	**0.0514**	544.042	0.093	47.0133
49	**0.04935**	501.19	0.127	49.042
50	6030	**0.001425**	0.2466	49.9284
52	0.005	133.757	**3.46**	52.0044
53	119.4	**0.0258**	0.0843	53.0205
54	0.041	**428**	3.3	53.5346
55	0.688	0.57	0.0186	54.8614

Table 3. Inputs with noise and ANN responses.

The generalization property of the network was verified by supplying noisy data to its inputs. This is presented in Table 3. 30 samples were examined. For each sample, one input

(boldfaced in the Table 3.) is incremented by +5% or -5%, representing noise generated during the measurement process. The responses of the network are given in the last column of the table. The ANN response was considered to be correct (i.e. acceptable) when its value was in the range [(m-0.5), (m+0.5)]. We can see that all faults can still be diagnosed though some with difficulties (for m=20, 35, and 54).

In the previous text we have presented our first results in the development of a new technique for fault diagnosis of nonlinear dynamic circuits. The method we proposed may be summarised as follows.

In applying ANNs to the diagnosis of nonlinear dynamic electronic circuits, as described, we have demonstrated the implementation of the method and a set of results. These results vindicate the technique. In further work we intend to resolve the elements of ambiguity groups. In addition, more complex systems will be considered and larger fault dictionaries generated. Consequently additional measurements will be needed in order to keep the number of test points low. This will, hopefully, allow for implementation of these ideas to diagnosis of mixed-signal circuits.

5. Fault diagnosis in digital part of sigma-delta converter

Further in this text we will show that feed-forward ANN may be applied to the diagnosis of non-linear dynamic electronic circuits (Andrejević & Litovski, 2006a; Andrejević et al., 2006; Andrejević, 2006) that are mixed with digital ones. Two types of defects in the digital part of the circuit will be considered: effects of rising and falling edge delays in logic gates and catastrophic defects that change the circuit topology. Similar procedure may be applied to diagnosis in analog part of the circuit (Andrejević & Litovski, 2006b).

The simulation before test concept was adopted. This means that after choosing the set of faults of interest (say the most probable ones), repetitive simulation is performed in order to create the system response for every fault. Codes are associated to the responses and used as part of the fault dictionary that, in addition, contains the faulty responses themselves. Of course, the responses are represented in a form that is easy to manipulate.

The ANN is first trained for modelling the look-up table. This means that faulty responses are repeatedly brought to the input, while the ANN is forced to present the fault codes at its output. Then, the ANN running with the given vector of stimuli (measured output signals of a faulty or, possibly, fault free system) may be viewed as search of the look-up table. The ANN response, if the network properly trained, will immediately find the fault and produce the fault code at its output.

The procedure applied is reminiscent to the one implemented to analog circuits in (Litovski et al., 2006). To our knowledge this is the first application of ANNs to diagnosis of mixed signal circuit.

6. Faults in the specific circuit design

As an example of a complex circuit, the sigma-delta modulator in Fig. 4. is chosen (Xu & Lucas, 1995).

This is a mixed-signal circuit, having both analogue and digital elements. Switches in the circuit are modelled as truly ideal switches, with zero resistance for closed switch and infinite resistance for open switch. Simulations are performed using *Alecsis* (Glozić, 1994) simulator.
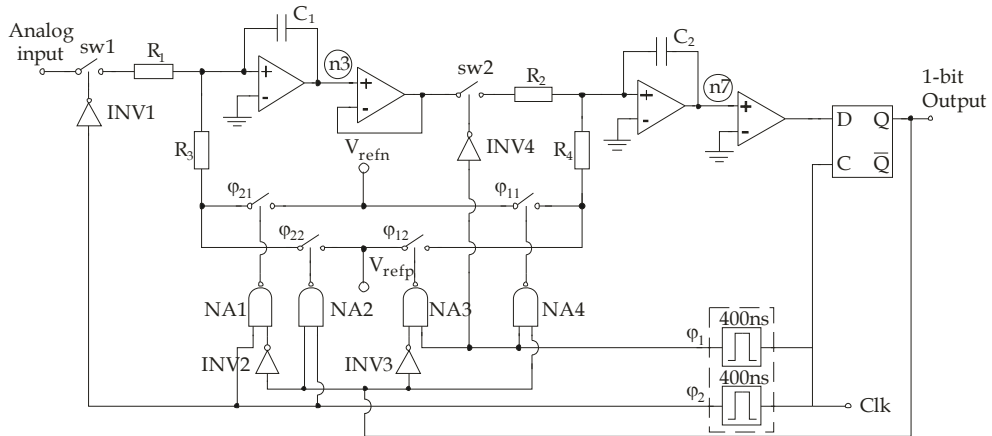
Fig. 4. Sigma-delta modulator architecture.

The integrator charging time is invariable with respect to clock rate in order to keep the gain constant. This means that the analog switch must be turned on for fixed time duration regardless of clock rate. This is achieved by using monostable multivibrator as a fixed-width pulse generator in the circuit.

The monostable multivibrator between the clock input and switch control block functions as a pulse generator to produce control signals of fixed time duration. Fig. 5 shows reaction of the system when the input is excited by a ramp signal.

We consider in this chapter defects only in the digital part of the circuit. There are two types of defects observed: catastrophic defects and delays of rising and falling edge of output digital signals. These delay defects are neither catastrophic, nor parametric, because there is no change in circuit topology, and no change in element values.

Digital signal can be "stuck-at-1" or "stuck-at-0". In the circuit in Fig. 4., analogue switches are controlled by digital signals, so there are pairs of the same fault effects, such as: the effect is the same when the switch is stuck at ON (OFF) and the logic circuit's output is "stuck-at-1" ("stuck-at-0"). So, we will consider hard faults (which refer to the analogue part of the circuit) as stuck switches (Andrejević et al., 2006). The cases when switches in the feedback loop (φ_{11}, φ_{12}, φ_{21}, φ_{22}) are permanently closed are excluded, because voltage references V_{refp} and V_{refn} would be shorted in such cases.

Having in mind that clock period in the circuit is 1.2μs (half period is 600ns), we examined effects of delays not greater than 400ns. In fact, effects of rising edge delay are simulated for values of delay: 100ns, 250ns, 400ns, and for falling edge, we simulated smaller values: 50ns, 100ns, 150ns. The goal was to determine how these delays influence the output, and whether different delay values produce different outputs (Andrejević et al., 2006). All digital gates are examined (4 inverters and 4 nand circuits). The first conclusion was that delays in the circuit of inverter 2 (INV2) do not influence output signal, meaning that output is not changed. Further, there exist groups of delays causing the same effect. Such groups are known as ambiguity groups, and they are listed in Table 4. The first four groups show the same effect of delays. In the second column of the Table 4, defects causing the same effects are named, and accordingly, third column presents that same effect (signature). The fifth ambiguity group is in a way different. The members of that group are both catastrophic and delay defects. Note that only one representative of each group is given in the fault dictionary, Table 5.

Ambiguity group	Defect type	Signature
1	na3(tf=50ns)	104108210
	na4(tr=250ns)	
2	na3(tr=400ns)	102104208
	na4(tr=400ns)	
3	na4(tf=100ns)	404210240
	inv3(tf=100ns)	
4	FF	20440480A
	inv2(tr=50ns)	
	inv2(tr=100ns)	
	inv2(tr=150ns)	
	inv2(tf=50ns)	
	inv2(tf=100ns)	
	inv2(tf=150ns)	
5	φ21OFF	000000000
	na1(tf=150ns)	
	sw1ON	

Table 4. Ambiguity groups.
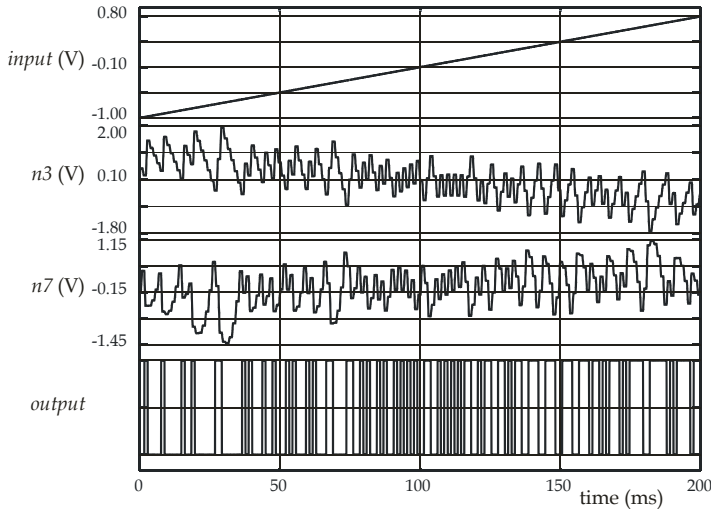
Fig. 5. Simulation results for ramp excitation

Fault dictionary is created using the response of the circuit to an input ramp signal. The circuit output value is registered after every clock period, so these output digital values form the output signature. These are then represented in more compact hexadecimal presentation. Accordingly, fault dictionary is created as shown in Table 5. It must be noted that defects are coded randomly, while it is very important that defects with similar signatures must not have similar fault codes. If this happens, it may be very difficult, or even impossible for ANN to recognize defects. In the second column of Table 5., defects are coded. First column describes the type of the defect, relative to notation given in Fig. 4. (*inv3*(*tf*=50ns) stands for the falling edge delay in inverter 3 and *na1*(*tr*=400ns) for the rising

edge in nand 1). FF stands for the fault free circuit. The third column contains the signature seen at the output.

Defect type	Defect code	Signature	Defect type	Defect code	Signature
FF	0	20440480A	inv4(tf=100ns)	23	220220821
sw1OFF	1	996999999	sw2ON	24	018018030
inv1(tr=150ns)	2	000010010	inv3(tr=150ns)	25	104208210
na1(tr=400ns)	3	31C352C66	inv4(tr=100ns)	26	050050110
inv1(tf=50ns)	4	811105024	na3(tr=250ns)	27	021041084
na1(tf=100ns)	5	008000010	inv4(tr=150ns)	28	0440900C0
na2(tf=50ns)	6	404811044	na1(tr=100ns)	29	844889112
na3(tr=400ns)	7	102104208	na3(tf=100ns)	30	082202208
na3(tf=50ns)	8	104108210	inv4(tf=50ns)	31	208210420
na3(tf=150ns)	9	030018090	na4(tf=100ns)	32	404210240
na4(tr=100ns)	10	204110210	sw2OFF	33	996696699
φ21OFF	11	000000000	inv1(tf=150ns)	34	092430918
inv1(tf=100ns)	12	848504890	na2(tf=150ns)	35	811104844
na2(tr=100ns)	13	102081042	inv1(tr=50ns)	36	040810108
na2(tf=100ns)	14	410842209	inv3(tf=150ns)	37	802408420
inv3(tr=100ns)	15	202404410	inv4(tf=150ns)	38	010840882
na3(tf=100ns)	16	808809021	na2(tr=250ns)	39	080408104
inv1(tr=100ns)	17	004020040	na1(tr=250ns)	40	149463131
na2(tr=400ns)	18	020101008	inv3(tf=50ns)	41	402804411
inv4(tr=50ns)	19	088108210	φ12OFF	42	300038003
na1(tf=50ns)	20	100110012	φ22OFF	43	925129252
na4(tf=50ns)	21	402804420	inv3(tr=50ns)	44	204208410
φ11OFF	22	001C00038	na4(tf=150ns)	45	802408811

Table 5. Fault dictionary.

ANN was trained for modelling the look-up table. It is a feed-forward neural network with one hidden layer. The signatures are inputs to the network, and the fault code is network output to be learned. It means that the neural network has 9 inputs (one input per hexadecimal digit) and one output neuron. Hexadecimal values are presented as decimal when they are inputs to the network. After learning was completed, the number of hidden neurons in the resulting ANN was 10, what was found by trial and error after several iterations starting with an estimation based on (Masters et al., 1993; Baum & Haussler, 1989). The structure of the obtained ANN is verified by exciting the ANN with faulty inputs. Responses of the ANN show that there were no errors in identifying the faults what is presented in Table 6. Only negligible discrepancies may be observed.

7. Conclusion

In applying ANNs to the diagnosis of nonlinear dynamic electronic circuits, as described, we have demonstrated the implementation of the method and a set of results.
In the second part of this chapter, effects of delay and catastrophic defects in sigma-delta modulator were examined. The diagnosis was successful.

Accordingly, we may conclude that ANNs are convenient and powerful means for diagnosis, and, what is important, realizable as a hardware that may be as fast as necessary to follow the changes of the system's response in real time.

Defect type	Defect code	ANN output	Defect type	Defect code	ANN output
FF	0	-0.000215	inv4(tf=100ns)	23	22.9998
sw1OFF	1	0.999861	sw2ON	24	23.9997
inv1(tr=150ns)	2	1.99969	inv3(tr=150ns)	25	24.9999
na1(tr=400ns)	3	2.99981	inv4(tr=100ns)	26	26.0009
inv1(tf=50ns)	4	3.99985	na3(tr=250ns)	27	27
na1(tf=100ns)	5	4.99988	inv4(tr=150ns)	28	28
na2(tf=50ns)	6	6.00003	na1(tr=100ns)	29	29.0001
na3(tr=400ns)	7	7.00006	na3(tf=100ns)	30	30
na3(tf=50ns)	8	7.99998	inv4(tf=50ns)	31	31.001
na3(tf=150ns)	9	8.99991	na4(tf=100ns)	32	32.0003
na4(tr=100ns)	10	10.0004	sw2OFF	33	33.0021
φ21OFF	11	10.9998	inv1(tf=150ns)	34	34
inv1(tf=100ns)	12	11.9997	na2(tf=150ns)	35	34.9998
na2(tr=100ns)	13	12.9994	inv1(tr=50ns)	36	36
na2(tf=100ns)	14	13.9997	inv3(tf=150ns)	37	37.0001
inv3(tr=100ns)	15	15	inv4(tf=150ns)	38	37.9998
na3(tr=100ns)	16	15.9996	na2(tr=250ns)	39	39.0024
inv1(tr=100ns)	17	16.9998	na1(tr=250ns)	40	40.0015
na2(tr=400ns)	18	18	inv3(tf=50ns)	41	40.9997
inv4(tr=50ns)	19	19.0018	φ12OFF	42	41.9996
na1(tf=50ns)	20	19.9997	φ22OFF	43	42.9998
na4(tf=50ns)	21	20.9999	inv3(tr=50ns)	44	43.9997
φ11OFF	22	22	na4(tf=150ns)	45	44.9996

Table 6. ANN output results.

8. References

Alippi, C., Catelani, M., Mugnaini, M. (2002). SBT Soft Fault Diagnosis in Analog Electronic Circuits: A Sensitivity-Based Approach by Randomized Algorithms, *IEEE Transactions on Instrumentation and measurement*, Vol. 51, No. 5, pp. 1116-1125.

Aminian, M., and Aminian, F. (2000). Neural-network based analog-circuit fault diagnosis using wavelet transform as preprocessor, *IEEE Transactions on CAS – II: Analog and Digital Signal Processing*, Vol. 47, No. 2, February 2000, pp. 151-156.

Aminian, F., Aminiam, M., Collins, H. W. (2002). Analog Fault Diagnosis of Actual Circuits Using Neural Networks, *IEEE Trans. On Instrumentation and Measurement*, Vol. 51, No. 3, June 2002, pp. 544-50, ISSN 0018-9456.

Andrejević, M., Litovski, V. (2004). ANN application in electronic diagnosis-preliminary results, *Proceedings of IEEE 24th International Conference on Microelectronics MIEL 2004*, pp. 597-600, Niš, Serbia, May 2004.

Andrejević, M., Litovski, V., Zwolinski, M. (2006). Fault Diagnosis in Digital Part of Mixed-Mode Circuit, *Proceedings of IEEE 25th International Conference on Microelectronics MIEL 2006*, pp. 437-440, ISBN 1-4244-0116-X, Niš, Serbia, May 2006.

Andrejević, M., Litovski, V. (2006a). Fault Diagnosis in Digital Part of Sigma-Delta Converter, *Proceedings of Neurel 2006 Conference*, pp. 177-180, ISBN 1-4244-0432-0, Beograd, Serbia, September 2006.

Andrejević, M., Litovski, V. (2006b). Fault Diagnosis in Analog Part of Mixed-mode Circuit, *VI Symposium on Industrial Electronics (INDEL 2006)*, pp. 117-120, Banja Luka, Bosnia and Herzegovina, November 2006.

Andrejević, M., Petrović, V., Mirković, D., Litovski, V. (2006). Delay Defects Diagnosis Using ANNs, *Proceedings of L Conference of ETRAN*, pp. 27-30., Belgrade, Serbia, June 2006.

Andrejević, M. (2006). Artificial neural networks application in electronic circuits diagnosis, *PhD Thesis*, University of Niš, Serbia, July 2006, (in Serbian).

Bandler, J., and Salama, A. (1985). Fault diagnosis of analog circuits, *Proceedings of the IEEE*, Vol. 73, No. 8, pp. 1279-1325, ISSN 0018-9219.

Baum, E. B., and Haussler, D. (1989). What size net gives valid generalization, *Neural Computing*, Vol. 1, pp. 151-60.

Bell, I. M., Camplin, D. A., Taylor, G. E., Bannister, B. R. (1991). Supply Current Testing Of Mixed Analogue And Digital ICs, *Electronics letters*, Vol. 27, No. 17, pp. 1591-1583, ISSN 0013-5194.

Benjamins, R., Jansweijer, W. (1990). Toward a competence theory of diagnosis, *IEEE Expert*, Vol. 9, No. 5, pp. 43-52, ISSN 0885-9000.

Catelani, M. and Gori, M. (1996). On the application of neural networks to fault diagnosis of electronic analog circuits, *Measurement*, Vol. 17, pp. 73-80.

Catelani, M., Giraldi, S. (1998). Fault diagnosis of analog circuits with model based techniques, *IEEE Instrum. Meas. Techn. Conference*, Vol. V 1, pp. 501-504.

Chang, Y.-H. (2002). Frequency-domain grouping robust fault diagnosis for analog circuits with uncertainties, *International Journal of Circuit Theory and Applications*, Vol. 30, pp. 65-86, ISSN 0098-9886.

Cherubal, S., Chatterjee, A. (1999). Parametric Fault Diagnosis for Analog Systems Using Functional Mapping, *Proceedings of Design, Automation and Test in Europe (DATE '99)*, p. 195, Munich, Germany.

Collins, P., Yu, S., Eckersaal, K. R., Jervis, B. W., Bell., I. M., and Taylor, G. E. (1994). Application of Cohonen and Supervised Forced Organization Maps to Fault Diagnosis in CMOS Opamps, *Electronics letters*, Vol. 30, No. 22, pp. 1846-1847, ISSN 0013-5194.

Cota, É. F., Carro, L., Lubaszewski, M. (1999). A Method to Diagnose Faults in Linear Analog Circuits Using an Adaptive Tester, *Proceedings of Design, Automation and Test in Europe Conference, DATE '99*, pp. 184-188, Munich, Germany.

Dai, Y., and Xu, J. (1999). Analog circuit fault diagnosis based on noise measurements, *Microelectronics and Reliability*, Vol. 39(8), pp. 1293-1298, ISSN: 0026-2714.

Dragic, S., and Margala, M. (2002). A 1.2 V Built-in Architecture for High Frequency On-Line Iddq/delta Iddq Test, *Proc. IEEE Computer Society Annual Symposium on VLSI*, pp. 148-153, April 2002, Pittsburgh, PA, USA.

El-Yazeed, M. F. Abu, Mohsen, A. A. K. (2003). A Preprocessor for Analog Circuit Fault Diagnosis Based on Prony's Method, *International Journal Electron. Commun.*, No. 1, pp. 16-22.

Glozić, D., Alecsis 2.1: An object-oriented hybrid simulator, *PhD Thesis*, University of Niš, Serbia, 1994, (in Serbian).

Golonek T., Rutkowski J. (2002). Use of Genetic Programming to Analog Fault Decoder Design, *Proceedings of the International Conference ICSES '2002*, Wrocław-Świeradów Zdrój.

Hayashi, S., Asakura, T., and Zhang, S. (2002). Study of Machine Fault Diagnosis System Using Neural Networks, *Proceedings of the International Joint Conference on Neural Networks*, pp. 233-238, Honolulu, Hawaii, May 2002.

He, Y.-G., Tan, Y.-H., and Sun, Y. (2002). A neural network approach for fault diagnosis of large-scale analog circuits, *Proceedings of IEEE ISCAS '02*, pp. I: 153-6, Phoenix, USA, May 2002.

Ho, C. K., Eberhardt, F., Tenten, W. (2001). Hierarchical fault diagnosis of analog integrated circuits, *IEEE Trans. on CAS – II: Analog and Digital Signal Processing*, Vol. 48, No. 8, pp. 921-929, ISSN 1057-7130.

Huang, J.-L., and Cheng, K.-T. (2000). Test point selection for analog fault diagnosis of unpowered circuit boards, *IEEE Trans. On CAS – II: Analog and Digital Signal Processing*, Vol. 47, No. 10, October 2000, pp. 977-987.

Liobe, J., and Margala, M. (2004). Fault diagnosis of a GHz CMOS LNA Using High-speed ADC-based BIST, *Proc. of the IEEE Defect-Based Testing Workshop*, pp. 85-89, Napa, CA, USA, April 2004.

Litovski, V., Andrejević, M., Zwolinski, M. (2006). Analogue Electronic Circuit Diagnosis Based on ANNs, *Microelectronics Reliability*, Vol. 46(8), August 2006, pp. 1382-1391, ISSN 0026-2714.

Liu, D., and Starzyk, A. (2002). A generalized fault diagnosis method in dynamic analogue circuits", *International Journal of Circuit Theory and Applications*, Vol. 30, pp. 487-510, ISSN 0098-9886.

Luchetta, A., Manetti, S., and Piccirilli, M. C. (2002). Critical comparison among some analog fault diagnosis procedures based on symbolic techniques, *Proc. of DATE'02*, p. 1105, Paris, France.

Maidon, Y., Jervis, B. W., Dutton, N., Lesage, S. (1997). Diagnosis of multifaults in analogue circuits using multilayer perceptrons, *IEE Proc.-Circuits Devices Systems,* Vol. 144, No. 3, June 1997, pp. 149-154.

Manetti, S., and Piccirilli, C. (2003). A singular-value decomposition approach for ambiguity determination in analog circuits, *IEEE Trans. On Circuits and Systems, -I: Fundamental Theory and Applications*, Vol. 50, No. 4, April 2003, pp. 477-487.

Margala, M., Dragic, S., El-Abasiry, A., Ekpe, S., Stopjakova, V. (2002). 1-V Fast IDDQ Current Sensor for On-Line Mixed-Signal/Analog Test, *Proc. IEEE Computer Society Annual Symposium on VLSI*, pp. 165-170, Pittsburgh, PA, USA, April 2002.

Masters, T. (1993). Practical Neural Network Recipes in C++, *Academic Press*, San Diego.

Materka, A. (1994). Neural network for parametric testing of mixed-signal circuits, *Electronics Letters*, Vol. 31, No. 3, February 1994, pp. 183-184, ISSN 0013-5194.

Milor, L., Visvanathan, V. (1989). Detection of Catastrophic Faults in Analog Integrated Circuits, *IEEE Tran. Computer-Aided Design,* Vol. 8, No. 2, Feb. 1989, pp. 114-130.

Milovanović, D., and Litovski, V. (1991). Fault models of CMOS transmission gate, *Int. Journal of Electronics*, Vol. 71, No. 4, October 1991, pp. 675-683.

Milovanović, D., and Litovski, V. (1994). Fault models of CMOS circuits, *Microelectronics Reliability*, Vol. 34, No. 5, pp. 883-896.

Mrčarica, Ž., Ilić, T., and Litovski, V. B. (1999). Time domain analysis of nonlinear switched networks with internally controlled switches, *IEEE Trans. on Circuits and Systems – I Fundamental Theory and Applications*, Vol. 46, pp. 373-378.

Papakostas, D. K., and Hatzopoulos, A. A. (1991). Supply current testing in linear bipolar ICs, *Electronics letters,* Vol. 30, No. 2, pp. 128-130, ISSN 0013-5194.

Pinjala, K. K., Kim, B. C., Varuyam, P. (2003). Automatic Diagnostic Program Generation for Mixed Signal Load Board, *Proc. International Test Conference,* pp. 403-409, Charlotte, NC, USA.

Pipitone, F., Dejong, K., and Spears, W. (1991). An artificial intelligence approach to analogue system diagnosis, In: *Testing and diagnosis of analog circuits and systems*, Liu, R.-W., (Ed.), pp. 187-215, Van Nostrand Reinhold, New York.

Pous, C., Colomer, J., Meléndez J., and de la Rosa, J. L. (2002). Introducing Qualitative Reasoning in fault dictionaries techniques for analog circuits analysis, *Sixteenth International Workshop on Qualitative Reasoning,* Barcelona, Spain, June 2002.

Rodrigez, C., Rementeria, S., Martin, J. I., Lafuente, A., Muguerza, J., Perez, J. (1994). A modular neural network approach to fault diagnosis, *IEEE Trans. on Neural Networks*, Vol. 7., No. 2., March 1996, pp. 326-340, ISSN 1045-9227.

Savioli, C. E, Calvano, J. V., de Mesquita Filho, A. C. (2005). Fault-Trajectory Approach for Fault Diagnosis on Analog Circuits, *Proc. Design, Automation and Test in Europe Conference, DATE '05*, pp. 174-177, March 2005, Munich, Germany.

Scarselli, F., and Tsoi, A. C., (1997). Universal approximation using feed-forward neural networks: A survey of some existing methods and some new results, *Neural Networks*, Vol. 11, No. 1, pp. 15-37.

Sheu, H. -T., Chang, Y.-H. (1997). Robust fault diagnosis for large-scale analog circuits with measurement noises, *IEEE Trans. CAS-I,* , Vol. 44, pp. 198-209, ISSN 1057-7122.

Soma, M., Huynh, S., Zhang, J. (2001). Hierarchical ATPG for Analog Circuits and Systems, *IEEE Design & Test of Computers,* Vol. 18, pp. 72-81, ISSN 0740-7475.

Spina, R., and Upadhyaya, S. (1997). Linear circuit fault diagnosis using neuromorphic analysers, *IEEE Trans. On CAS – II: Analog and Digital Signal Processing*, Vol. 44, No. 3, pp. 188-196, March 1997, ISSN 1057-7130.

Starzyk, J. A. and Liu, D. (2002). A Decomposition Method for Analog Fault Location, *IEEE Int. Symposium on Circuits and Systems*, pp. III-157-160, Scottsdale, Arizona, USA, May 2002.

Stopjakova, V., Malošek, P., Mišučik, D., Matej, M., Margala, M. (2004). Classification of Defective Analog Integrated Circuits Using Artificial Neural Networks, *Journal of Electronic Testing: Theory and Applications,* Vol. 20, February 2004, pp. 25-37, ISSN 0923-8174.

Tadeusuewicz, M., Halgas, S., and Korzybski, M. (2002). An algorithm for soft-fault diagnosis of linear and nonlinear circuits, *IEEE Trans. on CAS – II: Analog and Digital Signal Processing*, Vol. 49, No. 11, Nov. 2002, pp. 1648-1653.

Worsman, M., and Wong, M. W. T. (2002). Non-linear analog circuit fault diagnosis with large change sensitivity, *International Journal of Circuit Theory and Applications*, Vol. 28, pp. 281-303, ISSN 0098-9886.

Xu, X., and Lucas, M. S. P. (1995). Variable-Sampling-Rate Sigma-Delta Modulator for Instrumentation and Measurement, *IEEE Transactions on Instrumentation and Measurement*, Vol. 44, No. 5, October 1995, pp. 929-932.

Yang, Z. R., Zwolinski, M., Chalk, C. D., and Williams, A. C. (2000). Applying a robust heteroscedactic probabilistic neural networor to analog fault detection and classification, *IEEE Transactions on CAS of Int. Circuits and Systems*, Vol. 19, No. 1, January 2000, pp. 142-151.

Yoon, H., Hou, J., Chatterjee, A. and Swaminathan, M. (1998). Fault Detection and Automated Fault Diagnosis for Embedded Integrated Electrical Passives, *International Conference on Computer Design: VLSI in Computers and Processors*, pp. 588-593, Austin, USA, ISSN 1063-6404.

Yu, S., Jervis, B. W., Eckersall, K. R., Bell., I. M., Hall, A. G., and Taylor, G. E. (1994). Neural Network Approach to Fault Diagnosis in CMOS Opamps With Gate Oxide Short Faults, *Electronics Letters*, Vol. 30, No. 9, pp. 695-696, ISSN 0013-5194.

Zografski, Z. (1991). A Novel Machine Learning Algorithm and Its Use in Modeling and Simulation of Dynamical Systems, *Proceedings of 5th Annual European Computer Conference, COMPEURO'91*, pp. 860-864, Bologna, Italy.

Zwolinski, M., Bartt, A., Wilkins, B. R., Suparjo, B. S. (1996). Analogue Circuit Test using RMS Supply Current Monitoring, *IEEE International Mixed Signal Testing Workshop.*

Integration Verification in System on Chips Using Formal Techniques

Subir K Roy
1: Texas Instruments
Bangalore, India

1. Introduction

System on Chips (SoCs) have become an all pervasive component in many of the equipments - both the common placed and the sophisticated, that are relied upon by human beings in today's modern societies; ranging from mobile phones, personal computers, microwave ovens, high definition televisions, base stations for cellular mobile communication and automobiles. Their penetration into every day aspects of human life, and the range of applications and products in which SoCs are being deployed is increasing at a rapid pace. To keep up with this rapid pace it is imperative to design SoCs with reduced turn-around time and cost. Towards this, SoCs are being increasingly designed by integrating existing in house IPs, or third party IPs provided by external vendors. The integration process in realizing an SoC implementation consists of several different kinds of integration which can be classified as (1) static integration, which is essentially of a non-functional nature consisting of simple electrical connections (or hookup) of the inputs and outputs of different component IPs, (2) dynamic, and (3) functional integration; where, besides the pure electrical connectivity, a temporal and a functional dimension, respectively, needs to be taken into account [1]. Typical sizes of state of art SoCs range from fifty million to a few hundred million logic gates. Designing these SoCs involves an integration process consisting of tens of thousands of pure static connections that needs to be established between the input and output ports of the constituent IPs, and when carried out manually can result in introduction of inadvertent errors [1], involving wrong connections, or even, no conncections. The degree of the effects manifested by these errors, depends on when they are detected in the design verification cycle. The latter these are observed in the design cycle, the more difficult and expensive are these to detect, and consequently, to correct, in the implementation. While several approaches have been adopted to tackle the issue of integration verification of SoCs, in this chapter, we focus on the use of formal verification techniques to solve them.

While formal verification has been used in, rather, niche areas of functional validations of IPs and modules, it has found application in the domain of SoC functional validation only recently[13]. With increasing maturity of commercial offerings of formal verification tools by EDA vendors this area of application is expected to grow at a fair pace. The issue of which category of formal verification approaches needs to deployed, for different aspects of SoC functional validation, is however, largely left unanswered. In this chapter we give a glimpse, in Section 2, of the different formal verification techniques that are available, either

as academic tools, or as commercial offerings, and see their applicability to different aspects of SoC verification. We discuss the underlying concepts, the strengths and weaknesses of each approach, the justification for taking these approaches, so that the interested reader can make a judicious choice in their intended application domains. We also point to important references in each of the approaches, so that the interested reader can refer to them for more details.

In Section 3, we will briefly allude to existing methodologies using the formal verification approaches that have been reported in the literature to set the stage for presenting approaches that are not covered by them. More specifically, we will highlight an important aspect of SoC integration verification, vis-a-vis DFT logic, to show the manner in which re-usability is leveraged through automated generation of re-usable parameterized properties and constraints for DFT logic and the hookup or integration logic. And towards "ends justifying means" we will present data and results from their deployment on a real SoC design and show the benefits that can be derived from these approaches. In Sections 4, we will present one interesting scenario from the domain of DFT IP verification.

In Section 5 we will summarize the main contribution of our approaches, which are (1) effective use of formal techniques based on symbolic model checking in the top level verification of SoC integration, (2) effective use of abstraction and modeling of SoC sub-systems in enabling assertion based formal verification, (3) automated generation of assertions and constraints to detect integration errors, (4) automated generation of scripts to capture the SoC design information and invoke a formal verification tool on which to prove the validity or correctness of these assertions. We will end this section and the chapter by drawing conclusions from the presented approaches, data and results, respectively.
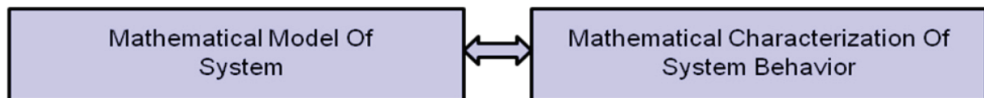
Fig. 1. Formal Models and System Behaviour
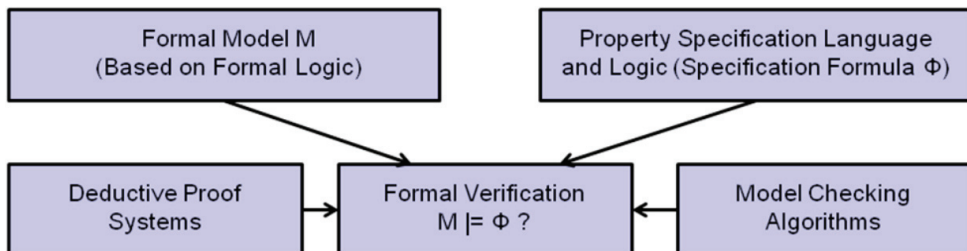
Fig. 2. Generic Structure of the Formal Verification Process

2. Formal approaches

In this section, a brief introduction to formal verification for hardware and a brief review of the different formal verification approach is given. For a detailed presentation and review of hardware formal verification techniques and their application to the problem of verifying IPs the readers are refered to the survey paper given in reference [3, Greenstreet]. The block diagram of the generic structure of the formal verification process, in Figures 1 and 2, succinctly explains the key components involved in formal verification. At the most abstract

level [Figure 1] formal verification essentially consists of having (1) a general mathematical model (**M**) , capturing abstractly the system being verified, (2) the system behavior, described abstractly, again, through a set of mathematically well characterized formulae (**Φ**), and finally (3) proving that the set of formulae (**Φ**) holds true on the mathematical model (M), represented symbolically by **M |= Φ**. This is further elaborated in Figure 2, where **M** is either a computational model or a formal logic model, **Φ** is a set of formulae from a formal logic system, and the proof techniques used for establishing the truth value of **M |= Φ** are either deductive or based on model checking. In deductive proof systems, **M** is decribed by a set of axioms (also known as invariants of the system), and the proof method essentially consists of establishing that the truth of **Φ,** in the underlying formal logic, by using only the given set of invariants (or axioms) of the system. The proof is largely driven by inputs provided by the user, and therefore not fully automated, though steps in the proof may lend themselves to full automation. On the contrary, in the model checking approach, the proof is fully automated for some of the underlying formal logic, as it is based on constructing the reachable set of states of the system.

2.1 Symbolic model checking

A hardware module is formally verified by stating a property on the design and then checking that the design satisfies the property. The most commonly specified property is an invariant, which expresses a condition on the hardware module that should never happen in a reachable state (or conversely, a condition that should always be true in a reachable state). Formally, an *invariant* is a boolean formula over the signals of the module. The module M satisfies the invariant I if every reachable state of M satisfies I . Thus, invariant verification on a module is performed by computing the set of its reachable states. However, this computation is difficult because the set of reachable states can be exponential in the number of signals in the module. This exponential growth in the number of states is known as the *state explosion problem.*

Model checking is one of the most popular approaches to formal verification. In model checking, a mathematical representation of a design in the form of a finite state machine (FSM) is first constructed. Any specified behaviour (or a specification) of the design is then formally stated in terms of a property, or a assertion, in unambiguous terms, both syntactically and semantically, in a formal temporal logic. The mathematical model, i.e. the FSM, is then analysed using different state traversal techniques starting from the set of initial states, to check whether it satisfies the formal temporal property, on all, or atleast, one computational path of the state transition graph that is implicitly generated by the above state traversal. This state traversal is known as *reachability analysis*. In case the temporal property is violated or falsified, a trace with respect to the primary inputs and state variables of the FSM, starting from its set of initial states, is generated up to the Kth set of states, where the property fails on one of its states. This is known as an *error trace*. This search is realized because every set of states that is reachable on each clock cycle starting from the set of initial states is stored internally by the model checker. The collection of such sets of reachable states is finite for a finite state machine. When each of the reachable state set is implicitly represented as a *binary decision diagram (BDD)*, the model checking technique is known as symbolic model checking(SMC). BDDs enable a compact representation of the set of states. In many situations, the negation of a desired property needs to be verified, so that the error trace generated automatically by the symbolic model checker when the stated

property is falsified, or the desired property satisfied because of the negation, will result in a sequence of input and state variable data values in the abstract FSM model. Thus, we implicitly use symbolic model checking as a sophisticated search engine. For a number of hardware designs, while it may be possible to construct the the BDD representation of a very large set of reachable states, it may be impossible and infeasible to explicitly enumerate such a set of states. Despite this, in most cases invariant verification based on SMC techniques is limited to a few hundred signals and states.

In symbolic model checking, properties are specified using different temporal logic, e.g. *Linear Temporal Logic (LTL)*, or *Computation Tree Logic (CTL)* [3]. Some of the temporal properties specified in LTL, or CTL, can be equivalently specified in the form of a finite state machine (FSM) using the same set of internal signals that were used to define them in LTL, or CTL. We, next, give a brief overview of CTL and LTL.

2.2 CTL model checking

The main purpose of a model checker is to verify that a model satisfies a user specified set of desired properties. Specifications to be checked can be expressed in two different temporal logics: the Computation Tree Logic (CTL), and the Linear Temporal Logic (LTL).

CTL is a *branching-time* logic. Its formulas allow for specifying properties that take into account the non-deterministic, branching evolution of a FSM. The evolution of a FSM from a given state can be described as an infinite tree, where the nodes are the states of the FSM and the branching is due to the non-determinism in the transition relation. The paths in the tree that start in a given state are the possible alternative evolutions of the FSM from that state. In CTL one can express properties that should hold for *all the computational paths* that start in a state, as well as, those that should hold only for *some of the computational paths*.

As an example, consider the following CTL formula - *AF p*. It expresses the condition that, for *all* the paths (*A*) starting from a state, *eventually in the future* (*F*) condition *p* must hold. Thus, in every possible single path of the computation tree over which the abstract model of the design, or system, evolves temporally, it will eventually reach a state in which the condition *p* is logically satisfied; i.e. in the considered temporal logic the formula will be asserted as a TRUE, in this state. Differently from this, the CTL formula *EF p*, has the semantics, that requires the *existence* (E) of any one, or some path that eventually, in the future, satisfies *p*. Similarly, formula *AG p* semantically implies that condition *p* is satisfied always (or *globally*), i.e. it is true in every state in every path that exists in the computation tree; while formula *EG p* requires that there is some path along which condition *p* is true in all states in that path. Other CTL operators are as follows,

- *A[p U q]* and *E[p U q]*, requiring condition *p* to be true *until* a state is reached that satisfies condition *q*;
- *AXp* and *EXp*, respectively, require that condition *p* is true in all, or in some of the next states reachable from the current state.

2.3 LTL model checking

In this, specifications or properties are expressed in linear temporal logic (LTL). LTL characterizes each linear path induced by the FSM (linear time approach). LTL has a different expressive power as compared to CTL. Typical LTL operators are :

- *Fp* ("in the future *p*"), stating that a certain condition *p* holds in one of the future time instants.

- *G p* (" globally *p*"), stating that a certain condition *p* holds in all future time instants.
- *p U q* ("*p* until *q*"), stating that condition *p* holds until a state is reached where condition *q* holds.
- *X p* ("next *p*"), stating that condition *p* is true in the next state.

Compared to CTL, LTL temporal operators do not have CTL path quantifiers **A** or **E**. LTL formulas are evaluated on linear paths, and a formula is considered true in a given state, if it is true for all paths starting in that state. Its performance is similar to CTL model check as described above. It has been shown that the complexity of a LTL symbolic model checking algorithm is higher than that of a CTL symbolic model checking algorithm.

2.4 Bounded model checking

In Bounded Model Checking (BMC) the model checker instead of evaluating CTL or LTL properties on paths over infinite time, does so over a finite time defined by a parameter *k* which represents *k* units of time. It tries to find a counterexample of increasing length, and immediately stops when it succeeds, declaring that the formula is false. The maximum number of iterations can be controlled by the parameter *k*. If the maximum number of iterations is reached and no counter-example is found, then the model checker exits, and the truth of the formula is not decided, i.e. it cannot be concluded that the formula is true, but only that any counter-example should be longer than the maximum length. The model checking engine in most implementations of BMC is based on a satisfiability (SAT) solvers instead of BDDs. The complexity of SAT solvers depend on the number of satisfiability constraints that need to be formulated, which in turn is directly dependent on the parameter *k*. For reasonable values of *k,* BMC based on SAT is computationally more efficient than SMC based on BDDs [4].

2.5 Checking invariants

BMC can be used, not only for checking LTL specification, but also for checking invariants. An invariant is a propositional property which must always hold. BMC tries to prove the truth of invariants via a process of inductive reasoning, by checking if (i) the property holds in every initial state, and (ii) if it holds in every state that is reachable from a state where the propositional property holds.

2.6 Newer approaches

Here, we highlight the need to look for other formal verification approaches. We present brief descriptions of some of the promising approaches that are from areas of ongoing research and development in formal verification, in both academic and industrial research circles.

Formal verification has been applied to many classes of designs [13]. We will discuss this aspect in some details in a later sub-section. The key drawback of the automated symbolic model checking based formal verification approaches has been the bane of state explosion faced by even moderately sized modules. Any module, in which the number of state elements or flip-flops exceeds 1000, is liable to face the issue of state explosion during the formal proof of the properties. Microprocessors with modest capabilities, such as the following - non-pipelined instruction stage, single stage instruction pipeline, four stage instruction pipeline, and a four stage instruction pipeline supporting jump and branch instructions - are known to result in state explosion. Typically, for the different SMC approaches the increasing order of performance are as follows,

- CTL, or LTL model checking,
- Invariance checking using CTL,
- Invariance checking with the CTL, or LTL temporal properties represented as FSMs,
- Bounded model checking, and
- Bounded model checking with CTL, or LTL temporal properties represented as FSMs.

One approach to addressing the state explosion problem in such designs is to use compositional formal verification techniques, at the module level of the design heirarchy. Compositional verification is enabled by the assume and guarantee approach [3]. This is shown in Figure 3 below.
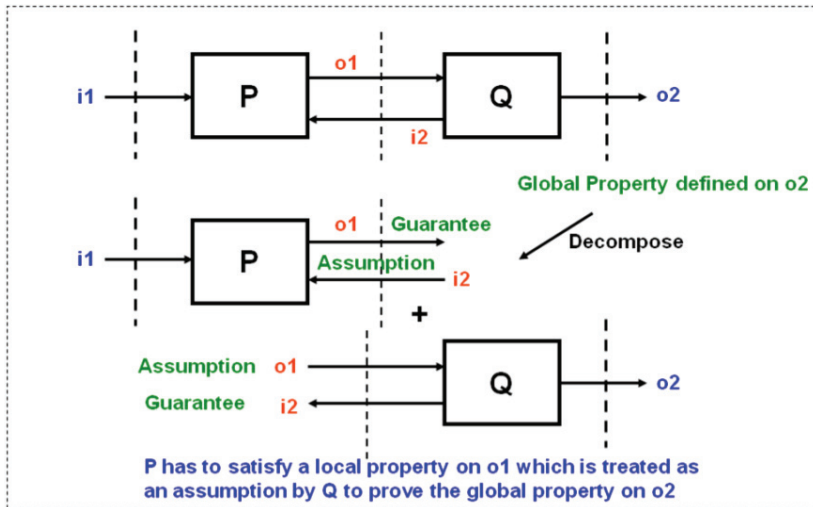
Fig. 3. Assume and Gaurantee approach to Compositional Verification
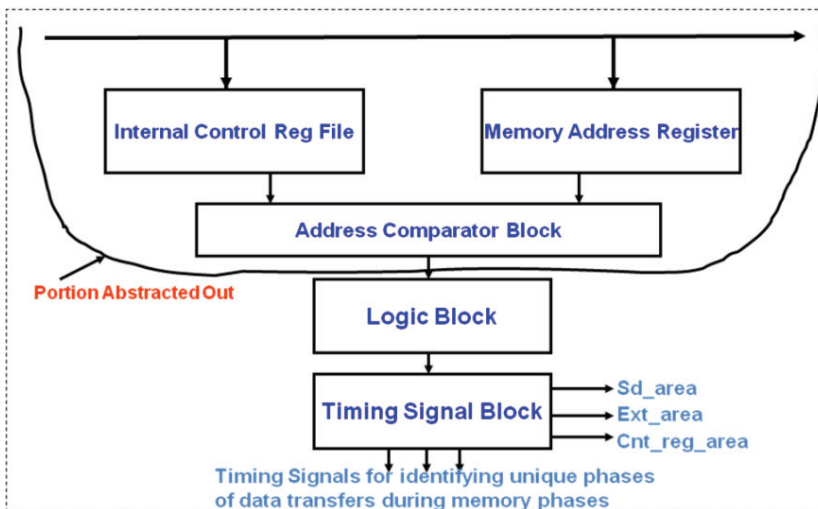
Fig. 4. Design Abstractions to Reduce Complexity of Formal Verification

Another approach is that of abstraction (see Figure 4 and Figure 5), where the design is abstracted (or simplified), to remove portions of design not needed to prove a property. This can result in a substantial reduction in the number of flip-flops, thereby enabling automated proof convergences of the formal properties.

However, for complex industrial RISC and DSP processors, or SoCs based on them, even these approaches are not be feasible. We will need newer formal verification approaches which are not limited by the state explosion problem.

Recent research carried out by different academic and industrial research groups address these capacity issues in formal verification. Though, no stable implementations of formal verification tools exist for such approaches, they serve as good pointers to pursue in the future to address difficult verification problems. We give below, very brief descriptions of some of the approaches.
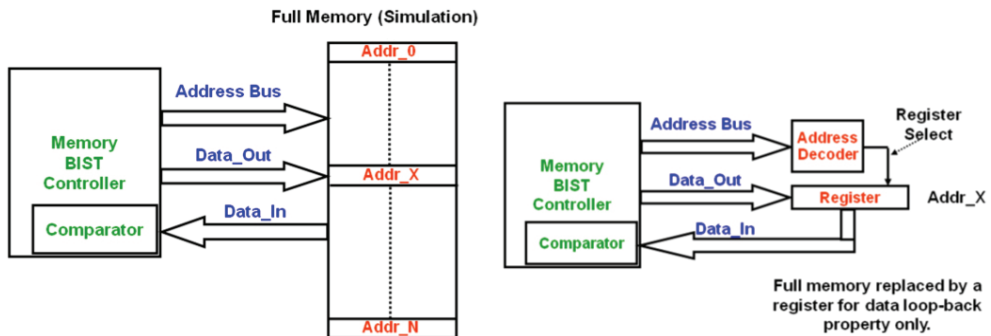
Fig. 5. Memory Abstraction to Reduce Complexity of Formal Verification (Each memory bit adds to a state bit in the verification process)

2.7 Generalized symbolic trajectory evaluations

Symbolic trajectory evaluation (STE) provides a means to formally verify properties of a sequential system by a modified form of symbolic simulation. In this the desired system properties or specifications are expressed in a notation combining Boolean expressions and the temporal *next-time* operator. If the state space of a system is a lattice, the behavior of the sytem can be expressed as a *trajectory*, a sequence of points in the lattice determined by the initial state and the system functionality. Formulas in a simple temporal logic express properties of the system. Given a formula, one can derive bounds that trajectories with the desired property must obey. In its simplest form , each property is expressed as an assertion [*A* => *C*], where the antecedent *A* is a trajectory formula which expresses some assumed conditions on the system state over a bounded time period, and the consequent *C* another trajectory formula which expresses conditions that should result. That is, it determines whether or not every state sequence satisfying *A* must also satisfy *C*. It does this by generating a symbolic simulation sequence corresponding to *A*, and testing whether the resulting symbolic state sequence satisfies *C*. A generalization allows simple invariants to be established and proved automatically.

The Boolean expressions provide a convenient means of describing many different operating conditions in a compact form. By allowing only the most elementary of temporal operators, the class of properties that can be expressed is relatively restricted, as compared

to other temporal logics. However, it has been found in [5] that many aspects of synchronous digital systems at various levels of abstraction can be readily expressed. It is adequate for expressing many subtleties of system operations such as instruction pipelining in modern processors.

The verifier operates on system models in which the state space is ordered by "*information content*". By suitable restrictions to the specification notation, it can be guaranteed [5] that for every trajectory formula, there is a unique weakest trajectory for *A* and testing adherence to *C*. Also, establishing invariants corresponds to simple fixed point calculations. STE implementation of [5] requires a comparatively small amount of simulation and symbolic manipulation to verify an assertion. In [5] it is shown that the length of the simulation sequence depends only on the depth of nesting of the temporal next time operators in the assertion and the speed of convergence of the fixed point calculations.

Formal verification techniques such as, symbolic model checking and theorem proving have met with limited success because of intrinsic problems related to state explosion and the need for manual intervention, respectively. Even though STE is less sensitive to state explosion problem and proven to be a viable methodology for large scale data path verification, it suffers from the problem of inexpressibility. Properties which are spread over infinite time intervals cannot be expressed in STE, let alone be verified [5,6]. GSTE constitutes a very significant extension to STE [8-10]. It has been used successfully by INTEL on its new generation microprocessor designs. GSTE addresses the drawbacks of STE and has the power to verify complex assertion graphs with which any ω-regular property can be equivalently represented, while at the same time it preserves the benefits of STE, like the insensitivity to state explosion, thereby capturing the expressiveness of classical model checking ([3-4] and [6]).

Verification of a complex pipelined data path designs and memories using GSTE model checking techniques have been reported in the literature. Complex properties which are spread over infinite time intervals are specified and verified. The verification time is improved by carefully reducing the number of precise nodes used to perform reachability analysis, while providing complete state information to the symbolic simulator. These results prove the viability of the GSTE methodology as a formal verification technique for control dominated designs such as large scale pipelined data paths. GSTE, therefore, seems a good candidate formal verification approach to use, as it appears to scale well with the actual implementation model of a processor.

2.8 Theorem provers

These are based on formal systems such as logic. For hardware verification, both the specification and implementation can be described in a formal logic, and the task of verifying the system is to prove that the implementation entails the specification. The core of a theorem prover is a set of axioms and inference rules. Using only these, the user can prove a theorem, with the system mechanically checking each step in a proof. One of the best known theorem provers is the HOL system [3], with which theories of different sorts can be built up in a rigorous way using a small number of primitive axioms and inference rules. One major advantage is that the proof can be checked mechanically. Another advantage is that it can be used to argue at different levels of abstraction. As theorem proving is structural rather than behavioral, one can exploit the structure of the system to manage complexity. A major disadvantage of theorem provers is that it can be extremely tedious to verify certain low level properties of systems.

In [6] the combining of theorem proving and trajectory evaluation is explored, with a motivation to gain the benefits of both the approaches. In their theorem proving approach the mathematical objects manipulated by the theorem prover are the trajectory assertions. VOSS is an implementation of these ideas in which STE is used to perform partial verification based on the decomposition of the original specification. Combinational theory is then used to combine these results through the use of the theorem prover framework.

2.9 Logic of Positive Equality with Uninterpreted Functions (PEUF)

This provides a means of abstracting the manipulation of data by a processor when verifying the correctness of its control logic. By reducing formulas in this logic to propositional formulas, one can apply Boolean methods such as BDDs and Boolean satisfiability checkers to perform the verification. In [7], two approaches have been shown to translate formulas in PEUF into propositional logic. The first interprets the formula over a domain of fixed length bit vectors and uses vectors of propositional variables to encode domain variables. The second generates formulas encoding the conditions under which pairs of terms have equal valuations, introducing propositional variables to encode the equality relations between pair of terms. In [7] techniques are presented to drastically reduce the number of propositional variables that need to be introduced and to reduce the overall formula sizes. This allows verification of microprocessors with load, store and branch instructions at both the RTL or the gate level model. This again makes the approach based on PEUF, a good candidate for solving many formal verification problems.

3. Existing formal based approaches

In this section, we first justify the need to resort to formal verification, then we will briefly allude to existing methodologies based on the formal verification approach that have been reported in the literature to set the stage for presenting newer approaches in latter sections.

As an example case study, we will present the methods and challenges in verifying the integration of Design For Testability (DFT) logic - both BIST and non-BIST, in complex SoCs using formal techniques. We will first present a generic architecture of the DFT logic that is typically present in state of art SoC designs. For this DFT logic, we will, next, list the validation task that needs to be accomplished, to ensure its proper integration into the functional logic portion being implemented in a SoC design. We will then identify the commonality that exists amongst the listed tasks from the perspective of verification. We will then show how such common verification tasks are amenable to automation. As some of these verification flow automations have already been reported in available literature, we will briefly describe them in the context of our adoption of these flows, and refer the interested reader to relevant reference papers for more details. As these flows are being applied in the regression mode, to the various revisions of the currently ongoing implementation design of an in-house SoC, we report recent data from our formal verification efforts, to show-case the value propositions brought in by these approaches.

To simplify the above discussion, we will assume that different DFT IPs present in the generic DFT architecture are pre-verified. However, in reality, this is not always the case. In many situations, it may be necessary to verify even the different DFT IPs, specially, if these are parameterized, configurable and auto-generated, to ensure that the version intended for integration into the SoC, is indeed being generated correctly. Towards this, we will briefly

discuss, how congigurable DFT IPs can be formally verified with a configurable set of generic properties, so that alongwith any desired IP configuration, the corresponding set of properties are auto-generated to verify the IP. This will demonstrate how re-usability is being leveraged through automated generation of re-usable parameterized properties and environmental constraints for DFT logic and the integration logic.

3.1 Justification for using formal approaches to integration verification

An exceedingly important design phase, which gets carried out in the background, and far from the lime-light of the functional features of any SoC, is the integration of DFT logic and the verification of this integration to other sub-systems and IPs in a SoC. While this does not feature as a prominent front end task in the design of any SoC, it does constitute a significant portion of the overall design and verification effort. Any savings in this design integration phase, and its subsequent verification, helps in reducing the overall SoC design cost. Some of the key components in DFT logic that need to be integrated into a SoC are those for 1) testing embedded memories and core logic, 2) control logic to enable different test modes to be set up during post-fabrication Silicon testing, 3) multiplexing and control logic to enable different test modes to selectively run tests, directly from SoC top level ports on different portions of functional logic by bypassing intervening logic blocks, 4) configuring scan chains to rout test vectors to different portions of functional logic.

A key to realizing the above mentioned cost reduction for the above tasks can be through their automation. A prominent factor that can help in facilitating automation is the fact that most DFT IPs possess behaviors and structures that are of an extremely canonical and regular nature, and that these are independent of the functional nature of the SoC. Besides this, the interconnection of the IPs to the rest of the logic in the SoC is also of a very generic nature. This has resulted in the consolidation of SoC level DFT logic architecture towards a highly standardized, and a highly configurable form (known as the DFT sub-system), enabling it to be auto-generated through software tools. Individual components within this sub-system are auto-generated using point commercial tools addressing the highly specialized requirements corresponding to each DFT task. Two such examples are, memory and logic BIST controllers. We briefly discuss below the DFT tasks of testing embedded memories and core logic in a SoC, in the context of these controllers.

Present generation SoC designs are built hierarchically with a large number of embedded memories of different kinds and sizes and different embedded cores (e.g. processors, peripherals, etc.). The embedded cores may themselves have different types of internal memories, functional logic blocks and different types of I/O ports. Built-in self-test (BIST) techniques are employed to reduce expensive ATE time for the post manufacturing silicon testing of these blocks; besides, they enable low pin count testing, and testing of embedded core of SoCs fabricated on low cost packages with fewer pin count. Application of test vectors to memories and core logic can span several million clock cycles depending on the size of the embedded memories, the core logic, and the testing algorithm employed to generate these vectors, resulting in exceedingly long verification times for the BIST controllers through simulation. A memory BIST tool needs a description of the embedded memories and memory test pattern generation algorithms to generate and integrate the different memory BIST controller logic needed for different types of memories. In a similar manner, the logic BIST tool requires a description of the gate level net-list representation of a design, to analyze and to extract the core logic portion, before generating the logic BIST

controller needed to configure them for testing and for generation of the test vectors specific to this configuration. The configuration hook-up logic and the logic BIST controller are then integrated into the netlist automatically by the tool. To meet performance, timing and power constraints specific to a SoC, and to support scan, self test and clocking, it is often the case that, such, auto-generated BIST logic have to be modified, thereby, necessitating their verification, to ensure that the modifications do not break the intended behavior or functionality.

Verification of proper integration of memory BIST logic into a design, and any modifications to it, to meet performance and timing constraints in the original design, has been traditionally done based on simulation techniques. This is often incomplete and time consuming, as the correctness of the integration is verified indirectly by running the entire test suite developed for memory BIST logic. Even a single change in the control or hookup logic and its integration into the rest of the design may necessitate re-running the entire set of simulation vectors. Besides being time and compute intensive, the time needed to analyze errors detected in these simulation runs and to correlate them to design integration problems can be correspondingly large. These simulation test benches are often created manually, and may be design specific, making them un-useable across different controller-memory, or controller-embedded logic configurations. Verification of memory BIST logic using formal techniques is appealing, as the behavior of the controller block is sequential, while the behavior of the hookup block is combinational. Writing re-usable formal properties for such blocks are easy, precise and less time consuming. It is possible to obtain comprehensive verification coverage across different environmental constraints, resulting in high quality and confidence in the verification process using formal techniques. Formal verification of BIST controllers, however, can be difficult, if we include models of embedded memories, as in simulation. This is due to the large number of register elements used to model memory, which leads to the problem of state explosion and can be overcome by effective modeling and abstraction techniques.

We next justify the need to verify even the auto-generated DFT logic sub-system and its integration into the SoC. DFT logic sub-systems have to be verified as different modular configurations arising out of generic customizable, configurable and parametrisable components may be needed for different SoCs. This implicitly enforces verification requirements on the integration of such configurable DFT logic modules into an SoC whose RTL itself could be auto-generated with a tool (for example, 1-Team-Genesis [11]) and with its own set of configurable functional IPs. While there is variability in the configurations, each configuration nevertheless, retains the above characteristics, thereby, rendering the verification of DFT logic and its integration into a SoC a very good candidate for formal approaches. To leverage the capabilities of FV in the context of auto-generated configurable modules, it is essential that the formal properties themselves be configurable and auto-generated, along with the formal verification environment. This enables high re-usability of properties developed during the tactical formal verification of each module present in the DFT logic subsystem in different SoCs. While the generation of DFT logic and its integration in a SoC is automated, our approach results in the automation of the verification task as well. This enables the complete automation of DFT logic in terms of verification and integration in a SoC at the RTL implementation level, resulting in a considerable reduction in the overall SoC design cost and design turnaround time. We briefly describe below the process by which we systematically achieved this automation.
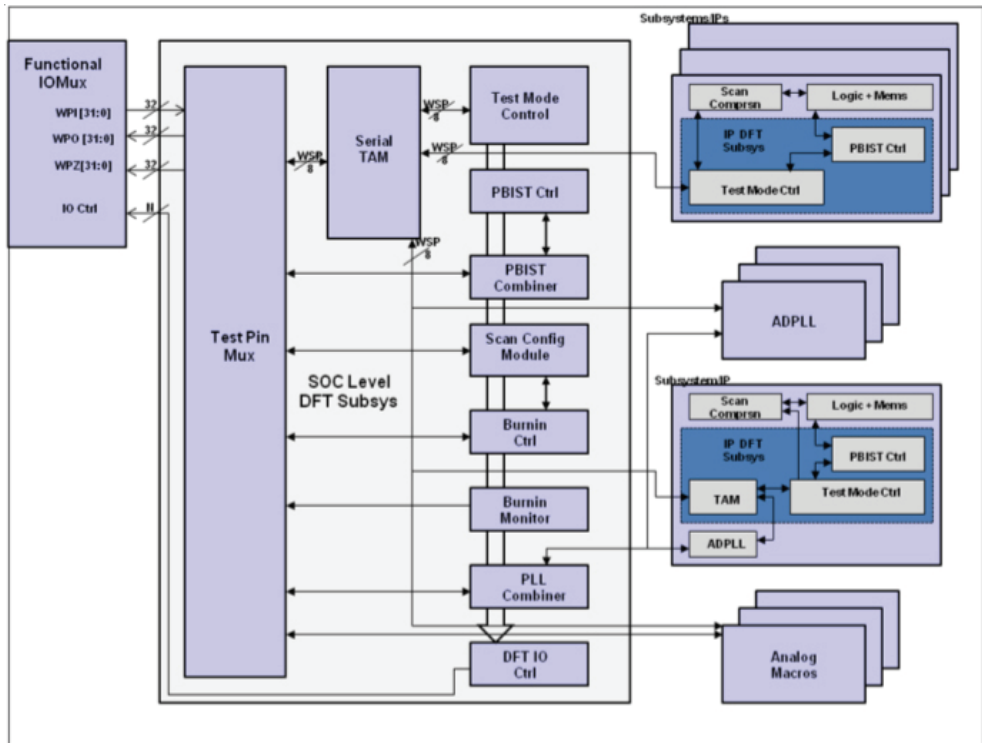
Fig. 6. Generic SoC DFT Logic Architecture

Verification IPs (VIPs) in the form of formal properties, verification environment and verification tool setup were developed tactically for the maximal configuration possible in each block within the standardized configurable DFT sub-system described above, using techniques given in [1,2]. These VIPs were then validated on several in-house driver SoC designs. Once these VIPs reached a level of maturity by way of functional specification coverage, their parameterization in the context of individual blocks were taken up, to enable different sets of VIPs to be auto-generated for a given set of parameters specific to a particular desired configuration of the DFT sub-system. VIPs necessary to check the correct generation of the sub-system includes the VIPs to check the correct integration of individual blocks within the sub-system. Different verification sub-tasks related to the validation of behavior of DFT logic and its interaction with functional logic under different test modes were identified, and corresponding VIPs along with their auto-generation scripts were developed tactically. These were then validated on several driver designs. The tactical development of these VIPs on driver SoC designs were then moved into a common infrastructure through which desired configurations of the DFT sub-system logic and VIPs specific to different SoC designs are generated, enabling high re-use and faster turn-around times.

We next give details of how some of the formal verification flows related to the different DFT verification tasks have been achieved. Towards we first briefly describe a generic DFT logic architecture typically found in any state of art SoC.
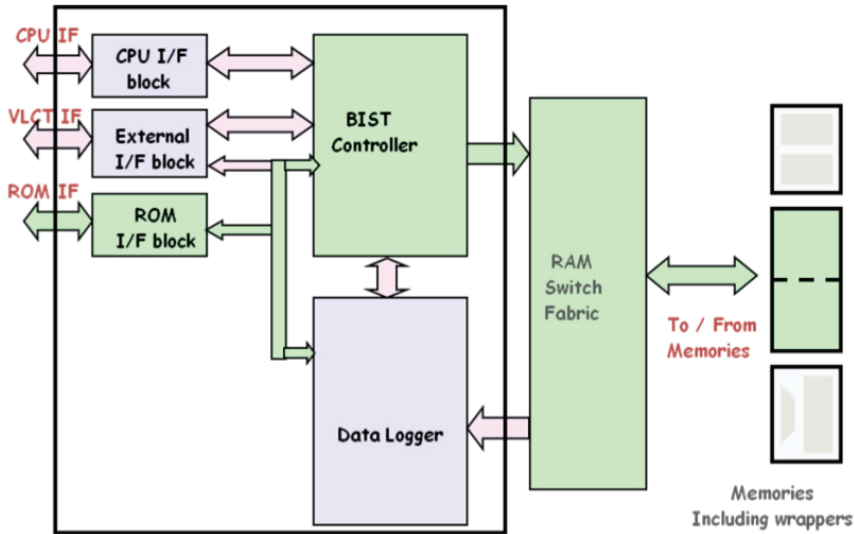
Fig. 7. Microcoded Programmable Memory BIST Controller Architecture.
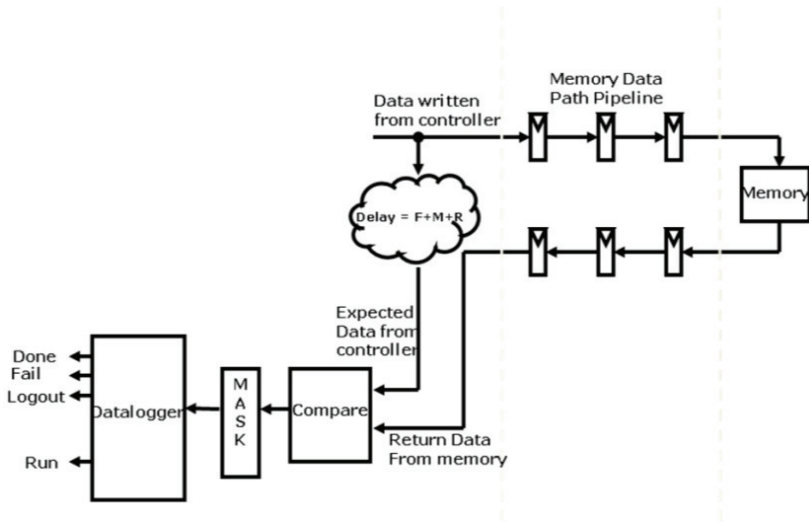
Fig. 8. Flow of Memory Data between PBIST Controller and Embedded Memories

3.2 Generic SoC DFT logic architecture

Figure 6 shows the DFT logic architecture typically found in any state of art SoC design. This SoC has a heirarchical DFT logic architecture characterised by a complex top level DFT sub-system, and depending on the complexity of the constituent IPs, several simpler IP level DFT sub-systems could be present. The *Functional_IO_Mux* block at the top level of the SoC routes external inputs to the SoC to either the functional core logic or to the DFT sub-sytems depending on the SoC operational modes, viz., functional or test modes. Under the test
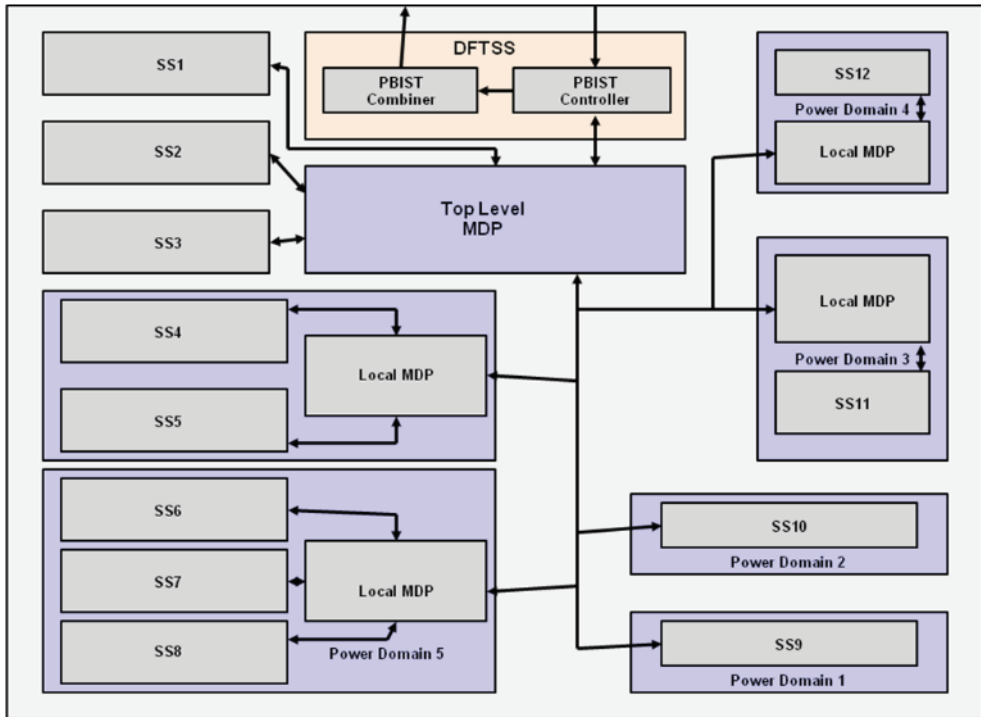
Fig. 9. Generic Top Level Heirarchical Memory Data Path Architecture in a SoC

mode, the external inputs are routed by the *Test_Pin_Mux* block, under different test modes to different modules within the top level DFT sub-system or the IP level DFT sub-system. The *Test_Pin_Mux* block achieves the routing to IP level DFT sub-system blocks through a IEEE1500 module in the *Test_Mode_Ctrl* block, which not only enables setting up of the various SoC level test modes, but also IP level test modes. The latter is achieved through a serial programmation of the *Serial_TAM* block under the control of the IEEE1500 module. The programmation of the IEEE1500 module is carried out by the ATE through top level SoC JTAG ports. The *Test_Mode_Ctrl* block exercises control over choice of, either serial test data, or parallel test data through the *Functional_IO_Mux* block, through the *DFT_IO_Ctrl* block based on the requirements imposed by the different SoC and IP test modes. The different IP level test modes are set by the programmation of the *Serial_TAM* block in the top level DFT sub-system under the control of the IEEE1500 module. Based on the value written into its control register the ports of the different IP level TAM blocks get connected to its corresponding ports. Serial data from the top level SoC ports can then be routed directly to the individual IP level *Test_Mode_Ctrl* blocks to set the desired test modes within the IPs.

The *PBIST_Ctrl* block in the top level DFT sub-system is the programmable BIST controller. Several PBIST controller blocks can also be present in the IP level DFT sub-systems as shown in Figure 6. Memory test data from these controllers and the top level PBIST controller, are routed sequentially to top level SoC test ports by the *PBIST_Combiner* block.

The *PLL_Combiner* block controls the generation of the various functional and test clocks with different frequencies needed by different IPs and test controllers under the different functional and test modes of operation. These clocks are generated from the top level system clock of the SoC. For the sake of brevity we will not discuss the remaining blocks present in the top level DFT sub-system.

The correctness of the SoC DFT logic architecture described above is established through different sets of verification checks carried out at different levels in its module hierarchy. We list below some of them.

- SoC Level Checks
    - Hook-up checks
    - IO Related Checks
    - Memory Data Path
    - P1500 Slave Verification
    - Test Mode Entry
- Module Level Checks
    - Burn in monitor module
    - Clock observation module
    - Test secure controller
    - Test clock management module

There are different categories of connectivity checks (*Hook-up Checks*) that need to be carried out at the top level. As listed below there are a large number of connectivity checks that need to be performed at the SoC top level under various categories.

- Hook-up checks
    - Test pin mux verification
    - Clock propagation checks
    - ATPG reset propagation checks
    - ATPG control signal checks for soft macros
    - Memory power management ports hook up
    - Power switch ports hook up
    - WPI/WPO connectivity from DFT-SS to complex IO's, analog macros, digital hard IPs
    - DFT-SS DFT read/write signals to control modules
    - Compression wrapper connectivity
    - Connectivity checks between DFTSS to IPs
    - Burnin monitor input/output connectivity
    - Clock observation/lock observation signal connectivity
    - PLLCM/ADPLL connectivity
    - Connectivity checks for PBIST, DPLL, SCM interface
    - IForce/VSense connectivity checks
    - Memory port connectivity checks
    - Margin mode pin checks
    - Memory power management ports hook up
    - Power switch ports hook up
- Direct Connectivity
    - TPM, DPLL, SCM, PBIST, ATPG Reset, PRCM Clock,
- Muxed Connectivity
    - Burnin monitor muxing logic

- Safe Value
    - IE, PU/PD, GZ checks
- Connectivity with inverted value
    - Slew Override checks
- Test Mode Entry
    - THBMode, TestMode
- Clock division based on a division factor
    - Clock Observation Module
- TAM Connectivity
- Memory Data Path connectivity
- Register loading through JTAG

Most of the above checks are simple point to point, static connectivity checks which have been discussed in details in references [1,2]. We will discuss below, briefly, one check which is more complex as compared to the other checks. More details on this check can be found in the references [1, 2]. This is the memory data path (MDP) check, in which, the correctness of the pipelined datapath connectivity between a PBIST controller and its corresponding set of embedded memories is established. This correctness has to be established individually between every possible pair of controller-memory combination. The setting up of each unique pair is achieved by a hierarchical mux logic structure known as the Memory Data Path (MDP). Each pair can have different numbers of pipelined registers along both the forward memory data path (from controller to the embedded memory) and the return path (from embedded memory to the controller), to account for the different path delays due to different geographical seperations of the embedded memories vis-à-vis the controller (Figure 7 and Figure 8).
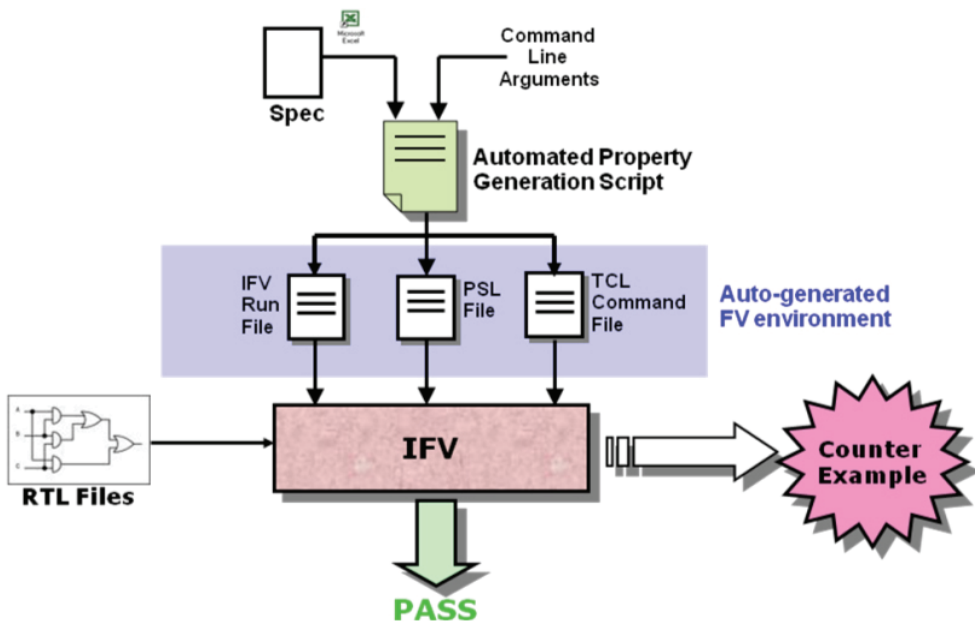
Fig. 10. Automation Flow For Memory Data Verification Using Formal Properties

The MDP check consists of the following – correct establishment of a pair, correct temporal to and fro transportation of the address, memory and control data between the corresponding memory ports and controller ports. Establishing the correctness of a pair under a unique control value issued by the PBIST controller is done by ensuring that only the desired pair is chosen, and that for this pair no other pair is chosen. As will be seen below typical SoCs can have tens of embedded memories of differing types. Besides, each IP can have its own local PBIST controller. Figure 9 shows the hierarchical MDP structure in a typical SoC, with the grouping of the local MDPs being based on the different power domains that each IP belongs to in the SoC. Based on the argument presented earlier in the section, formal verification of the MDP structure for an ongoing SoC implementation is being carried out using an automated formal verification flow, shown in Figure 10. Details of this flow can be found in [1, 2]. We discuss the data obtained from our formal verification efforts for this check.

Of the 14 functional sub-subsystems being integrated into the SoC, 9 are soft IPs, whereas, 5 are hard, pre-verified, third party IPs, requiring only simple connectivity checks. Full set of MDP connectivity checks are carried out on the soft-IPs. The number of memories and their corresponding ports on which connectivity checks are performed are listed in Table 1 below.

IPs	Number of Memories	Ports checked (IP/Mem End)
Sub-System1	3	td, ta, taw, tar, q, twen, tm, twrenz
Sub-System2	1	td, taw, tar, q, twen, tm, twrenz
Sub-System3	28	td, ta, taw, tar, q, twen, tm, twtz, twz, twrenz
Sub-System4	3	td, taw, tar, q, twen, tm, twrenz
Sub-System5	23	td, taw, tar, q, twen, tm, twrenz
Sub-System6	2	td, taw, tar, q, twen, tm, twrenz
Sub-System7	10	td, taw, tar, q, twen, tm, twrenz
Sub-System8	3	td, ta, taw, tar, q, twen, tm, twrenz, twtz, twz, tez0
Sub-System9	1	a, ta, q, ez tez, tm
Sub-System10	Hard IP – connectivity checks	csr, rgs, rds, rdata*, wdata*, addr*, wtz*, ms*, tm, wz*, twrenz
Sub-System11	Hard IP – connectivity checks	csr, rgs, rds, rdata*, wdata*, addr*, wtz*, ms*, tm, wz*, twrenz
Sub-System12	Hard IP – connectivity checks	wpi_memory_bist*
Sub-System13	Hard IP – connectivity checks	wpi_memory_bist*
Sub-System14	Hard IP – connectivity checks	wpi_memory_bist*
Total	74 + connectivity checks	NA

Table 1. Sub-systems, Their Memories and Signals for Formal Verification in Example SoC

The total number of formal properties for each IP is given in Table 2. This table also shows the progression of the checks on different RTL versions released by the design team at different points in the temporal evolution of the SoC implementation. Several useful bugs were caught by the formal verification runs in each release. As can be clearly seen, over each iteration there is a reduction in the number of bugs caught by formal verification.

Towards the formal verification runs, the set up time needed for the first RTL release using our automated flow was approximately 36 hours for all the 14 sub-systems. Most of this

time was devoted towards establishing the correct environmetal constraints to be applied at the SoC top level for formal verification runs, and the right heirarchical paths of each functional IP in the SoC, and each module in the heirarchical DFT logic architecture, to enable black-boxing of un-necessary modules. This results in efficient and faster convergence of the properties during formal runs. This is a one time effort. Set up times in subsequent regression runs are drastically reduced to around an hour. A PERL based script is under development to completely automate the above.

	1st iteration			2nd iteration			3rd iteration			4th iteration		
IPs	Prps	Pass	Fail	Prps	Pass	Fail	Prps	Pass	Fail	Prps	Pass	Fail
SubSys1	148	144	4	148	148	0	148	148	0	149	148	1
SubSys2	68	67	1	68	68	0	68	68	0	69	68	1
SubSys3	1344	1312	32	1344	1344	0	1344	1344	0	1372	1372	0
SubSys4	158	155	3	158	158	0	158	158	0	158	158	0
SubSys5	1363	1324	37	1363	1363	0	1363	1363	0	1386	1386	0
SubSys6	48	46	2	48	48	0	48	48	0	54	54	0
SubSys7	670	660	10	670	670	0	670	670	0	690	690	0
SubSys8	172	168	4	172	168	0	172	172	0	175	175	0
SubSys9	38	5	33	38	38	0	38	38	0	38	38	0
SubSys10	29	21	8	29	29	0	33	33	0	35	35	0
SubSys11	29	4	25	29	29	0	33	33	0	35	35	0
SubSys12	NA	NA	NA	3	3	0	3	3	0	3	3	0
SubSys13	NA	NA	NA	3	3	0	3	3	0	3	3	0
SubSys14	NA	NA	NA	3	3	0	3	3	0	3	3	0
Total	4067	3906	161	4076	4076	0	4076	4076	0	4150	4148	2

Table 2. Data For FV Regression Runs on SoC Sub-system Memories For Different RTL releases

The MDP checks on the different sus-sytems/IPs varies from 5 minutes to 10 minutes, with an overall verification time of 90 minutes over different regression runs for different RTL releases. Thus, for each RTL release a regression run of the MDP checks can be completed within 150 minutes (2.5 hours). Simulation based regressions runs need atleast a day to report similar results. This has been consistently observed with respect to other formal verification flows developed to carry out the different SoC level integration checks listed earlier. In Table 3 we report some data based on these checks performed on the latest RTL implementation release of the SoC discussed above.

In the next section we take up the task of verifying formally some interesting aspects of one of the DFT IP discussed above.

4. Formal verification of protocols for transfer of programs and data in programmable DFT controllers

An oft repeated claim in the context of formal verification in the verification community, both academic, as well as, industry has been that model checking based formal approaches do not work well for designs that have behaviors involving aperiodic events with long latencies, such as found in Ethernet MAC interfaces and Elastic Buffers. In this section we discuss a strategy devised to formally verify one such design which involves huge sequential depths.

SL. No.		Properties	Passes	Fails
1	Test Pin Muxing Connectivity	984	984	0
	+ Safe Value Checks	272	272	0
2	SCM Interface Connectivity	518	518	0
3	Burn in Monitor	134	134	0
4	Clock Observation Module Hookup	48	48	
5	Clock divider	1536	1536	0
6	IO Checks –THBMode	475	475	0
7	IO Checks – HiZ instruction	475	475	0
8	IO checks – IDDQ	777	777	0
9	DPLL Interface Connectivity	90	90	0
10	Compression Wrapper Connectivity	153	143	10
11	Boundary Scan Register Connectivity + Override Checks	1610	1592	18
12	EFuse Connectivity	7	5	2
	+ LDO/BG DFT Checks	45	45	0
13	Test Secure Controller Hookup	15	15	0
14	Clock Connectivity Checks	112	112	0
15	Burn-In Monitor Connectivity	95	95	0
	+ Module	161	161	0
16	IEEE1500 TAM Connectivity Checks	550	550	0
17	Memory Margin Mode Checks	21	21	0
18	ATPG Reset Checks	126	73	53
19	Test Mode ATPG Checks	96	96	0
20	DFT Mux Mode	40	40	0
	+ DFT Read/Write Checks	16	16	0

Table 3. Formal Verification Run Statistics on Different SoC DFT Logic Integration Checks

In many critical SoCs (with stringent and low DPPM values) post silicon fabrication verification of embedded memories using programmable built in self test (BIST) controllers involves downloading of memory testing algorithms (for different memory types) in the form of microcoded instructions from an external ROM into the internal memory of the BIST controller. Besides the algorithms, critical information related to them, such as, the embedded memory types and their grouping are also downloaded to enable the controller to execute the memory-testing algorithm on each memory in a group. There is a predetermined grouping of the algorithms and their memory related information, both, within the external ROM and within the internal memory of the controller which enforces a strict protocol with branching semantics to be followed during downloads. Due to limited capacity of the internal memory, the downloading is interleaved with the execution of the memory-testing algorithm by the controller, until each algorithm is downloaded and executed on each memory of their target memory groups. It is, therefore, imperative that the interface implementing the downloading protocol with branching semantics be verified comprehensively for the correct execution of the memory testing algorithms on memories in

the targeted groups. In this section, we show how one can effectively use symbolic model checking based formal approach to verify a complex protocol involving long sequence of events until completion of testing of each embedded memory in the SoC.

4.1 The microcoded programmable memory BIST controller architecture

The design under verification here is a ROM Interface which is a block in a programmable memory BIST controller IP (Programmable BIST, or PBIST), as shown earlier in Figure 7. This figure shows the architecture of the PBIST controller. The path through which the data from the external ROM flows into the controller and the embedded memories are highlighted in green. The different memory testing algorithms (ALGO), the information on the memory type (RAM data – RAMD) and the background patterns (BGPs) specific to the algorithm, are all downloaded from the external ROM. The external ROM communicates with the microcoded PBIST memory controller through the ROM Interface, whereby, the relevant data to be downloaded is transferred sequentially to data type specific registers in a program register file within the controller. The memories to be tested are grouped into RAM Groups (RAMG) (Figure 10). PBIST can be instructed to selectively test specific RAMG's *(the targeted RAMG)* using either a single ALGO, or a set of ALGOs applicable to the different memories in the RAMG. The maximum number of RAM groups and the maximum number of different memory testing algorithms supported in the latest version of the PBIST controller is 64 (with a maximum of upto 51 memories in each memory group) and 32 (with a maximum of upto 14 back ground patterns for each algorithm), respectively.
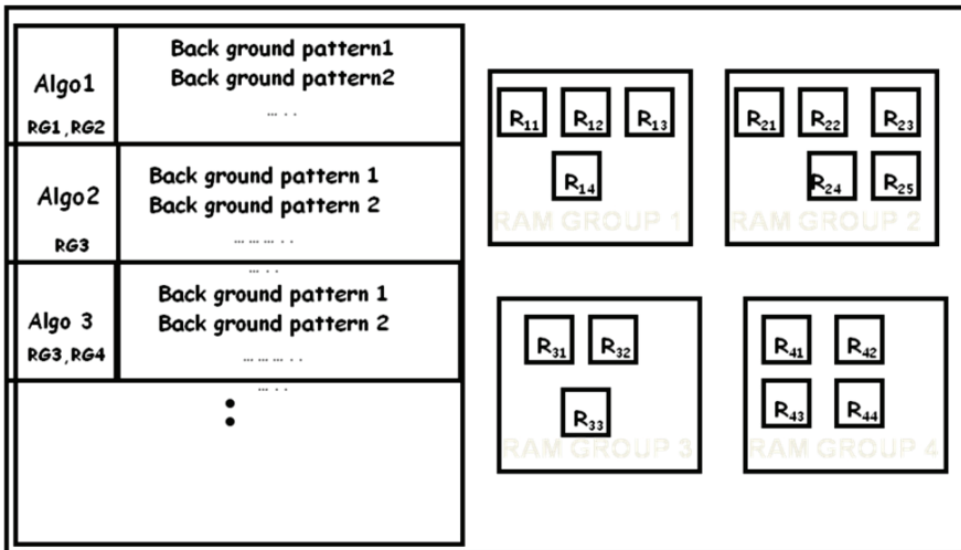
Fig. 10. Memory Test Algorithms and Their Mapping To Memory Groups.

4.2 Source of enormous sequential depth in the PBIST controller's ROM interface behaviour

For the maximum number of algorithms, the maximum number of RAM groups and the maximum number of background patterns that can be supported by a single PBIST

controller, we can easily calculate the maximum number of clock cycles it takes for the controller to assert its MDONE (or PASS) signal in case no memory errors are detected by any of the algorithms executed in each memory in each RAM group. To simplify this calculation we will assume the following relevant set of data values :

- There are 32 ALGO, 14 BGP in each ALGO, 64 RAMG and 51 RAMs in each RAMG.
- Each ALGO targets all the RAMG.
- The ROM has a data read latency of 1 clock (we ignore all clocks during which the controller does not attempt to fetch any data from the ROM; for example, during a switch over from the ALGO section to RAMD section, during the execution of a memory testing algorithm on a specific memory in a specific RAM group).

For the above set of assumed values, the number of clock cycles required to just fetch all the relevant data from the ROM into the PBIST controller based on the transaction protocol shown in Figure 11 alone can be easily seen to be,

$$[34+\{2+(64*51*10)\}*14]*32 = 14.2 \text{ million cycles!}$$
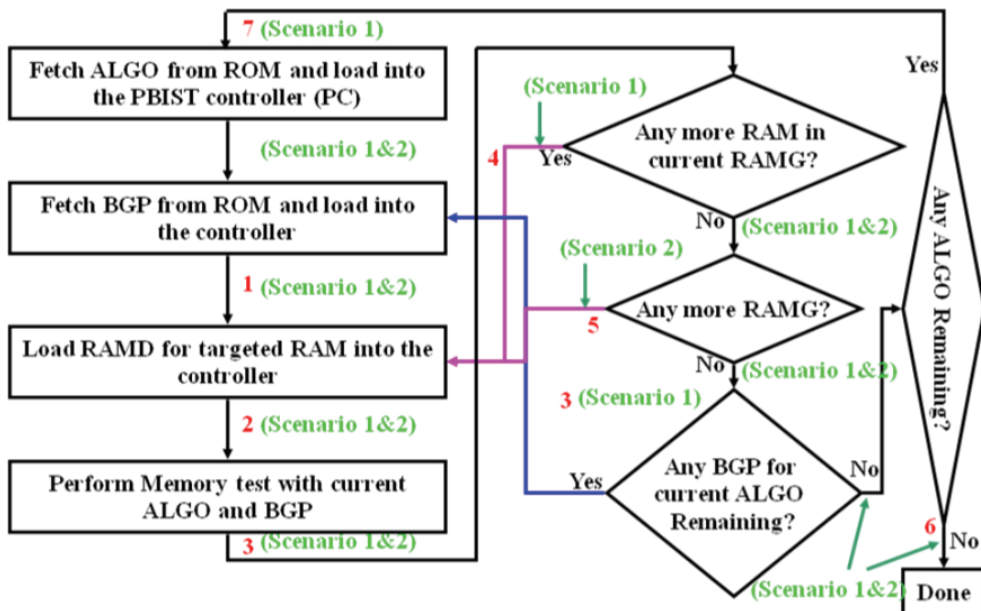
Fig. 11. Transaction Protocol followed by the ROM Interface logic and its functional classification. (Each functional category is numbered in red, while the test case covering it is numbered in green.)

As can be easily noted, this is a rather conservative figure, as we ignore all clock cycles consumed during suspension of data downloads. Besides, if the latency is higher, than the optimistic value of 1, the sequential depth would be even larger. Symbolic Model Checking tools, such as IFV, are incapable of handling functional behaviors with such enormous sequential depths. This was borne out by the fact that even simple properties written to validate the ROM interface behavior exhibited state explosion.

4.3 Verification strategy

To verify design behavior involving exteremly large sequential depths we cannot take recourse to structural abstraction techniques based on module heirarchies to reduce the complexity of the verification effort. A close look at the root cause of the issue reveals the following - during the download process of a data element from the external ROM, one part of a switching logic block is repeatedly exercised every time the control jumps from downloading data from the memory testing algorithm section to the memory data section. Therefore, for a maximum of N memory testing algorithms that are supported by the protocol, this logic will be exercised N times. This also implies that a property written to verify the sequence of events associated with this switching, would be triggered N times in the antecent of the property and, therefore, the final pass status, depending on the satisfaction of the consequent of the property will be declared, after an extremely large sequential depth with respect to the set of initial states is traversed. A simple startegy of reducing N, to say 5, not only exercises the switching logic to check for any corner case arising from the switch in data transcation from the algorithm portion to the memory data portion, and vice-versa; but also results in a smaller sequential depth. This simple idea is similarly used to reduce the number of background test patterns assigned to each ALGO, the number of memory groups, and, finally, the number of memories in each group. Towards this, we chose values of 5 for the number of algorithms, 5 for the background patterns, 5 for memory groups and 5 memories in each group, respectively.

We simplified the verification task, further, by splitting the environment to enable verification of two different cases:

i.   1 algorithm, 1 BGP, 1 RAMG and 5 RAMs in each RAMG.
ii.   5 algorithms, 5 BGP, 5 RAMG and 1 RAM in each RAMG.

The two cases have been carefully chosen to further reduce the sequential depths traversed by IFV to prove the corresponding properties, as well as, exercise complementary portions of the corresponding logic in the RTL. For example, in test case 1, logic enabling transition to a new memory testing algorithm will not be exercised, as only one algortihm is assumed to be present; while in test case 2, logic enabling transition to a new memory in a memory group will not be exercised, as only one memory is assumed to be present in each memory group. The gaurantee on the exhaustiveness of the verification process with respect to the entire functional behavior of the ROM interface logic is based on the following considerations. A complete analysis of the RTL functionality results in its being classifiable into the following seven categories:

1. Branching into the RAMD section once an algorithm and the first BGP have been fetched from the external ROM and transferred to the PBIST controller.
2. Branching from the RAMD section into a wait mode and remaining in that mode until an external signal flags the completion of a memory testing algorithm test on the corresponding RAM.
3. Once a RAM has been tested, the information for the next RAM within the same RAMG needs to be fetched provided the currently chosen RAM is not the last RAM within the present RAMG.
4. Once a chosen testing algorithm has been executed completedly on all the RAMs in a RAM group, control should revert back to the BGP section to enable fetching the next BGP, corresponding to the memory test algorithm to be run on the next memory group.
5. After a RAM has been tested, the information for the first RAM of the next RAMG needs to be fetched, provided the current RAM is the last RAM within the current RAMG

6. After all the memory testing algorithms with all their respective BGPs have run on all their targeted RAMs, the control should revert back to the idle state of the underlying control FSM of the ROM interface logic.

7. Once a RAM has been tested the control should revert back to the ALGO section to enable fetching the next ALGO, in case all the RAMs targeted by the current algorithm have been tested.

The above categories of logic are marked on a process flow diagram to ensure that none of the interface functionality is missed by the above classification. This flow diagram is shown below in Figure 11. In this figure, the verification test case which covers one of the above sub-functionality is marked in green and red, respectively. The overall coverage for each test case is captured in Table 4 below.

Verification Test Cases	Targeted Functionalities
Test Case 1	1,2,4, 6 & 7
Test Case 2	1,2,3, 5 & 6

Table 4. Functional category coverage by the different test cases.

4.4 Results from formal verification runs

The results from different formal verification runs based on the approach discussed above are shown in Table 5. A significant improvement is observed in the run-times of the different properties - many of the properties, which suffered state-space explosions earlier, converged; while, many converged properties from earlier runs report significant reduction in their run-times. We report results from IFV runs on two properties in Table 6.

Property Types	Constraint Applied	No of Properties Passed	CPU Time for Passed Props (hrs)	Avg FF/Latch Count	No of Properties Explored	CPU Time for Explored Props (hrs)
Check for sequence of events during download of algorithms from external ROM.	Set A	69	45	168/0	3	74.67
	Set B	35	40	197/0	3	21.95
	Set C	31	6.66	197/0	4	78.33
All 34 words of an algo are transferred to the targeted registers sequentially.	Set D	38	4	171/0	0	0
Check for events during download of RAM Section Information from external ROM.	Set E	36	17	170/0	1	14.53
	Set F	37	318	170/0	5	71.08
	Set G	8	7.17	174/0	0	0
	Set H	3	33.7	174/0	5	112
Initialization Properties for ROM I/f FSM		13	39	102/0	0	0
Retention Mode		6	1.15	162/0	0	0
MISR Mode		10	7	163/0	0	0

Set A: 5Algos,5BGP,5RAMG,1RAM per RAMG / Set B: 5Algos,5BGP,5RAMG,1RAM per RAMG / Set C: 1Algo,1BGP,1RAMG and no. of RAMs per RAMG unconstrained / Set D: 1Algo, 1BGP, 1RAMG. Number of RAM in each RAMG is unconstrained / Set E: 1Algo, 1BGP, 1RAMG. No. of RAMs per RAMG is unconstrained / Set F: 5Algos, 5RAMG, 1RAM per RAMG. no of BGP unconstrained / Set G: 1Algo, 1BGP, 1RAMG. No of RAM per RAMG unconstrained / Set H: 5Algos, 5BGP, 5RAMG and 1 RAM in each RAMG.

Table 5. Formal verification results from proposed approach.

Property Name	Targeted Functionality	Result before	Result After
ramgroup_start	To check the FSM state transition and other events that are expected during the control transfer to a new RAMG	Passed in 12.7 hrs	Passed in 3.95 hrs
ram_addr_write_str	The last word in the RAMD for each RAM is STR. It is a mnemonic for the start instruction issued to the controller to start the memory testing. This property checks the transfer of this instruction to the corresponding register of the controller and the associated events.	Exploded in 12.7 hrs	Passed in 8.96 hrs

Table 6. Comaprison of Formal Verification Results from different approaches.

4.5 Another useful methodology based on functional compositional verification

While the above proposed approach significantly improved the convergence of the property set needed to verify the ROM interface functional behavior with reduced runtimes, a few properties continued to suffer state-space explosions, as seen from the results presented in Table 5. Fortunately, the convergence issue related to such properties was much simpler to analyse and resolve. The simple startegy of splitting the original property into several smaller sub-properties resolved convergence issues. As an example, consider the property which verifies the sequential transfer of the first 36 words in a ALGO section, to their respective registers in the program register file of the controller. This property took 17 hours in IFV to converge. It was then split into 36 different properties, with each one dedicated to verifying just one word in the sequence of 36 words. This entire set of 36 properties took less than 8 hours to converge.

Functional behaviors involving extremely large sequential depths can pose a formidable challenge to existing automated formal verification approaches. However, analysis of such behavior usually lend themselves to prudent partitions; while these, in most cases suffer from specificity, usually result in convergence of formal verification runs on the partitioned behaviours.

5. Summary and conclusion

To summarize, the key motivation of our approach has been to automate integration verifications of IPs and DFT logic towards, 1) cycle time reduction by a factor of two in the DFT logic verification task by minimizing usage of simulation based chip level verification requirements, 2) improvement in Silicon quality by elimination of all DFT logic and its SoC integration related bugs and 3) deployment of DFT logic generation, its integration in SoC and its verification through a common infrastructure to facilitate re-use of these tasks across different SoC designs. One of the key contributions in the automation of the DFT logic verification task has been the deployment of formal verification techniques, as justified above.

Based on our experience in deploying the proposed approach, good insight has been developed into the DFT verification problem for comparison of simulation based and formal approaches. Experimental data using a commercial formal verification tool IFV [12] show that the proposed approach is an order of magnitude faster than approaches based on simulation. Though we report our results based on IFV, our approach is independent of any FV tool and can work with any FV tool which supports the Property Specification Language (PSL), or the System Verilog Assertion (SVA) language.

6. Acknowledgement

7. References

[1] Subir K. Roy, "Top Level SoC Interconnectivity Verification using Formal Techniques", International Workshop on Microprocessor Test and Verification, Austin, Texas, USA, 2007.

[2] Subir K. Roy and R. A. Parekhji, "Modeling Techniques for Formal Verification of BIST Controllers and Their Integration into SoC Designs", International Conference on VLSI Design, Bangalore, India, 2007.

[3] C. Kern and M. R. Greenstreet, "Formal Verification in Hardware Design: A Survey", *ACM Transactions on Design Automation of Electronic Systems,* Vol. 4, April 1999, pp. 123 - 193.

[4] A. Biere, A. Cimatti, E. Clarke, and Y. Zhu. Symbolic model checking without bdds. In *Tools and Algorithms for Construction and Analysis of Systems, In TACAS'99*, March 1999.

[5] C.J.H. Seger and R.E. Bryant, "Formal Verification by Symbolic Evaluations of Partially – Ordered Trajectories", *Formal Methods in System Design*, 6, 147-189, 1995.

[6] S. Hazelhurst and C. J. H. Seger, "A Simple Theorem Prover Based on Symbolic Trajectory Evaluation and BDDs", *IEEE Transaction on Computer Aided Design of Integrated Circuits*, 14, 4 (April 1994), 413-422.

[7] R.E. Bryant, S. German and M. N. Velev, "Processor Verification Using Efficient Reductions of the Logic of Uninterpreted Functions to Propositional Logic", *Technical Report CMU-CS-99-115*, May 1999, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213.

[8] Yang, J. and Seger, C.: "Introduction to Generalized Symbolic Trajectory Evaluation," *IEEE Trans. on VLSI Systems, 11(3)*, pp. 345-353, 2003.

[9] Jin Yang and C.-J. Seger, "Generalized symbolic trajectory evaluation - Abstraction in action," *LNCS: Proc. Of FMCAD2002,* November 2002.

[10] Jin Yang, "GSTE: An illustrative and comparative introduction," *5th International Conference on ASIC,* Volume 1, pp. 41-44, October 2003.

[11] 1-Team-Genesis – Tool for Architecture Generation, Atrenta Inc., 2008.

[12] IFV – Incisive Formal Verification Tool, Cadence Design Systems Inc., 2009.

[13] Bill Murray, "Mixing Formal and Dynamic Verification – Part 1 and 2", Special Technology Report, SCDsource, http://www.scdsource.com/article.php?id=333 and 341, 2009.

Test Generation based on CLP

Giuseppe Di Guglielmo, Franco Fummi,
Cristina Marconcini and Graziano Pravadelli
University of Verona
Italy

1. Introduction

The complexity of designs continues to rise, driven by technology advances, while time-to-market imposes always shorter time. Moreover, the increasing of design complexity implies that design verification becomes one of the most cost-dominating phase in design production.

Functional Automatic Test Pattern Generators (ATPGs) based on simulation (Corno et al., 2001; Fin & Fummi, 2003a) are fast, but generally, they are unable to cover corner cases, and they cannot prove untestability. On the contrary, functional ATPGs exploiting formal methods (Ghosh & Fujita, 2001; Zhang et al., 2003; Xin et al., 2005a), are exhaustive and cover corner cases, but they tend to suffer of the state explosion problem when adopted for verifying large designs. In this context, a functional ATPG is presented, that relies on the joint use of pseudo-deterministic simulation and constraint logic programming (CLP) (Jaffar & Maher, 1994), to generate high-quality test sequences for solving complex problems. Thus, the advantages of both simulation-based and static-based verification techniques are preserved, while their respective drawbacks are limited. In particular, CLP, a form of constraint programming in which logic programming is extended to include concepts from constraint satisfaction, is well-suited to be jointly used with simulation. In fact, information learned during design exploration by simulation can be effectively exploited for guiding the search of a CLP solver towards design under verification (DUV) areas not covered yet.

Therefore, this work is focused on the use of CLP for addressing corner cases during functional test pattern generation. In particular, a CLP-based fault-oriented ATPG engine is proposed to be adopted, after simulation, learning and random-walk/backjumping, as the last step of the incremental test generation flow showed in Figure 1.

According to such a flow, the ATPG framework is composed of three functional ATPG engines working on three different models of the same DUV: the hardware description language (HDL) model of the DUV, the set of concurrent EFSMs extracted from the HDL description, and the set of logic constraints modelling the EFSMs. The EFSM paradigm has been selected since it allows a compact representation of the DUV state space (Lee & Yannakakis, 1992) that limits the state explosion problem typical of more traditional FSMs.

In the proposed framework, the first engine is random-based, the second is transition-oriented, while the last is fault-oriented. This approach quickly covers the greater part of DUV faults, typically easy-to-detect faults. Then, the transition-oriented engine and fault-
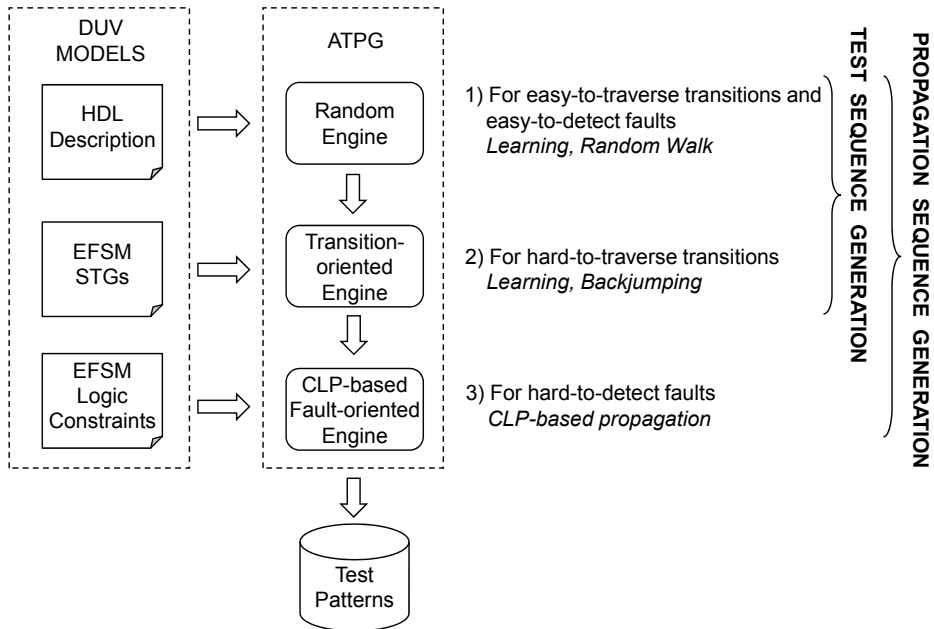
Fig. 1. The incremental test generation flow.

oriented engine permit to deterministically generate test patterns for the remaining uncovered hard-to-detect faults. Therefore, the test generation is guided by means of transition coverage (Li & Wong, 2002) and fault coverage (Abramovici, 1993). In particular, 100% transition coverage is desired as a necessary condition for fault detection, while the bit coverage (Ferrandi et al., 1998a) functional fault model is used to evaluate the effectiveness of the generated test patterns by measuring the related fault coverage.

This Chapter is organized as follow. Section 2 reports the state of art of CLP techniques applied for test generation. Section 3 describes the EFSM as the adopted computational and the strategies defined for generating a particular type of EFSM suited for test generation. Section 4 presents the functional ATPG engine that relies on learning, random-walk/backjumping and constraint logic programming to deterministically generate test vectors for traversing all transition of the EFSM and activate the fault injected into the DUV. Section 5 describes the fault-oriented ATPG engine that exploits CLP techniques to propagate hard-to-detect faults activated but not observed to the outputs. Experimental results are presented in Section 6, and conclusions are drawn in Section 7.

2. Background

Constraint programming (Wallace, 1997) is a paradigm that is tailored to search problems. The main application areas are those of planning, scheduling, timetabling, routing, placement, investment, configuration, design and insurance. Constraint programming incorporates techniques from mathematics, artificial intelligence and operations research, and it offers significant advantages in these areas since it supports fast program development, economic program maintenance, and efficient runtime performance.

Constraint logic programming (CLP) combines logic, which is used to specify a set of possibilities explored via a very simple in-built search method, with constraints, which are used to minimize the search by eliminating impossible alternatives in advance.

The programmer can state the factors which must be taken into account in any solution (the constraints), state the possibilities (the logic program), and use the system to combine reasoning and search. The constraints are used to restrict and guide search. The whole field of software research and development has aims to optimize the task of specifying, writing and maintaining correct functioning programs.

CLP-based ATPGs have been already proposed in the literature (Vemuri & Kalyanaraman, 1995; Pauli et al., 2000; Xin and Harris, 2002; Ferrandi et al., 2002b), however, existent approaches differ in several aspects from the ones presented in this work. In (Pauli et al., 2000), CLP is used to generate test sequences according to a path coverage-based criterion. However, this approach is oriented only to the control part of circuits. Control flow paths are also the target of the approach presented in (Vemuri & Kalyanaraman, 1995; Ferrandi et al., 2002b). In (Vemuri and Kalyanaraman, 1995), constraints are derived from a preprocessing of a VHDL description to enumerate all the target paths. On the contrary, in (Ferrandi et al., 2002b), constraints are generated by enumerating paths of concurrent FSMs describing the DUV. However, path enumeration is a very hard and time-consuming task, since paths of sequential circuits are generally infinite. In (Xin and Harris, 2002), the authors propose to use CLP for generating test sequences targeting synchronization/timing faults in hardware/software models described as a network of co-design FSMs. This work identifies sequences to trigger synchronization faults but the observability of the fault effect is not considered.

Finally, the previous approaches propose neither strategies for completely modelling a design by means of CLP (just some paths are modelled), nor approaches for avoiding the risk of state explosion when a CLP solver is asked to generate a test sequence to activate the target path.

The CLP solver adopted and integrated into the proposed ATPG is ECLiPSe (ECLiPSe Common Logic Programming System) (Wallace & Veron, 1994; Apt & Wallace, 2007). ECLiPSe is a Prolog based system whose aim is to serve as a platform for integrating various Logic Programming extensions, in particular Constraint Logic Programming. The kernel of ECLiPSe is an efficient implementation of standard (Edinburgh-like) Prolog as described in basic Prolog texts. It is built around an incremental compiler which compiles the ECLiPSe source into WAM-like code, and an emulator of this abstract code. ECLiPSe is now an open-source project, with the support of Cisco Systems.

3. EFSM manipulation to exploit CLP based techniques

The EFSM model is a generalization of the classical FSM model that provides a compact representation of local data variables and preserve properties of the traditional state machine models.

In the current work, a digital system is represented as a set of concurrent EFSMs, one for each process of the DUV. In this way, according to the following Definition 1, the main characteristics of state-oriented, activity-oriented and structure-oriented models are captured (Gajski et al., 1997). In fact, the EFSM is composed of states and transitions, thus it is state-oriented, but each transition is extended with hardware description language (HDL) instructions that act on the DUV registers. In this sense, each transition represents a set of

activities on data, thus, the EFSM is a data-oriented model too. Finally, concurrency is intended as the possibility that each EFSM of the same DUV changes its state concurrently to the other EFSMs to reflect the concurrent execution of the corresponding processes. Data communication between concurrent EFSMs is guaranteed by the presence of common signals. In this way, structured models can be represented.
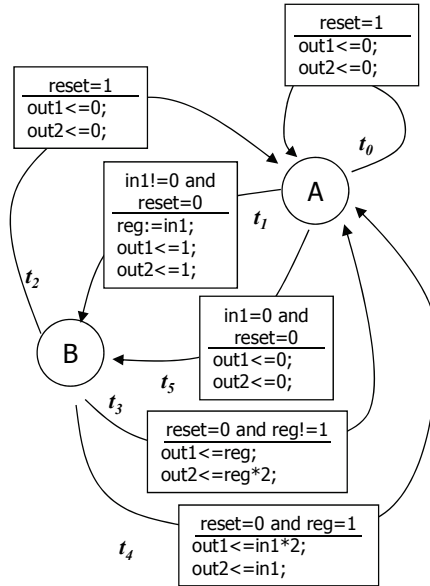
Fig. 2. State transition graph of an EFSM.

Definition 1 *An EFSM is defined as a 5-tuple M=<S,I,O,D,T> where: S is a set of states, I is a set of input symbols, O is a set of output symbols, D is a n-dimensional linear space $D_1 \times \ldots \times D_n$, T is a transition relation such that $T:S \times D \times I \rightarrow S \times D \times O$. A generic point in D is described by a n-tuple $x=(x_1,\ldots,x_n)$ and models the values of the registers of the DUV. A pair $<s,x> \in S \times D$ is called configuration of M.*

An operation on *M* is defined in this way: if *M* is in a configuration $<s,x>$ and it receives an input $i \in I$, it moves to the configuration $<t,y>$ iff $((s,x,i),(t,y,o)) \in T$ for $o \in O$.

The EFSM differs from the classical FSM, since each transition does not present only a label in the classical form *(i)/(o)*, but it takes care of the register values too. Transitions are labelled with an enabling function *e* and an update function *u* defined as follows.

Definition 2 *Given an EFSM M=<S,I,O,D,T>, $s \in S$, $t \in T$, $i \in I$, $o \in O$ and the sets $X=\{x \mid ((s,x,i),(t,y,o)) \in T$ for $y \in D\}$ and $Y = \{y \mid ((s,x,i),(t,y,o)) \in T$ for $x \in X\}$, the enabling and update functions are defined respectively as:*

$$e\,(x,i) = \begin{cases} 1 & \text{if } x \in X; \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

$$u(x,i) = \begin{cases} (y,o) & e(x,i) = 1 \text{ and } ((s,x,i),(t,y,o)) \in T; \\ \text{undef.} & \text{otherwise.} \end{cases}$$

An update function $u(x,i)$ can be applied to a configuration $<s_1,x>$ if there is a transition $t:s_1 \rightarrow s_2$, labelled e/u, such that $e(x,i)=1$. In this case we say that t can be *traversed* by applying the input i. Figure 2 shows the state transition graph of a simple EFSM.

Many EFSMs can be generated starting from the same HDL description of a DUV. However, despite from their functional equivalence, they can be more or less easy to be traversed. Easiness of traversal is a mandatory feature for a computational model used in CLP-based test pattern generation, and it is a desirable condition to activate and propagate faults.

Stabilization of EFSMs improves the easiness of traversal (Lee & Yannakakis, 1992), but it can lead to the state space explosion. Thus, in (Di Guglielmo et al., 2006a), a set of theoretically-based automatic transformations has been proposed to generate a particular kind of semi-stabilized EFSM (S^2EFSM). This particular kind of EFSM allows the ATPG to easily explore the state space of the corresponding DUV reducing the risk of state explosion. The S^2EFSM presents the following characteristics:

- It is functionally and timing equivalent to the HDL description from which it is extracted, i.e., given an input sequence, the HDL description and the corresponding S^2EFSM provide the same result at the same time.
- The update functions contain only assignment statements. This implies that all the control information, needed by a CLP-based ATPG to traverse the DUV state space, resides in the enabling functions of the S^2EFSM.
- The S^2EFSM is partially stabilized to reduce the state explosion problem that may arise when stabilization is performed to remove inconsistent transitions. Only transitions not leading to state explosion are stabilized.

The S^2EFSM is referred simply as EFSM in the following.

4. CLP-based technique for generating test sequences

The ATPG engine proposed in this Section relies on random and pseudo-deterministic simulation and CLP techniques to generate test sequences for traversing the system represented as a collection of EFSM. A two step approach, implementing respectively the random engine and the transition-oriented engine, is depicted in Figure 1.

First the DUV state space is explored by performing a simulation-based random-walk (Section 4.3). This allows to quickly fire easy-to-traverse (ETT) transitions and, consequently, to quickly cover easy-to-detect (ETD) faults. However, the majority of hard-to-traverse (HTT) transitions remain, generally, uncovered.

Thus, the backjumping-based strategy is applied to cover the remaining HTT transitions by mean of transition-oriented ATPG (Section 4.4). Backjumping, also known as non-chronological backtracking, is a special kind of backtracking strategy which rollbacks from an unsuccessful situation directly to the cause of the failure. Thus, the engine deterministically backjumps to the source of failure when a transition, whose guard depends on previously set registers, cannot be traversed. Next it modifies the EFSM configuration to satisfy the condition on registers and successfully comes back to the target state to activate the transition.

The transition-oriented engine generally allows achieving 100% transition coverage. However, 100% transition coverage does not guarantee to explore all DUV corner cases, thus some hard-to-detect (HTD) faults can escape detection preventing the achievement of 100% fault coverage. Therefore, the CLP-based fault-oriented engine, as described in Section 5, is finally applied to focus on the remaining HTD faults.

4.1 ATPG architecture

The EFSM model is specially suited to be used with CLP-based ATPGs that generate test sequences by deterministically activating the enabling functions of the transitions. According to this observation, in this section the functional ATPG framework depicted in Figure 3 is described. The framework is composed of two main modules: the DUV-dependent component generator (DCG) and the run-time engine (RTE).

Given an HDL functional description of the DUV, the DCG modules generates the state-transition-graph (STG) representations of the corresponding EFSMs (EFSM STG Generator), the faulty description of the DUV and the related fault list (Fault Injector), and the file containing the constraints involved in the EFSM enabling and updating functions (Constraint Generator). The constraint generation for transition-oriented CLP-based ATPG is described in Section 4.2

The RTE module is composed of the EFSM navigator, the CLP Interface, and the Simulation Engine. The RTE navigates the STG representation of the EFSM to generate test sequences. An external CLP solver (ECLiPSe) is used to generate values for primary inputs of the DUV which allow firing the enabling functions of transitions that the EFSM navigator wants to traverse. Then, the generated test sequences are provided to the Simulation Engine, and the behaviour of the fault-free and faulty DUVs is compared. Test sequences that highlight discrepancies between the primary outputs of the fault-free and faulty DUVs constitute the final test patterns.

4.2 Constraint generation for transition-oriented CLP-based ATPG

Starting from the in-memory representation of the DUV, the Constraint Generator automatically creates the CLP-constraint file to allow the RTE module to evaluate the enabling functions when the EFSM is navigated. For example, Figure 4 shows the constraint representation for ECLiPSe solver related to transitions of the EFSM of Figure 2.
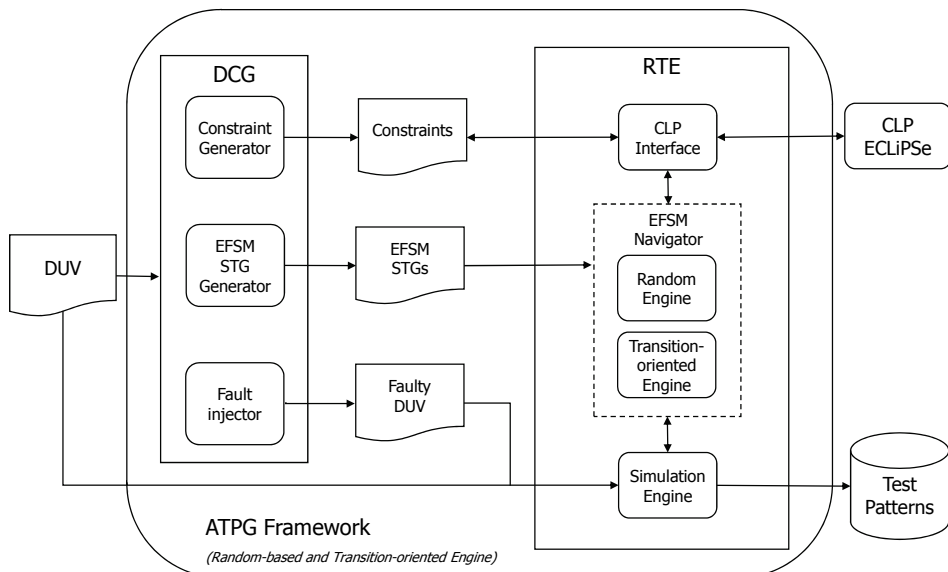
Fig. 3. The ATPG framework.

During the test pattern generation such constraints are exploited. In particular, each test vector is randomly initialized. Then, it is modified accordingly with the values provided by the CLP solver. In particular, if the enabling functions of the currently evaluated transition are satisfied, the part of the test vector related to the primary inputs involved in the enabling function is modified accordingly to the values returned by the CLP solver.

Moreover, the EFSM Navigator accesses to structures pointing to DUV internal registers. The actual values of the registers are used to instantiate the constraints in the current system status.

```
t0 :- reset::0..1,
    (reset#=1),
    indomain(reset,random).
t1 :- lb is -(2^32), rb is (2^32)-1,
    in1::lb..rb, reset::0..1,
    (in1#\=0)and(reset#=0),
    indomain(in1,random), indomain(reset,random).
t2 :- reset::0..1,
    (reset#=1),
    indomain(reset,random).
t3 :- lb is -(2^32), rb is (2^32)-1,
    reg::lb..rb, reset::0..1,
    (reg#\=1)and(reset#=0),
    indomain(in1,random), indomain(reset,random).
t4 :- lb is -(2^32), rb is (2^32)-1,
    reg::lb..rb, reset::0..1,
    (reg#=1)and(reset#=0),
    indomain(in1,random), indomain(reset,random).
t5 :- lb is -(2^32), rb is (2^32)-1,
    in1::lb..rb, reset::0..1,
    (in1#=0)and(reset#=0),
    indomain(in1, random), indomain(reset,random).
```

Fig. 4. Enabling function representations for ECLiPSe solver.

4.3 Random-walk

During the random-walk phase, the ATPG randomly walks across the transitions of the EFSMs representing the DUV. Thus, ETT transitions are very likely traversed.

Starting from a reset condition, the ATPG randomly selects a transition from each EFSM according to a scheduling policy. The EFSM-scheduling policy aims at maximizing the ATPG capability of exploring the whole state space. Then, it tries to satisfy the enabling function of each selected transition by exploiting the CLP solver invoked by providing it with the corresponding constraints. When it succeeds, the values for the primary inputs, provided by the solver, are used to generate a test vector. Finally, a simulation cycle is performed, by using the generated test vector, to update the internal registers of the DUV and to move to the destination state. Then, another transition is selected, and the cycle repeats.

Each time a test vector is generated, the traversed transition is labelled with the test sequence number and the test vector number. A list of pairs of parametric length is saved

for each transition. In this way, the backjumping mode can exploits such lists to quickly recover the prefix of a test sequence which allows the ATPG to move from the reset state to an already visited target state.

4.4 Backjumping

The ATPG automatically changes to the backjumping mode when the computation time assigned to the random-walk expires, or no coverage improvement is provided for long time. Thus, the transition-oriented ATPG works as represented in Figure 5. Let us assume t is a not fired transition, out-going from state S_t already visited during the random-walk phase. Let us also assume that the unsatisfiability of the enabling function of t depends on clauses involving a single register reg. If the unsatisfiability of t depends on more than one register, the backjumping procedure is repeated for each of them. Then extract an already visited transition t_u from the set of transitions T_u^{reg} whose update function updates reg. Load the test sequence, previously generated during random-walk mode, to move from the reset state to S_{tu} (source state of t_u). Thus, the ATPG backjumps from S_t to S_{tu}. Use the Dijkstra's shortest path search algorithm to provide a path π from S_{tu} to S_t starting with t_u. Satisfy the enabling function of t_u according to the constraints derived from the enabling function of t as follows. Let us suppose that e_{tu} is the enabling function of t_u and $e_t|reg_{tu}$ is the part of the enabling function of t which involves the clauses depending on reg, where each occurrence of reg has been substituted with the right-side expression of the assignment that updates reg in the update function of t_u. Invoking the CLP solver to satisfy the constraint $e_{tu} \wedge e_t|reg_{tu}$ allows us to obtain a test vector which satisfies e_{tu} and sets the value of reg in such a way that when simulation reaches transition t, following π, its enabling function will be correctly fired. The last observation may be false if there is a transition $t'_u \neq t_u \neq t$ in π, such that t'_u updates reg after t_u did. In this case, the ATPG moves the problem from t_u to t'_u requiring a solution for $e_{t'u} \wedge e_t|reg_{t'u}$. Finally satisfy the enabling function of transitions included in π by iteratively applying the constraint solver to generate the corresponding test vectors. The test sequence obtained by joining s, to move from the reset state to S_{tu}, and the test vectors generated to traverse π allows to fire t.
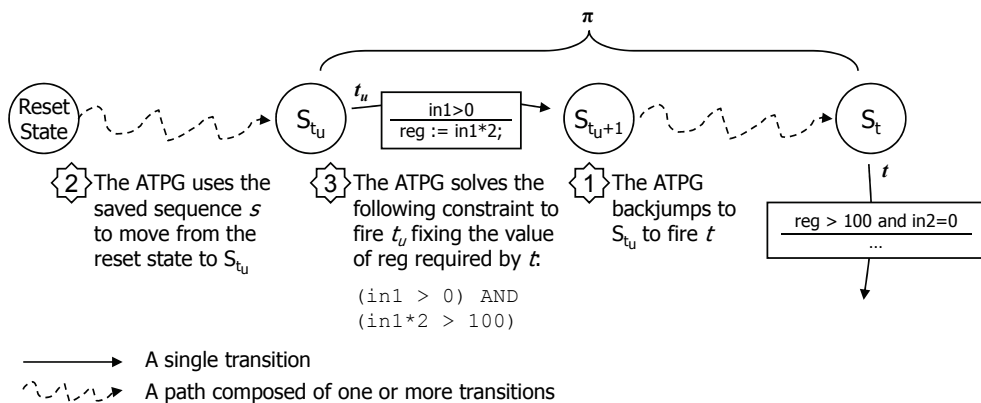
Fig. 5. The backjumping strategy.

5. CLP-based technique for generating propagation sequences

In this section the CLP solver is used to deterministically search for sequences that propagate the faults observed, but not detected by means of the previously presented random and transition-oriented engines. In particular, Section 5.1 presents the technique defined to model the EFSM in CLP, Section 5.2 describes the implemented CLP-based fault-oriented ATPG engine, and Section 5.3 presents some optimization strategies defined to deal with the complexity typical of static techniques like CLP.

5.1 EFSM modelling by CLP

At first, the concept of time steps is introduced, required to model the EFSM evolution through time via CLP. Then, techniques are presented to model logical variables and constraints describing the enabling functions and update functions of the EFSM.

Hardware description languages easily allow modelling the DUV time evolution by means of processes, implicit or explicit wait statements, sensitivity lists and events. On the contrary, CLP does not provide an explicit mechanism to model the time. To overcome this limitation, a logical variable N is introduced to represents the total number of time frames on which the EFSM can evolve. The domain of N is *[1,Max]*, where *Max* is defined according to the sequential depth of the DUV. Then, the CLP variables, used to model the EFSM behaviour, are defined as arrays of size N. Thus, for example, let us consider a variable V defined in the HDL description of the DUV. When CLP is adopted, an array *V[]* is used to model the evolution in time of variable V. In this context, a CLP constraint of the form *V[T]#=0* indicates that at time T the variable V of the DUV has value *0*.

Three kinds of arrays of logical variables must be defined to describe, respectively, states, transitions, and registers of an EFSM.

- Each state of the EFSM is modelled by an array of boolean variables of size N. When the EFSM is in the state S at time T, the Tth element of the array *S[]*, corresponding to the state S, is assigned to *true*. For example, two arrays, *A[]* and *B[]*, are required to model states A and B of the EFSM of Figure 2. At every time step T, either *A[T]* or *B[T]* is *true*, thus indicating the current state of the EFSM.

- Each transition of the EFSM is associated to an array of boolean variables of size N. If a transition is fired at time T, the Tth element of the array corresponding to the transition is assigned to *true*.

- Registers are modelled as array of size N respecting their original data type.

The CLP code of Figure 6 exemplifies the constraints required to model logical variables for states, transitions, and registers of the EFSM of Figure 2. Moreover, in Figure 6, the predicate *bool(X,N)* defines the boolean data type used to model states and transitions. It means that the logical variable X is an array of size N (for modelling of time), whose elements can assume the values included in the list *Xlist*, i.e., *0* or *1*. In a similar way, the predicate *int32(X,N)* defines an arrays of N 32-bit integers used as data type to deal with primary inputs, primary outputs, and registers.

The functional behaviour of the EFSM is represented by means of enabling functions and update functions labelling the transitions between states. Thus finally, a way for modelling such functions and their relation with states and transitions is proposed. In particular, two kinds of constraints have been defined to model the current state of the EFSM, and the relation between the enabling function and the corresponding update function.

```
% data types
bool(X,N) :- dim(X,[N]),term_variables(X, Xlist), Xlist::[0,1].
int32(X,N) :- dim(X,[N]), X[1..N]:: -2147483648..2147483647.
% states
bool(A,N), bool(B,N),
% transition
bool(T0,N), bool(T1,N), bool(T2,N), bool(T3,N), bool(T4,N),bool(T5,N),
% primary inputs, primary outputs, and registers
int32(IN1,N), int32(REG,N), int32(OUT1,N), int32(OUT2,N).
```

Fig. 6. Constraints for modelling state, transition and register variables.

5.1.1 Current state modelling

Two constraints must be defined for each array of state variables to specify the current state of the EFSM. The first constraint specifies that, at each time step T, the T th element of an array $S[]$, modelling a state S of the DUV, is *true*, if and only the Tth element of one of the transition arrays corresponding to the transitions in-going in S is *true*. The second constraint specifies the dual situation, i.e., if the Tth element of the transition array is true at time T, then the Tth element of the array associated to the destination state of the corresponding transition must be true at time step $T+1$ (*NEXT_T*). For example, let us consider the EFSM of Figure 2. The constraints in Figure 7 must be defined for specifying that the current state of the EFSM at time step $T+1$ is A , if and only if one of the transitions in-going in A has been fired at time T.

```
A[NEXT_T] #= T0[T] xor T2[T] xor T3[T] xor T4[T],
T0[T] xor T2[T] xor T3[T] xor T4[T] => A[NEXT_T].
```

Fig. 7. Constraints for modelling next-state relation.

Finally, a further constraint is introduced to explicitly force the system to be in a single state and transition at each time step. Thus, the Tth element of arrays corresponding to states of the EFSM are put in *xor* each other as shown in Figure 8.

```
A[T] xor B[T],
T0[T] xor T1[T] xor T2[T] xor T3[T] xor T4[T] xor T5[T].
```

Fig. 8. Constraint for modelling transition and state mutual exclusion.

In fact, at a particular time step, only one transition of the EFSM can be traversed and, obviously, the EFSM can have only one state active. Designers implicitly include such a constraint, when they model the DUV by means of an HDL. However, the explicit presence of such a constraint, when the EFSM is provided to the CLP solver, allows the solver to immediately prune the solution space by ignoring configurations where more than one state variable is concurrently true, thus drastically reducing the number of backtracking steps.

5.1.2 Enabling and update function modelling

Firing a transition at time T implies that its enabling function is satisfied at time T, its update function is executed at time T, and the state of the EFSM at time T is the source of the transition. Thus, for example, if Ti is a transition out-going from state S, whose enabling function and update function are modelled, respectively, by the predicates EF and UF (described later), the constraints in Figure 9 are used to model Ti.

EF[T] and S[T] => Ti[T],
Ti[T] => EF[T] and S[T],
Ti[T] and (EF[T] and S[T]) => UF[T].

Fig. 9. Constraint for correlating the enabling and update functions to transitions.

The first two constraints represent a double implication for imposing that the transition variable $Ti[T]$ is *true* (i.e., the transition Ti is fired at time T) if and only if the predicate of the corresponding enabling function $EF[T]$ and the variable $S[T]$, associated to the state from which Ti is out-going, are *true*. On the contrary, the predicate of the update function $UF[T]$ does not require a double implication, because it is possible that $UF[T]$ is *true* even if $EF[T]$ is *false*. However, in this case the transition is not fired and the update function is not executed. The predicate $EF[T]$ is directly derived from the condition involved in the corresponding enabling function. Its modelling requires only a syntactical conversion from the syntax of the HDL used to model the DUV towards the syntax accepted by the CLP solver.

On the contrary, modelling the predicate $UF[T]$ associated to an update function requires more attention. In particular, an update function involves assignments to registers and primary outputs. Let us use an example to show how to model such a kind of statement. Consider, for example, the statement $SIG := SIG + IN$, where SIG is an internal signal and IN is a primary input. The corresponding CLP constraint is $SIG[Next\_T]\#=SIG[T]+IN[T]$.

However, registers and primary outputs, that do not require to be updated, are not assigned in the update function when a design is modelled by using an HDL. Indeed, they implicitly preserve their previous value. Unfortunately, the CLP solver assigns random values to variables that are not explicitly assigned. Thus, when an update function is modelled by means of constraints, it has to ensured that a constraint is explicitly added to preserve the value of signals, registers and primary outputs that do not require to be updated.

According to the previous rules, for example, the transition t_4 in Figure 2 is modelled as depicted in Figure 10.

% Enabling function
((REG[T] #=1) and B[T]) => T4[T],
T4[T] => ((REG[T] #= 1) and B[T]),
% Update function
(T4[T] and ((REG[T] #= 1) and B[T]) =>((REG[Next_T] #= REG[T]) and
(OUT1[Next_T] #= IN1[T]*2) and (OUT2[Next_T] #= IN1[T])).

Fig. 10. Constraint for representing transition t_4 of the EFMS in Figure 2.

5.2 Fault-oriented CLP-based engine

The transition-oriented engine described in Section 4 pseudo-deterministically generates sequences for firing HTT transitions on EFSMs. In this way, the majority of faults are detected as a consequence of transition traversal, but some HTD faults can remain uncovered. On the contrary, the CLP-based fault-oriented engine exhaustively searches for test sequences targeting specific faults. It exploits the CLP-solver to explore the CLP-based representation of the DUV extracted from the EFSM model. The exhaustiveness, guaranteed by CLP, is paid in terms of execution time, but such an engine is applied to a small number of faults: those not detected neither by the random-based engine nor by the transition-oriented one.

Let us consider a fault f that has not been detected yet by these engines. This may depends on two different reasons:

1. the ATPG has been unable to find an activation sequence, i.e., in the case of the bit coverage fault model, a sequence that causes the bit (or the condition) affected by f to be set with the opposite value with respect to the one induced by f;
2. the ATPG activated f, but it has been unable to find a propagation sequence, i.e., a sequence that propagates the effect of f to the primary outputs of the DUV.
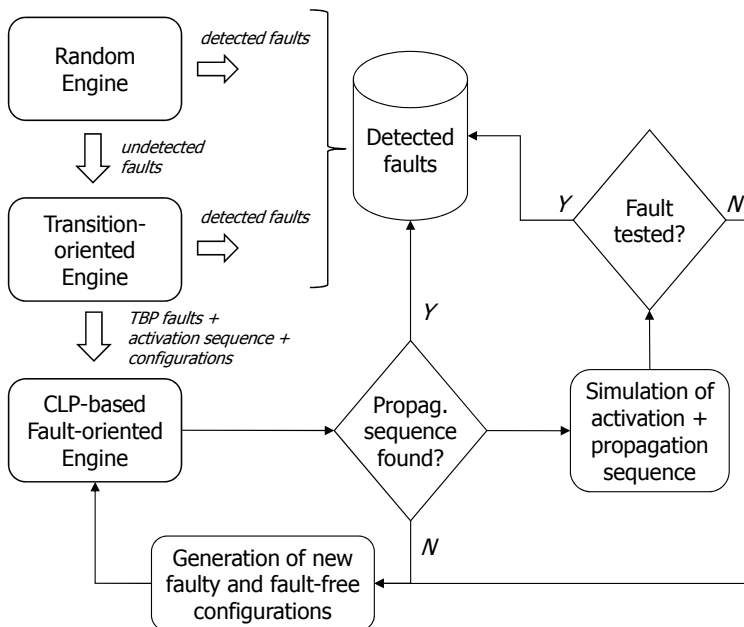
Fig. 11. The role of the CLP-based fault-oriented engine.

To distinguish between the previous alternatives, the ATPG observes the effect of each fault on both primary outputs and internal registers, during the simulation of test sequences generated by the random-based and transition-based engines. From this observation, the following well-known definitions derive.

Definition 3 *A fault f is said to be observable on primary outputs, i.e., detectable, if there exists a test sequence s such that, by concurrently applying s to the faulty and the fault-free DUVs, the value*

of at least one primary output in the fault-free DUV differs from the value of the corresponding primary output in the faulty DUV, at least once in time.

Definition 4 *A fault f is said to be observable on a register if there exists a test sequence s such that, by concurrently applying s to the faulty and the fault-free DUVs, the value of at least one register in the fault-free DUV differs from the value of the corresponding register in the faulty DUV, at least once in time.*

According to the previous definitions, if the fault is observed on primary outputs, it is marked as detected and the corresponding test sequence is saved. Otherwise, if the fault is observed only on registers, the fault is marked as to be propagated (TBP). Finally, if the fault is observable neither on primary outputs nor on registers, this is due to the difficulty of finding an activation sequence. Thus, the fault is marked as to be activated (TBA).

The current work addresses TBP faults (see Figure 11); future works will address the problem of TBA faults. In the following, Section 5.2.1 describes how submit TBP problem to the CLP-solver, Section 5.2.2 introduces searching functions to generate test sequences and finally Section 5.2.3 proposes how to managing problem complexity.
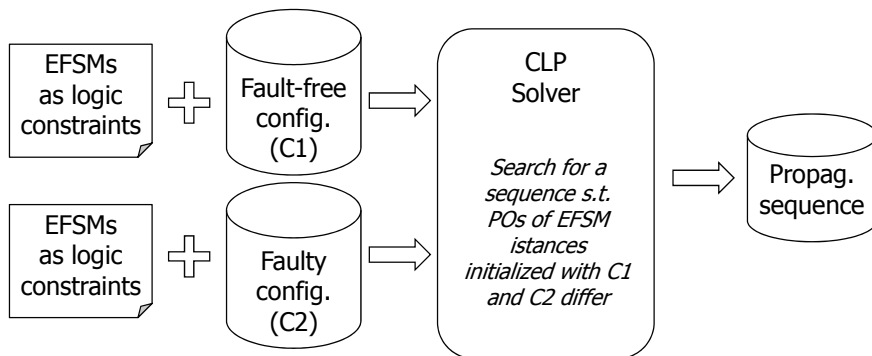
Fig. 12. Use of the CLP solver for finding propagation sequences.

5.2.1 Propagation sequence generation

The propagation sequence for a TBP fault is generated by providing the ATPG engine with two instances of the CLP-based representation of the DUV. This representation consists of the CLP model of the generated EFSM. The instances are initialized with the EFSM configurations (the faulty and the fault-free ones) that allow the random or the transition-based engine to observe the fault on at least one register.

Note that the two DUV instances, provided to the CLP solver, are exactly equal, but their registers are initialized with different values (the faulty and the fault-free ones). This is the only reason such instances should behave in different ways. In the following, the terms faulty and fault-free are used to distinguish the DUV instances initialized, respectively, with the faulty and the fault-free configuration. Such configurations consist of the values of registers (included the value of the state register) in the faulty and fault-free DUVs at the moment the fault has been observed. As an example, consider the constraints in Figure 13. They state that, at time $T=1$, *REG* has the same value (i.e., *5941*) in both the faulty and fault-free DUV instances, but the two EFSMs are in different states (i.e., the fault-free DUV is in state A, while the faulty DUV is in state B).

```
% Fault free
REG[1] #= 5941, A[1] #= 1, B[1] #= 0,
% Faulty
REG_F[1] #= 5941, A_F[1] #= 0, B_F[1] #= 1.
```

Fig. 13. How to specify faulty and fault-free configuration for EFSM models.

After the set-up of the faulty and fault-free configurations, the CLP-solver is asked to find a sequence that, starting from such configurations, propagates the effect of the fault towards the primary outputs (see Figure 12). Therefore, a constraint is defined that forces the arrays of primary outputs to be different at least in a position as shown in Figure 14.

```
˜((OUT1 = OUT1_F) and (OUT2 = OUT2_F)).
```

Fig. 14. Asking for a propagation sequence.

If the solver finds a solution, it consists of a propagation sequence that can be appended to the activation sequence, previously generated by the random-based or transition-based engine to observe the target fault on the internal registers. The so obtained sequence is very likely a test sequence for the target fault, but this must be definitely proved via simulation (see Figure 11). In fact, the propagation sequence is generated by initializing a fault-free instance of the DUV with a faulty configuration, which is not the same as using a real faulty DUV instance directly affected by the fault. However, experimental results showed that in very few cases the propagation sequence generated by the CLP solver according to the proposed strategy, fails to propagate the corresponding fault when it is simulated on the faulty DUV.

5.2.2 Definition of search procedures

Constraints described in the previous subsection are used to set up the problem of finding a propagation sequence as a CLP problem. Then, some search procedures, that exploit search strategies and heuristics, must be defined to force the solver to provide the solution (i.e., the set of values to be assigned to the logic variables for satisfying all the problem's constraints), when it exists. Therefore, a predicate that exploits the *search/6* function of the ECLiPSe's IC library have been defined (Figure 15).

```
search_func(A), search_func(B), search_func(IN1),..
search_func(L):- search(L,0,input_order,indomain,complete,[]).
```

Fig. 15. A search procedure is defined for each variable of the DUV.

Such a function, whose signature is *search(Vars, Arg, Select, Choice, Method, Options)*, is a generic search routine which implements different partial search methods. It instantiates the variables *Vars* by performing a search based on the parameters provided by the user. In our case the search method performs a complete search routine which explores all alternative choices for each variable. The choice method *indomain* tries to find a solution by analyzing

the variable values in increasing order, from the lower value in the variable range to the upper value. The predicate *search_func* is called on each variable of the DUV. In this way, if a solution exists, the solver provides a value for each variable for each time step, thus generating the required propagation sequence.

5.2.3 Managing the CLP complexity

Tools that exhaustively search for a solution of NP-hard problems frequently run out of resources when the state space to be analyzed is too large. The same happens for the CLP solver, when it is asked to find a propagation sequence on large sequential designs. To limit such a problem, heuristics is generally used for pruning the state space. However, this may prevent the solver from finding a solution (even if it exists), if the pruning is too restrictive. Thus, choosing a good heuristics is a very challenging task.

In this context, three strategies for managing the complexity of the CLP solver exploited by the fault-oriented engine have been defined. Two of them have been already presented in previous sections; however, for convenience of the reader, we summarize them here:

- The Tth element of arrays corresponding to states (and transitions) of the DUV are put in *xor* each other, to avoid that the solver wastes time to analyze configurations where more than one state variable is concurrently true. This drastically decreases the number of backtracking steps, especially for designs with many states and many registers.

- A constraint on DUV registers is defined to assure that at least one register of the faulty DUV differs from the corresponding register of the fault-free DUV at each time step $T>1$. On the contrary, the search is immediately stopped, and no solution is reported. Such a constraint avoids situations where the solver spends uselessly efforts, as it cannot lead to the observability on primary outputs if starting from different configurations, the faulty and fault-free DUVs evolve in the same configuration.

A further strategy for managing the complexity of the CLP solver, that can be jointly used with the previous ones, consists of asking the solver to find a solution (i.e., a propagation sequence) starting with a small state space, that is incrementally enlarged until a solution is found (or execution time expires). Thus, the state space to be analyzed by the solver is restricted by limiting the range of the DUV PIs, similarly to what has been proposed for limiting the size of binary decision diagrams in the test generation strategy proposed in (Ferrandi et al., 2002a). At the beginning, the ATPG statically fixes the values of all bits, but two, for each PI. In this way, only two bits can be changed by the CLP solver during the search, independently from the PIs range declared on the HDL description of the DUV. Then, the solver is asked to find a solution. If it fails, the ATPG opportunely increases the number of free bits of PIs. In particular, the ATPG engine searches for the constraints that induce the failure, and it frees the bits of the PIs involved in such constraints. Then, a new search session is launched. Such a process is iterated until a solution is found or execution time expires.

5.3 Optimizations exploiting EFSM model features

This section describes some heuristics to reduce the complexity problem for the solver. Two different approaches have been defined to improve the CLP-based ATPG engine by reducing the problem complexity.

The first approach is based on the EFSM manipulation as described in Section 5.3.1. A new EFSM is generated for the solver that has to deal with a reduced version. This technique exploits two phases. In the first phase, all the transitions that delete the observability

property of the given configuration are removed. Then, all that parts of the EFSM that cannot be traversed are eliminated. This strategy removes all that constraints that are useless and could also avoid the propagation of observability on the primary outputs. In fact the new generated EFSMs are pruned by all that transitions are not needed to model any possible behaviour that can lead to fault observability. Note that the complexity of the proposed algorithm is linear with the number of transitions and that, above all, the number of transition is limited and is not affected by the state explosion problem with the EFSM model.
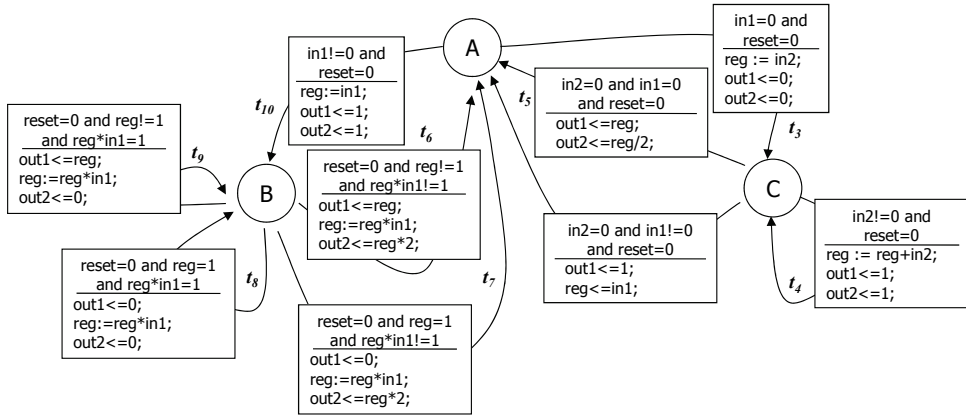
Fig. 16. Non-optimized EFSM example.

```
// let E be an EFSM and let ob_reg be the observed register
reduce_efsm (register ob_reg, efsm E) {
  //retrieve the sets of transitions and states of E
  T = transitions_set(E);
  S = states_set(E);
  for each transition t in T {
    // delete a transition if the conditions that preclude
    // observability are matched
    if (write_on_register(t, ob_reg) and not(read_register(t, ob_reg)))
      delete_transition(t);
  }
}
delete_transition(transition t) {
  // retrieve the in-going state for transition t
  state s= in_going_state(t);
  // to delete a state, it must be different from the state
  // of the current configuration and there are no other
  // transitions, except the current one t, going into it
  if (not(is_configuration_state(s)) and (number_of_ingoing_transitions(s) == 1)
    delete_state(s);
  // remove transition t from E
  t.remove();
}
```

Fig. 17. Algorithm for removal of EFSM transition precluding observability.

Then, another method is proposed in Section 5.3.2 to reduce the complexity problem for the solver. The idea is that part of the work performed by the solver to find a sequence could be done earlier, and then invoke the solver on the reduced constraints set. In this case, no constraints are actually removed, but a part of the solution is provided to the solver. In such way, different constraints are already satisfied at beginning, and this is equivalent to reduce the number of constraints.
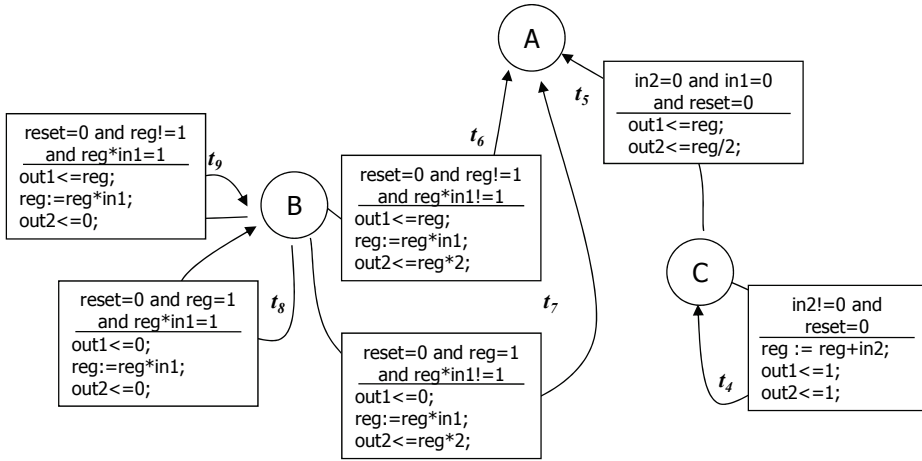
Fig. 18. Reduced EFSM.

5.3.1 Removal of useless transitions and of isolated states

Once a fault is observed on the registers, the ATPG saves a configuration. A configuration is composed by the state register and all the other registers of the design. Then, faulty and fault free configurations can be distinguished on the current state or on some registers value. Consider, for example, the EFSM represented in Figure 16 and that a particular fault *f* is observed on register *reg*. In faulty configuration *reg* is *0* and the EFSM in state *A*, and in the fault free configuration *reg* is 2 and state is always *A*. Then, if transition t_3 is traversed, a new value would be assigned to the register *reg*, losing the possibility of propagate the faulty configuration to the primary outputs. In fact, if the update function of transition t_3 is executed, then the faulty and fault free configuration would be equivalent.

Therefore, an algorithm has been defined to automatically prune this kind of transition from the EFSM model that is given in input to solver to generate a sequence. The algorithm is presented in Figure 17.

The algorithm uses the information collected during the learning phase to identify the transitions that cannot propagate a difference in the configuration and eliminates them.

Once the EFSM has been pruned from all the transitions that prevent the observability of registers configuration on primary outputs, some parts of the EFSM can remain isolated. Given the configuration state *s*, it is possible to be in a self-pointing state. This mean that there is no transitions outgoing from that state, but at most only transition in-going in *s*. Thus, the states of the EFSM that cannot be reached from the initial configuration are

```
// let E be a reduced EFSM
// let config_state be the configuration state
optimize (state config_state, efsm E) {
  state_set reached_states;
  // build the state set reached from the configuration state
  reached_states.insert(config_state);
  states_reached_from_state(config_state, reached_states);
  // remove all states that are not reached from the configuration
  for each state s in S {
    if ( not(reached_states.contains(s)) ) {
      s.remove();
      // remove all transitions moving out from state s
      transition tl = out_going_from_state(s);
      while ( not(tl.current_item() == NULL)) {
        tl.current_item().remove();
      }
    }
  }
}

// return the set of states reachable form sate s
states_reached_from_state(state s, state_set reached_states) {
  transition_list tl = out_going_from_state(s);
  while ( not(tl.current_item() == NULL)) {
    state in_state = in_going_state(t);
    // check whether current transition does not return to
    // current state and it does not reach a state which
    // has already been traversed in the visit
    if (not(in_state == s) and not(reached_states.contains(in_state))) {
      reached_states.insert(in_state);
      // continue exploring out-going paths
      // from state in_state
      states_reached_from_state(in_state, reached_states)
    }
    tl.move_to_next();
  }
}
```

Fig. 19. Algorithm for EFSM optimisation.

removed as they represent only constraints that cannot be satisfied. Figure 19 describes the algorithm that has been defined, to remove all the parts of the EFSM that are not reachable from the current configuration state. Let's consider the example in Figure 18 and say that configuration state is *B*. Then, starting form *B*, it is not possible to generate a sequence to reach state *C*. Therefore state *C* and all its out-going transition can be removed. The EFSM after the application of the optimization algorithm is depicted in Figure 20(a). Then, the Figure 20(b) presents the EFSM generated after optimization starting from state *C*.
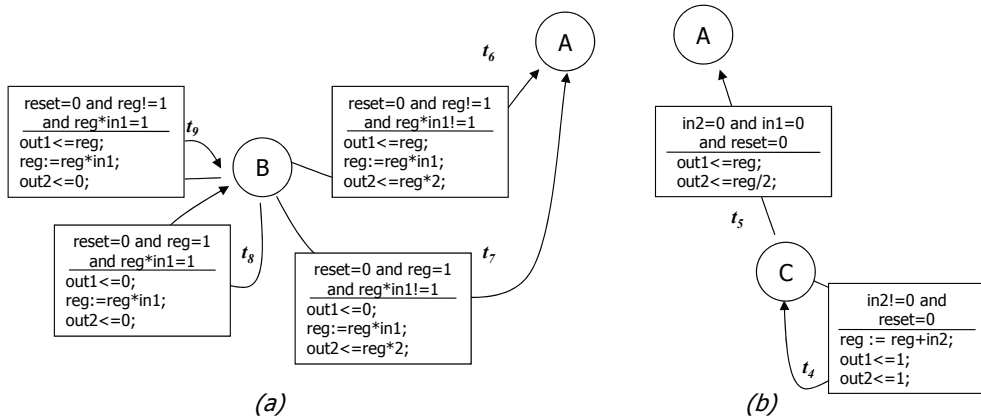
Fig. 20. Optimized from configuration's state B and C.

5.3.2 Pruning based on paths

The CLP solver tries to find a propagation sequence starting from the configuration that activates the fault. Such effort can be reduced if the solver is provided with a hint about the paths that are more profitable to be traversed. In this case, no constraints are actually removed from the CLP-based representation of the EFSM, but we provide the solver with part of the solution, thus reducing the number of satisfiable constraints.

Before invoking the solver to generate the propagation sequence, the ATPG applies an algorithm that searches for paths connecting the fault-free EFSM configuration with a state where at least one of update functions of its out-going transitions writes on the primary outputs. In fact, if a path allows updating the primary outputs, it is probable that the register value of the faulty configuration would be propagated on such outputs too. Then, all constraints related to the transition that are not involved in the generated paths are removed. These paths are generated such that they include transitions which allow the propagation of the values of registers involved in the faulty configuration towards the primary outputs. These transitions are identified and marked by learning phases performed during the EFSM generation and the subsequent removal of useless transitions. In particular, for every transition t marked as useful, a path is generated from the current configuration state to the out-going state of transition t.

The solver would try to check if the constraints are satisfiable and it would find a solution. If either it is not able to find a propagation sequence within a given timeout, or it returns that the problem is non satisfiable, a different path is generated and passed to the solver as initial configuration. Future works are related to associate a weight to each transition to generate paths maximizing the total weight of the involved transitions. Possible ideas for weighting transitions are: preferring transitions leading to a state, whose out-going transitions update the primary outputs, or transitions whose update functions use a large number of faulty registers, or transitions whose enabling functions consist of small conditions.

6. Experimental results

The CLP-based techniques for generating test sequences and propagation sequences has been applied to the benchmarks described in Table 1, where columns report the number of

primary inputs (PIs), primary outputs (POs), flip-flops (FFs) and gates (Gates). Column Trns. shows the number of transitions of the EFSM modelling the DUV and GT (sec.) the time required to automatically generate the EFSM. Then, column BC reports the number of bit-coverage faults injected into the designs to check the fault coverage.

DUV	PIs	POs	FFs	Gates	Trns.	GT (sec.)	BC
b00	66	64	99	1692	7	0.1	1182
b04	13	8	66	650	20	0.3	408
b10	13	6	17	264	35	0.3	216
b11m	9	6	31	715	20	0.2	725
b00z	66	64	99	11874	9	0.2	1439
fr	34	32	100	1475	10	0.2	1041

Table 1. Benchmarks properties.

Such benchmarks have been selected because they present different characteristics which allow analyzing and confirming the effectiveness of the proposed approach. *b04, b10* have been selected from the well known ITC-99 benchmarks suite (ITC, 1999). *b11m* is a modified version of *b11*, included in the same suite, created by introducing a delay on some paths to make it harder to be traversed. The HDL descriptions of *b04, b10* and *b11m* contain a high number of nested conditions on signals and registers of different size. *b00, b00z* and *fr* contain conditional statements where one branch has probability $1-(1/(2^{-32}))$ of being satisfied, while the other has probability $1/(2^{-32})$. Thus, they are very hard to be tested by a random ATPG. In particular, *b00* and *b00z* are internal benchmarks, while *fr* is a real industrial case, i.e., it is a module of a face recognition system.

6.1 Test sequence generation

The effectiveness of the CLP-based transition oriented ATPG has been evaluated by comparing to a genetic algorithm-based high-level ATPG (Fin & Fummi, 2003a), which outperforms pure random-based ATPGs but it is not aware about the EFSM structure, and with a pseudo-deterministic ATPG, which uses only the random-walk mode to traverse the DUV state space. Stopping criterion is defined in term of the number and length of the generated test sequences. Table 2 reports the transition coverage (TC%), the statement coverage (SC%), the fault coverage (FC%), and the test generation time (T (sec.)), by using respectively the genetic algorithm-based ATPG (GA-ATPG), the pseudo-deterministic ATPG (RW-ATPG), and the proposed ATPG (RW+BJ-ATPG). It can be observed that RW+BJ-ATPG outperforms both the GA-ATPG and the RW-ATPG. The very low transition coverage achieved by the GA-ATPG for some benchmarks is due to the presence of transitions out-going from the initial states, whose enabling functions have an infinitesimal probability of being traversed by randomly fixing the values of primary inputs. Such a problem is partially solved by the RW-ATPG which is aware about the enabling functions of the EFSM, and definitely solved by the backjumping-based RW+BJ-ATPG that reaches 100% transition and statement coverage for all benchmarks. Then, also the achieved fault coverage for all benchmarks is sensibly increased.

6.2 Propagation sequence generation

The efficiency of the CLP-based fault oriented ATPG for propagation sequence generation has been evaluated by applying the testing flow of Figure 1.

DUV	GA-ATPG				RW-ATPG				RW+BJ-ATPG			
	TC%	SC%	FC%	T (s.)	TC%	SC%	FC%	T (s.)	TC%	SC%	FC%	T (s.)
b00	28.6	26.7	1.1	3.0	85.7	87.0	48.7	2.6	100.0	100.0	52.5	2.9
b04	80.0	90.2	94.9	23.2	85.0	95.0	99.0	8.7	100.0	100.0	99.0	9.1
b10	37.1	66.7	87.0	5.7	40.0	69.7	93.0	5.7	100.0	100.0	94.0	6.8
b11m	90.0	80.0	37.0	5.7	95.0	82.2	39.0	5.1	100.0	100.0	54.6	16.3
b00z	22.2	31.0	13.7	4.1	66.6	75.9	44.3	5.0	100.0	100.0	51.8	12
fr	20.0	13.3	0.86	10.3	80.0	86.7	70.4	4.9	100.0	100.0	84.0	5.2

Table 2. Comparison between a GA-based ATPG, a pseudo-deterministic ATPG and proposed CLP-based approach.

DUV	RW+BJ-ATPG					CLP		CLP pure		RW+BJ-ATPG+CLP		
	FC%	TBP	TBA	SL	T (s.)	Prop	T (s.)	Prop	T (s.)	FC%	SL	T (s.)
b00	52.5	64	498	3	2.9	84	2.5	0	aborted	59.6	7	5.9
b04	99.0	4	1	6	9.1	4	3.6	0	aborted	99.8	10	13.4
b10	94.0	13	0	11	6.8	12	3.3	0	aborted	99.5	18	10.1
b11m	54.6	117	124	59	16.3	313	36.7	0	aborted	66.8	71	56.7
b00z	13.8	131	497	6	12	613	9.1	0	aborted	56.4	30	18.5
fr	84.0	63	82	42	5.2	22	6.0	0	aborted	86.1	80	11.2

Table 3. Experimental results of fault-oriented ATPG.

Columns RW+BJ-ATPG of Table 3 report results achieved by applying transition-oriented ATPG of the incremental test generation flow of Figure 1. In particular, these columns show the achieved fault coverage (FC%), the number of faults observed but not detected (TBP), the number of faults not activated (TBA), the average length of the generated test sequences (SL) and the test generation time (T (sec.)).

Then, the CLP-based fault-oriented engine has been applied to find propagation sequences for TBP faults. Columns Prop. and T (sec.), below CLP, reports, respectively, the number of TBP faults for which the CLP-based engine was able to generate a propagation sequence, and the corresponding execution time.

The column CLP pure reports the results achieved by using the CLP-based engine without applying the strategies described in Section 5.2.3 for managing the CLP complexity. In this case, all TBP faults were aborted, since the CLP solver always run out of resources. This highlights the effectiveness of the strategies proposed for managing the CLP complexity.

Finally, the last three columns of the table report, respectively, the total fault coverage (FC%), the average length of test sequences (SL) and the total generation time T (sec.) obtained by adopting all steps of the incremental testing flow shown in Figure 1.

Results show that the fault-oriented engine increased the fault coverage for all benchmarks without requiring long computation time. No fault has been aborted (i.e., the engine never run out of resources), even if some TBP faults remained untested, because no propagation sequence was found. The analysis of TBP faults not propagated highlighted the fact that many configurations allow TBP faults to be observed on internal registers (i.e., there exist many activation sequences), but very few of them allow TBP faults to be propagated. Moreover, such few configurations are difficult to be generated by using the transition-oriented ATPG, since they are not fault-oriented. To solve such a problem, in the future, the

CLP-based fault-oriented engine will be extended for the generation of activation sequences too.

The effectiveness of the optimization strategies, proposed in Section 5.3, are summarized in Table 4. This methodology has been applied also to another benchmark, *Prawn*, that is a RISC processor with the instruction set having been enhanced to include interrupt handling and conditional branches.

Columns St. and T. report, respectively, the number of states and transitions of the corresponding EFSMs. Column TBP shows the number of faults to be propagated which have been activated by using the RW+BJ-ATPG. Column TOut s. shows the timeout provided to the CLP solver for finding a propagation sequence. Columns PSEQ, Abort and Time s. under No optimization shows, respectively, the number of propagation sequences generated by the CLP solver, the number of faults aborted (i.e., the number of faults for which the solver was unable to provide a response, either positive or negative), and the total time required for generating the CLP constraints to model the EFSM and running the search. The same parameters have been computed after applying the optimization techniques presented in Section 5.3 (Optimized column). Experimental results show that the proposed optimization techniques sensibly improve the effectiveness of the solver in searching for propagation sequences. The improvement is particularly evident in the case of *Prawn*, whose EFSM is very large. Without optimizations the solver always aborted, while after optimizations were applied, it succeeded in generating propagation sequences for all the TBP faults. Moreover, it can be observed that the sequence generation time was sensibly decreased, for all benchmarks, but *b04*. In the case of *b04*, optimizations did not provide benefits, since its EFSM is composed of very few states that cannot be further reduced by applying the proposed optimizations.

DUV	St.	T.	TBP	TOut s.	No Optimization			Optimized		
					PSEQ	Abort	Time s.	PSEQ	Abort	Time s.
b04	3	20	4	10	4	0	16.23	4	0	18.51
b10	11	35	13	10	13	0	26.08	13	0	7.43
b11m	9	20	117	10	117	0	63.17	117	0	16.68
prawn	61	160	66	14	0	66	998.12	66	0	105.89

Table 4. Experimental results of fault-oriented ATPG with optimization.

7. Conclusions

This work defines a functional ATPG framework that exploits a particular kind of EFSM which has been theoretically showed to allow a more uniform traversing of the DUV state space. Determinism is obtained by interfacing with CLP solver that adopts formal methods to solve the conditions of the enabling functions.

The effectiveness of the proposed ATPG compared with a genetic-based ATPG is evident. It greatly benefits from the fact that, by using the EFSM model, all conditional statements included in the DUV are under its control. The adoption of the EFSM model joint to the learning/random-walk/backjumping-based mechanisms allows to accurately addressing hard-to-traverse transitions. Then the fault-oriented engine has been proposed together with. This is the first work addressing the problems of entirely modelling an EFSM by means of CLP, and generating functional test patterns by combining the use of EFSM and

CLP. Moreover, some strategies have been implemented to manage the CLP complexity, and experimental results showed that, in this way, the proposed engine is able to generate propagation sequences and increase the fault coverage without running out of resources.

8. References

Abramovici, M. (1993). Dos and don'ts in computing fault coverage. In *Proc. of IEEE ITC*.

Apt, K. R. & Wallace, M. G. (2007). *Constraint Logic Programming using Eclipse*. Cambridge Univeristy Press.

Cheng, K. & Krishnakumar, A. (1996). Automatic generation of functional vectors using the extended finite state machine model. *ACM Transactions on Design Automation of Electronic Systems*, 1(1):57–59.

Corno, F., Cumani, G., Reorda, M. S. & Squillero, G. (2001). Effective techniques for high-level atpg. In *Proc. of IEEE ATS*, pages 225–230.

Di Guglielmo, G., Fummi, F., Marconcini, C. & Pravadelli, G. (2006a). EFSM Manipulation to Increase High-Level ATPG Efficiency. In *Proc. of IEEE ISQED*, pages 57–62.

Di Guglielmo, G., Fummi, F., Marconcini, C. & Pravadelli, G. (2006b). Fate: a functional atpg to traverse unstabilized efsms. In *Proc. of IEEE ETS*.

Di Guglielmo, G., Fummi, F., Marconcini, C. & Pravadelli, G. (2006c). Improving Gate-Level ATPG by Traversing Concurrent EFSMs. In *Proc. of IEEE VTS*.

Di Guglielmo, G., Fummi, F., Marconcini, C. & Pravadelli, G. (2007). Improving high-level and gate-level testing with FATE: a functional ATPG traversing unstabilized EFSMs. *IEE Computers and Digital Techniques*, 1(3):187–196.

Dijkstra, E. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271.

Ferrandi, F., Fummi, F. & Sciuto, D. (1998a). Implicit test generation for behavioral vhdl models. In *Proc. of IEEE ITC*, pages 587–596.

Ferrandi, F., Fummi, F. & Sciuto, D. (1998b). Implicit test generation for behavioral vhdl models. In *Proceedings of IEEE International Test Conference (ITC)*, pages 436–441.

Ferrandi, F., Fummi, F. & Sciuto, D. (2002a). Test generation and testability alternatives exploration of critical algorithms for embedded applications. *IEEE Transactions on Computers*, C-51(2):200–215.

Ferrandi, F., Rendine, M. & Sciuto, D. (2002b). Functional verification for systemc descriptions using constraint solving. In *Proceedings of IEEE Design Automation and Test in Europe (DATE)*, pages 744–751.

Fin, A. & Fummi, F. (2003a). Genetic algorithms: the philosophers stone or an effective solution for high-level TPG? In *Proc. of IEEE HLDVT*, pages 163–168.

Fin, A. & Fummi, F. (2003b). Genetic Algorithms: the Philosopher's Stone or an Effective Solution for High-Level TPG? In *Proc. of IEEE HLDVT*, pages 163–168.

Fummi, F., Harris, I. G., Marconcini, C., and Pravadelli, G. (2007). A CLPbased Functional ATPG for Extended FSMs. In *Proc. of IEEE MTV*.

Gajski, D., Zhu, J. & Domer, R. (1997). Essential issue in codesign. Thecnical report ICS-97-26, University of California, Irvine.

Ghosh, I. & Fujita, M. (2001). Automatic test pattern generation for functional register-transfer level circuits using assignment decision diagrams. *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, 20(3):402–415.

Guarnieri, V., Fummi, F., Marconcini, C. & Pravadelli, G. (2008). An Optimized CLP-based Technique for Generating Propagation Sequences. In *Proc. of IEEE EWDTS*, pages 25–28.

Ier, M., Parthasarathy, G. & Cheng, K.-T. (2005). Efficient conflict-based learning in an RTL circuit constraint solver. In *Proc. of IEEE DATE*, pages 666–671.

ITC (1999). High time for high-level test generation. Panel at IEEE ITC.

Jaffar, J. & Maher, M. J. (1994). Constraint logic programming: A survey. *Journal of Logic Programming*, 19/20:503–581.

Kowalski, R. (1979). Algorithm = logic + control. In *Communications of the ACM*, pages 424–436.

Lee, D. & Yannakakis, M. (1992). Online minimization of transition systems. In *Proc. of ACM STOC*, pages 264–267.

Li, J. & Wong, W. (2002). Automatic test generation from communicating extended finite state machine (cefsm)-based models. In *Proc. of IEEE ISORC*, pages 181–185.

Lin, X., Pomeranz, I. & Reddy, S. M. (1999). Techniques for improving the efficiency of sequential circuit test generation. In *Proc. of ACM/IEEE ICCAD*, pages 147–151.

Lingappan, L., Ravi, S. & Jha, N. (2003). Test generation for non-separable RTL controller-datapath circuits using a satisfiability based approach. In *Proc. of IEEE ICCD*, pages 187–193.

Marconcini, C. (2008). A Functional ATPG as a bridge between Functional Verification and Testing. In *Ph.D. Thesis*.

Myers, G. (1979). *The Art of Software Testing*. Wiley - Interscience, New York.

Myers, G. (1999). *The Art of Software Testing*. Wiley - Interscience.

Padmanabhuni, S. (1999). Extended analysis of intelligent backtracking algorithms for the maximal constraint satisfaction problem. In *Proc. of IEEE CCECE*, pages 1710–1715.

Pauli, C., Nivet, M. L. & Santucci, J. F. (2000). Use of constraint solving in order to generate test vectors for behavioral validation. In *Proc. of IEEE HLDVT*, pages 15–20.

Russel, S. & Norvig, P. (2002). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Vemuri, R. & Kalyanaraman, R. (1995). Generation of design verification tests from behavioral vhdl programs using path enumeration and constraint programming. *IEEE Trans. Very Large Scale Integr. Syst.*, 3(2):201–214.

Wallace, M. & Veron, A. (1994). Two problems-two solutions: one system-ECLIPSE. In *IEE Colloquium on Advanced Software Technologies for Scheduling*, pages 1–3.

Wallace, M. G. (1997). Constraint programming. In *The Handbook of Applied Expert Systems*. CRC Press.

Wu, Q. & Hsiao, M. (2004). Efficient ATPG for design validation based on partitioned state exploration histories. In *Proc. of IEEE VTS*, pages 389–394.

Xin, F., Ciesielski, M. & Harris, I. (2005a). Design validation of behavioral vhdl descriptions for arbitrary fault models. In *Proc. of IEEE ETS*, pages 156–161.

Xin, F., Ciesielski, M. & Harris, I. (2005b). Design validation of behavioral VHDL descriptions for arbitrary fault models. In *Proc. of IEEE ETS*, pages 156–161.

Xin, F. & Harris, I. G. (2002). Test generation for hardware-software covalidation using non-linear programming. In *Proc. of IEEE HLDVT*, pages 175–180.

Zhang, L., Ghosh, I. & Hsiao, M. (2003). Efficient Sequential ATPG for Functional RTL Circuits. In *Proc. of IEEE ITC*, pages 290–298.

New Concepts of Asynchronous Circuits Worst-case Delay and Yield Estimation

Miljana Milić and Vančo Litovski
University of Niš, Faculty of Electronic Engineering
Serbia

1. Introduction

Although the benefits of asynchronous design style are undeniable, this style is still a road that designers rather avoid. There are, however, serious advantages of this digital design concept that are making it favourable for many applications. Asynchronous circuits need no clock generation and distribution (Sparso, 2006; Martin & Nystrom 2006), which leaves the problems related to clock skew behind and saves a lot of chip area. Asynchronous circuits are characterized with much easier technology migration and good modularity. Very low EMI occurs during operation, while achieving high noise immunity (Lewis & Brackenbury 2001). Power is consumed only when useful work is done. The absence of the clock itself reduces the power consumption. These issues are very important while designing portable systems where battery size and lifetime are important.

Synchronous circuit design styles have enormous commercial practice and very significant pedigree, and those are the major reasons for the lack of motivation to apply asynchronous circuit techniques (Davis & Nowick, 1997). Nevertheless, the motivation to pursue the study of asynchronous circuits is based on the simple fact that all high-performance ''synchronous'' design styles are ''asynchronous in the small" (Cortadella et al. 1999) Because of that, some techniques for desynchronization of synchronous circuits have appeared lately (Cortadella et al. 2006; Andrikos 2007).

Beside their benefits, some problems related to asynchronous circuit design are still waiting to be solved. One of the most important is the estimation of asynchronous circuit performances. That is, determining the delays of the paths in a particular asynchronous circuit. Early evaluation of the path delays in the circuit helps avoiding early timing problems as well as circuit performance characterization (Sokolovic, Litovski & Zwolinski 2009). Precise paths delays, of course, can be estimated only in the final steps of the design process. That is because the delay is extracted from the circuit after layout synthesis. If the delays do not satisfy the required speed of the circuit, the circuit has to be redesigned. The same conclusion stands when timing problems occur. This strongly suggests that new methods are to be offered enabling delay estimation to be performed during the early phases of digital system design. Our aim here is to establish the application of a standard logic simulator in asynchronous circuit path delay analysis and parametric yield estimation.

The simplest way to determine the circuit delay is simulation. At the transistor level complex circuits' simulation becomes inefficient. To verify the logic function and the timing

specifications of the circuit, logic simulators dealing with gate level descriptions are used. However, the delay of the circuit obtained by a logic simulator depends on the input test vectors. In order to determine the longest and the shortest possible delays in a combinational circuit, it has to be simulated using all 2^n possible input vectors, where n stands for the number of inputs. Therefore, the simulation is not an efficient solution for most circuits. On the other hand, since logic simulation is one of the first design steps, embedding worst-case delay analysis into a logic simulator would ensure early detection of incorrect design solutions.

The causes of the parameter values' variations are following: temperature and environment; technology and process; and specific phenomena within components (such as electromigration). These variations affect the circuit behaviour over time. Because of them, a 100% parametric yield is not achievable since the responses of all the manufactured circuits do not satisfy the required timings. The nature of parameter variations is statistical in the sense that all parameter values are random within a probability interval. As a result, the response values in particular delays are also randomly distributed within an interval that depends on the nature of the mapping of parameter tolerances onto response tolerances (Litovski & Zwolinski 1997).

Timing analysis of a circuit consisting of primitive gates is performed by a timing analysis tools. These tools can calculate circuit delays that are the result of parametric variations. The aim of static timing analysis is to verify that a logic circuit satisfies its timing constraints, i.e. that the logic circuit will function correctly when run at an intended speed (Spence & Soin 1988).

Commercial timing analysis systems are based on statistical timing verification. A problem that often appears in statistical timing analysis is that the longest paths often become false paths. Moreover, when considering DSM (deep-submicron) technologies, almost every path can be considered critical. The critical path of a digital circuit is the longest path or the path with the largest delay between the input and the output of the circuit (Mak et al. 2004).

Direct methods and sampling methods can be used instead of worst-case and statistical timing analysis. Direct methods are using formulae that map the parameter tolerances into the response tolerances. These methods are applied to cases where the parameter variations are small. Nevertheless, this analysis is, although more accurate, still very time consuming, since it requires all gate sensitivities with respect to all possible variations to be calculated.

2. How to analyze timings

The purpose of the timing analysis is to determine the following timing constraints:
- Do the signals arrive at ports in time?
- Do the signals stay long enough at the required state to be useful?
- Will the signals propagate with a proper slope?
- Can the hardware run with a specified speed?
- Are there any paths which need additional analysis and modification?

Timing measurements, as already mentioned, can be performed using a circuit simulation, but such an approach is too slow to be practical. There are two simulation alternatives for delay estimation in logic circuits. The first approach is based on static timing analysis (STA) and statistical static timing analysis (SSTA), while the second includes Monte-Carlo analysis. STA methods evaluate digital circuit timing without simulation. For nanometre manufacturing processes, which have increased parameter variability, a corner-based STA

has become inadequate. To avoid this problem, a statistical approach has been proposed: statistical static timing analysis (SSTA). As a result of SSTA analysis probability density functions (pdfs) are obtained. The percentage of fabricated dies which meet a required delay, can then be calculated or conversely, the expected performance for a particular parametric yield (Maksimović 2000). SSTA aims to determine the distribution of the delay of a design by accounting for actual statistics of process variations.

In practice, however, complete and exact statistical information is not always available or might be difficult to obtain from the foundries. In (Sokolović & Litovski 2005) a method for worst-case estimation of timing-yield that deals with those difficulties is presented. This method is based on distributional robustness theory (DRT). Nevertheless, since finding the worst-case estimate is defined as an optimization problem, it requires comparable simulation and data processing time. The probabilistic nature of the timing behaviour of a circuit, for selection of the critical path, imposes implementation of statistical analysis and simulation. However, even with their clear advantages, developing and using statistical models and methods requires considerable effort. The complexity of the statistical techniques is still significantly high. These can be reasons for avoiding statistical methods, but higher process integration and increasing operational speed also make them inevitable (Mak et al. 2004).

The Monte-Carlo method is based on a large number of circuit simulations (analyses). As a result, it gives the mean and standard deviation of the delay at the output of the circuit. A Monte-Carlo simulation cycle has two steps: a sampling step and an analysis step. In the sampling step, for a given set of parameters (delays of gates in this case) a single random value for every parameter is produced according to the given probability distribution.

An analysis of the circuit must be performed for each new version of the parameter vector. In this way a set of different parameter values (properties) of the circuit's output signals are obtained. The analysis step, in addition, utilizes these sampled values to derive the arrival times of all output signals for the given circuit instance.

The desired accuracy determines convergence criteria of the process i.e. the number of cycles. In fact, there should be no convergence, since the simulation results do not approach to any particular value. This is why the statistical properties of the responses are monitored. Once the mean or variance converges to the desired precision range, the procedure terminates. It takes from a several tens to a few hundreds of Monte-Carlo cycles to achieve convergence of the results. This means that the timing analysis step should be repeated that many times (Lin & Davoodi 2008). If, however, each iteration of Monte-Carlo analysis involves a transistor level simulation of the entire circuit (or the entire circuit path), this approach will have an unacceptable run time (iscas89.html --).

The design and analysis can significantly be accelerated by application of a logic simulator for the timing analysis in the Monte-Carlo analysis. A way for timing analysis with a VHDL logic simulator based on the research in (Maksimović 2000; Maksimović & Litovski 1999; Maksimović & Litovski 2002) will be presented next. It simplifies the delay evaluation procedures and speeds them up. In this way a good base for evaluation of asynchronous circuits' performances is established. This method will now be explained in more detail.

3. Estimating delays with a logic simulator

Our method for path-delay estimation in asynchronous nonsequential digital circuits is based on a robust delay estimation algorithm. It makes sense to analyze the circuit paths

only in one operating sequence. Because of that, and to be able to implement the suggested method to sequential circuits, one needs to pay special attention when dealing with circuits that have feedback loops. When sequential circuits are analysed at the circuit level, the feedback loops, need to be broken, while, for analysis at the gate level, complex models of element descriptions are required. The proposed concept can enable acceleration of Monte-Carlo analysis if it is embedded within the analysis step of the Monte-Carlo loop. The sampling step of the Monte-Carlo analysis is performed in the usual manner.

To perform a timing analysis that is, a delay estimation of all the paths in a circuit using a logic simulator, the logic simulation mechanism needs a small modification (Maksimović 2000). Neutral events that do not change the logic value of the signal in a standard logic simulator are ignored. If the signal description is extended to have a few additional attributes, such as event, delay value, etc. (Mak et al. 2004; Maksimović & Litovski 1999), then a change in any of those attributes will be considered as a non-neutral event. Simultaneous propagation of all input vectors through the circuit is assumed. The values of delay attributes are accumulated along structural paths, starting from the primary inputs and ending at the primary outputs or, if necessary, at any particular node inside a circuit. At the end of this very fast delay estimating process, after only one run of the logic simulator, all delays of both output signal edges are available.

3.1 Models of gates and signals

For each output signal of the circuit, S, four delay values are estimated:

$d1mn(S)$ – the shortest path delay for a rising edge at S,

$d0mn(S)$ – the shortest path delay for a falling edge at S,

$d1mx(S)$ – the longest path delay for a rising edge at S,

$d0mx(S)$ – the longest path delay for a falling edge at S.

In order to evaluate all worst-case path delays to all the signals in the circuit with just one simulation, it is necessary to perform simultaneous simulation of the circuit for all input vectors. To enable this, signals that connect logic gates within the circuit must contain two types of information: events on the signal, and the shortest and the longest path delays to the signal. This information is stored within a signal description in the form of two types of attributes: attributes that contain the delay information, as listed above, and the attributes for triggering the delay calculation processes in a gate. For a signal, S, the four attributes for triggering the calculation are:

$arr1mn(S)$ – rising transition of shortest path arrival flag,

$arr0mn(S)$ – falling transition of shortest path arrival flag,

$arr1mx(S)$ – rising transition of longest path arrival flag,

$arr0mx(S)$ – falling transition of longest path arrival flag.

It should be noticed that the signal now does not contain any logic values, as would be the case in a standard logic simulation.

To process signals described in this way and to perform the delay estimation, the gate model must include two modes: the activation – propagation mode and the delay calculation mode. Moreover, the gate description must contain two separate processes; first to calculate the maximal delay of the falling and rising transitions, and the second to calculate the

minimal delay of the falling and rising transitions. The activation – propagation mode of the model in each of these processes in a gate is sensitive to every change of the signal triggering attribute. After the delay calculation level of the model is activated, it then updates the output signal delay according to the input signal delay attributes and gate delay parameters. When the resulting output delay type (delay attribute of the output signal) is calculated, the output signal changes the particular triggering attribute to trigger processes in the following gates.

```
generic (ifo_izl: integer:= 1;
tpd_hlmn : real := 0.9e-9;
tpd_lhmn : real := 1.0e-9;
tpd_hlmx : real := 0.95e-9;
tpd_lhmx : real := 1.05e-9);
     .            .              .
p1: process (in1.d0mn, in1.d1mn, in1.arr0mn, in1.arr1mn,
        in2.d0mn, in2.d1mn, in2.arr0mn, in2.arr1mn,
        in3.d0mn, in3.d1mn, in3.arr0mn, in3.arr1mn)
        variable r,p: real;
        variable multipl : real;
begin
        multipl := real(ifo_izl);
        r:= ((multipl*1.0) + (0.03*(gauss_rng)));
        p:= (multipl*0.9 + (0.03*(gauss_rng)));
        if (in1.arr0mn or in2.arr0mn or in3.arr0mn) then
                out1.d1mn  <= min(min(in1.d0mn, in2.d0mn), in3.d0mn) + r;
                out1.arr1mn <= true;
        end if;
        if (in1.arr1mn and in2.arr1mn and in3.arr1mn) then
                out1.d0mn   <= min(min(in1.d1mn, in2.d1mn), in3.d1mn) + p;
                out1.arr0mn<= true;
        end if;
end process p1;

p2: process (in1.d0mx, in1.d1mx, in1.arr0mx, in1.arr1mx,
        in2.d0mx, in2.d1mx, in2.arr0mx, in2.arr1mx,
        in3.d0mx, in3.d1mx, in3.arr0mx, in3.arr1mx)
        variable r,p: real;
        variable multipl : real;
begin
        multipl := real(ifo_izl);
        r:= (multipl*1.05 + (0.03*(gauss_rng)));
        p:= ((multipl*0.95) + (0.03*(gauss_rng)));
        if (in1.arr0mx or in2.arr0mx or in3.arr0mx) then
                out1.d1mx   <= max(max(in1.d0mx, in2.d0mx), in3.d0mx) + r;
                out1.arr1mx <= true;
        end if;
        if (in1.arr1mx and in2.arr1mx and in3.arr1mx) then
                out1.d0mx   <= max(max(in1.d1mx, in2.d1mx), in3.d1mx) + p;
                out1.arr0mx<= true;
        end if;
end process p2;
```

Fig. 1. Process for assigning minimal and maximal delay of the falling and rising edges for a tree input NAND gate.

An example of the process for assigning the minimal and maximal delay of the falling and rising edges for a tree input NAND gate is shown in Fig. 1. The gate inputs are denoted as

in1, in2 and in3, and the output as out1. The gate propagation delays for the rising and falling edges at the output out1 in both processes are denoted by tpd_lhmn/mx (minimal or maximal time propagation delay from low to high) and tpd_hlmn/mx, respectively (minimal or maximal time propagation delay from high to low). Each falling transition at an input of the gate means that one of the falling transition flags at one of the input signals (in1.arr0mn/mx or in2.-arr0mn/mx or in3.arr0mn/mx) becomes ''true''. This sets a rising transition flag at the output signal attribute of the gate (out1.-arr1mn/mx) to ''true''. This corresponds to an OR function. Simultaneously with setting the output flag, the gate model calculates a new value for the shortest and the longest path delays. The resulting output shortest and longest path-delay attributes for the rising edge is denoted by out1.d1mn/mx and is calculated after taking into account the arriving shortest/longest path delays for all tree gate input signals (in1.d0mn/mx, in2.d0mn/mx, in3.d0mn/mx), the minimal/maximal delay of the rising edge for this gate (a separate function assigns this value) and the function f which depends on the gate fanout value. Conversely, a rising transition flag at any of the gate input signals (in1.arr1mn/mx, in2.arr1mn/mx, in2.arr1mn/mx) produces a falling transition at the output only if a rising transition has previously arrived at all other gate inputs (Spence & Soin 1988; Agarwal et al. 2003). This corresponds to an AND function. The resulting output shortest/longest path-delay attributes for the falling edge, denoted by out1.d0mn/mx, take into account arrived shortest of longest path-delay input signal attributes (in1.d1mn/mx, in2.d1mn/mx, in3.d1mn/mx), the minimal/maximal delay of the falling edge for this gate and the fanout dependent function f. A process denoted with p1 calculates minimal path delays for rising and falling edges while the process denoted with p2 calculates maximal path delays for rising and falling edges. In this figure, it is also shown that the delays of a particular gate can be generated by different random functions which can take into account different input signal slopes, loading capacitances and other parameters that influence the ranges of rising and falling gate delays, tpd_lhmx and tpd_hlmx.
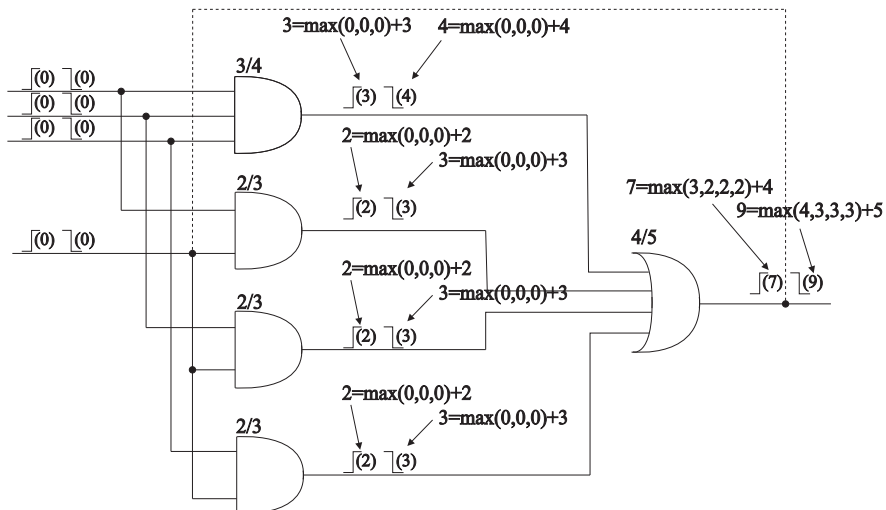
Fig. 2. Illustration of maximal delay estimation method for a possible implementation of a three input C-element.

The basic principle of delay accumulation is described in Fig. 2. The figure describes the maximal delay calculation of all paths in the simple asynchronous circuit which represents one possible implementation of the tree input C-element. Here, both rising and falling transitions are applied to all inputs of the circuit. Both, rising and falling transition delays are updated by each gate. The delay estimation of the circuit stops when all transitions reach the primary outputs. To analyze delays of the paths the feedback line had to broken.

3.2 Dealing with isolated asynchronous sequential elements

It may be of interest to estimate maximal delays of the rising and falling edges for the path that is going through the feedback in this circuit. The method that we suggest can be extended to be able to deal with the elements with loops or feedbacks. In these circuits longest paths will probably be ones that go through the feedback branches and return to the input of the circuit. The approach is similar to one used to generate test vectors in sequential digital circuits described in Ref. (Cheng & Agrawal 1989). Since more sequences must be observed, the circuit is replicated more times, representing each time sequence, and the delay along the paths is estimated for as many sequences as necessary. Fig. 3 shows how the principle would be used to estimate the delay of the longest path for the circuit of a simple asymmetric C-element. Assume that elements' delays in this figure are given in nanoseconds. In the implementation sense, the circuit will not be replicated many times. Instead of complicating the circuitry, only the results of one estimating simulation will be applied to the circuit input signals that are connected to the feedback loop. After that the circuit can be simulated again, with new initial delay parameters. Table 1 shows the results of repeated estimation for tree sequences.
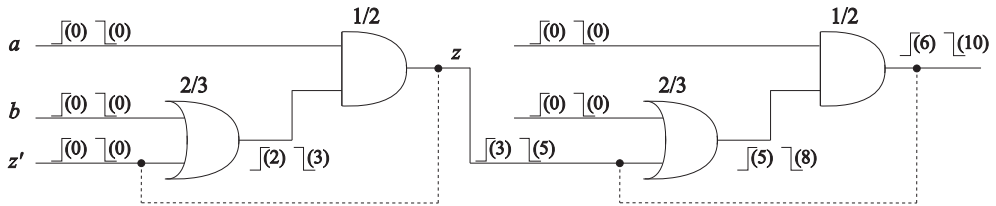
Fig. 3. Illustration of maximal delay estimation method in an isolated sequential element.

time sequence	$t_{mxr}(z)$ [ns]	$t_{mxf}(z)$ [ns]
I	3	5
II	6	10
III	9	15

Table 1. Results of the delay analysis for the asymmetric C gate in tree sequences.

3.3 Assigning the delay to a gate

In order to calculate four different worst-case delays for all paths in one digital circuit, the gate models must contain all four types of delays. It means that each gate in a circuit is characterized with four parameters: the maximal delay of the rising edge, the maximal delay of the falling edge, the minimal delay of the rising edge and the minimal delay of the falling edge. Nevertheless, assigning a particular delay to the gate and calculating the result is a

complex task. There are two delay components in each of gate delay functions (they are denoted with r and p in Fig. 1). One takes into account the fact that we want to use the gate models for statistical worst-case delay estimation, and the second must take care of the netlist of the entire digital circuit, that is the fanout information for each gate in the circuit. This is expressed by Eq. (1)

random_value_of(tpd_lhmx/tpd_hlmx/tpd_lhmn/tpd_hlmx)*fanout_func(output)=
func(tpd_lhmx/tpd_hlmx/tpd_lhmn/tpd_hlmx)   (1)

where the slash sign (/) represents the OR function.

Considering the first component, we must introduce randomness into the delay assignment process. This is the crucial step which enables Monte-Carlo analysis. The need for statistical delay analysis comes from the variations of the circuit parameters. Therefore, the delay estimation method must also include the influence of parameter variance on minimal and maximal delays of all signals in the circuit. If the delays were modelled as fixed values, the worst-case delay values would not be the real worst-case values due to the process variations. If we consider this fact, we conclude that the solution to this problem can be delay estimation in the usual manner, while all delay ranges in each gate instance are generated randomly. Each time the calculation process is activated in a gate, new worst-case delay values are considered in the signal delay calculations. Since simple models are used and the calculations are still very time-efficient, the simulations can be performed a few hundred times to enable statistical delay analysis. In this way a method for statistical static timing analysis using a standard logic simulator has been developed (SSTA for SLog).

The given delay probability density function determines the delay randomness. Hence, the gate parameters are the mean values of the probability density function for a particular gate type. The gate delay information given as parameters incorporates the real fabrication variations of a particular technology since worst-case delay distributions are characterized with mean and variance values that should be given as the fabrication technology parameters. Each gate randomly generates the maximal and minimal delays of the rising and falling edges with a Gaussian distribution, with the mean and standard deviation defined according to Eq. (2)

$$\varphi(p) = \frac{\exp\left[-\left(p - \mu_p\right)^2 / \left(2\sigma_p^2\right)\right]}{\sigma_p \sqrt{2\pi}}$$ (2)

where μ_p represents the mean value and σ_p^2 is the variance of the random variable p (Litovski & Zwolinski 1997). This function can, of course, be changed if necessary.

The second component of Eq. (1) deals with the real position of the particular gate in the netlist of the entire circuit. It is well known that the delay of the output signal for a single gate depends on the number of gates that are driven by that particular gate. If a gate has to drive two gates, the delay is larger than in the case of driving a single gate. In order to increase the accuracy of the gate delay model and the entire delay estimation algorithm, the fanout information of each gate in the circuit netlist must be included in the delay calculations. To do this, two major modifications must be introduced. One modification affects the logic gate descriptions. The second must be performed on the digital circuit netlist. In this way, the real implementation of the circuit is taken into account. For example,

if one gate output drives two inputs of following gates, it means that all delays of the particular gate will be increased according to the approximation function. Also, the technology has a large impact on the fanout_func(output) value, since the function that gives the fanout dependence of the delay is specific to each technology and each gate type, and would be given by the manufacturer. The VHDL implementation of this idea will be shown later.

3.4 The algorithm for path-delay estimation

For statistical estimation of worst-case delays, that is SSTA using a standard logic simulator – it is necessary to perform a few hundred estimation simulations. The exact number of simulations, thanks to the nature of the Monte-Carlo process, is not influenced by the number of parameters but by the required precision and accuracy of the results.

Table 2 gives a description of estimation phases for one sample.

Input:	-Ranges of delays for rising and falling transition for each gate
	-circuit netlist
	-library of circuit elements described to support the timing analysis
Output:	-Ranges of delays for rising and falling transition for all circuit output signal
step 1:	-Set all signals in the circuit to be a composite type consisting of the following attributes:
	4 different delay information (maximal and minimal delays of rising edge and maximal and minimal delays of falling edge)
	4 different flags for triggering each delay type calculation in a particular gate
step 2:	-Initialize all signals triggering flags to "false" value. Setting them to value "true" starts the calculations.
step 3:	-Initialize all signals to have zero values of the delay attributes
initializa tion	-Initialize the calculation process by setting the primary input signal triggering flags to "true"
step 4:	-Until all signals and gates are processed (all signal triggering flags should be set to "true") perform the following steps by moving through the topological levels of the circuit:
	-The delay calculation is activated when all input signals of the gate have triggering flags set to "true". When the delay calculation depends on the logic function of the gate, it is taken into account. For example, each falling transition (flag for falling transition is set to "true" at the input of the gate) at an AND gate input, produces a falling transition at its output (sets the gate output signal triggering flag for falling transition to "true"), but a rising transition at an input is able to produce a rising transition at the output only if the rising transition had previously arrived at all other gate inputs.
	-In order to complete all attributes of the gate output signal, before activating its output transition flag, the corresponding gate delay should be calculated by processing delays that arrived with input signals of the gate. The particular delay of the chosen gate is also added to the resulting corresponding delay.
	-The estimation terminates when all triggering flags for all output signals are set to "true".

Table 2. Path-delay estimation algorithm with a logic simulator.

The delay values of a particular gate have standard deviation r, which is in our case set to be 3%. This can be varied if necessary. This value is derived from the parameter tolerances for an integrated circuit fabrication technology.

The circuit is described and simulated at the structural (gate) level, while having available delay ranges values (minimum and maximum delays) of all building blocks for both rising and falling edges. When this estimation process is embedded in a Monte-Carlo loop, the delays for a gate in a circuit will be characterized with a mean and a variance and then randomly chosen in each estimation process. At the start of the simulation, the circuit is excited with both rising and falling transitions at all primary inputs. This is referred to as the initialization phase, where all triggering attributes of all signals at the primary circuit inputs are set to ''true'', that is the transitions at all primary inputs are initialized. All these transitions initiate the estimation processes in the gates at the first topological level of the digital circuit. When these processes are completed, the processes in the first topological level gates activate the transitions at their outputs to enable the calculation processes of the gates in the second topological level. As the transitions propagate from primary inputs towards primary outputs, the gate delays are accumulated along the paths, since an activation transition for the gate output signal is possible only if the delay of that gate has been already estimated. Signal attributes for the delay calculation and the calculation activation are used by the processes in the gate models and their values are dynamically updated, while the wave of activation shifts from the input to the outputs of the gates and the entire circuit. Once the circuit calculation activity is exhausted, the shortest and the longest path delays are available in signal attributes d1mn, d1mx, d0mn and d0mx of each output signal in the circuit.

It should be mentioned that these simulations, in addition, do not require any kind of stimuli selection, since they take into account all possible signal transitions. Only initialization is needed for the calculation processes in the entire circuit.

4. VHDL implementation

As already mentioned, the proposed concept is implemented using the VHDL hardware description language and simulator. Matlab was used for processing data obtained after simulations.

In order to have statistical simulation results, a random number generator is needed. Fig. 4 shows a VHDL implementation of the random number generator with a Gaussian distribution (Zwolinski 2004).

```
function gauss_rng return real is
        variable u1, u2, v1, v2, r, q, p: real;
        begin
            loop            u1:=rand;
                            u2:=rand;
                            v1:=u1*2.0 -1.0;
                            v2:=u2*2.0 -1.0;
                            r:=v1*v1 + v2*v2;
                            exit when r<1.0;
            end loop;
            q:=log2(r);
            p:=(sqrt((0.0-2.0)*q/r))*v1;
        return p;
end function gauss_rng;
```

Fig. 4. Gaussian random function generation implementation.

This function is executed 4 times within each gate, once for each delay type. Function rand in this description generates random numbers in the interval [0,1], with a uniform distribution.

In order to verify the efficiency of the applied gate models, we created a simple test circuit that calls the delay random generation function 600 times. The result of this can be used for random delay generation of rising or falling edges at the output of this circuit. The resulting histogram is shown in Fig. 5. In this case, the mean in the distribution is set to 1 ns, while the standard deviation is 3%. The x-axis shows the delays in [ns] units, and the y-axis represents the number of particular delay appearances within the corresponding range.
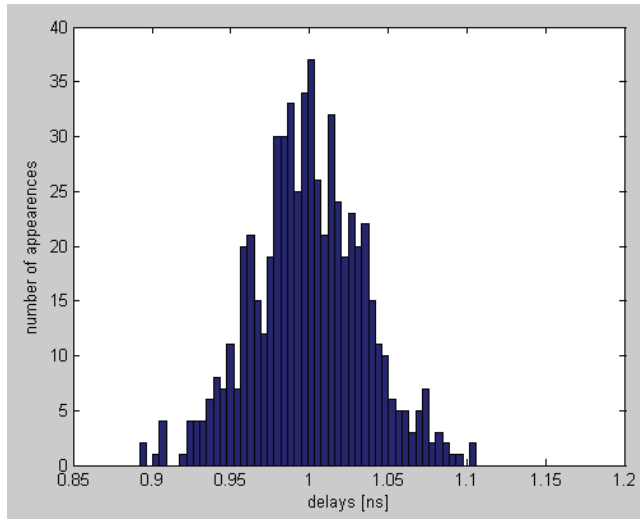
Fig. 5. Histogram of the function for random delays generation.

The shape of the particular delay type distribution function can be of arbitrary complexity. It is even possible to implement the spatial correlations within those random processes. Nevertheless, since we speak about a pre layout analysis no data are available for doing that at this stage.

VHDL models of primitive logic gates and simple asynchronous elements are kept in a VHDL library. Figs. 6–8 show VHDL modelling of a D-latch, RS-latch and two input C-element, respectively. It should be noted that D-latch circuit does not have specific conditions for activating a calculation process, because these circuits have only one data input.

Considering RS-latch circuit, the situation is far more complicated. It is assumed that this circuit has two inputs, R and S, and two outputs Q and its complement NQ. At the beginning of the description, it can be noticed that this circuit requires 18 different generics. There are 16 delay generics that are related to minimal and maximal delays through the latch, of four possible input–output combinations (R–Q, R–NQ, S–Q, and S–NQ), and to both possible signal transitions. The first and the second generics consider fanout values for the first and the second output. As at the other gate descriptions, these two values are initially set to one. In the latch instantiation, and after the circuit netlist is being processed, these values become their actual values that correspond to their actual topological position

```
entity DLatch is
        generic (ifo_izl: integer:= 1;
                tr_en_qmn : real := 1.0e-9;
                tf_en_qmn : real := 0.9e-9;
                tsu_d_enmn : real := 0.45e-9;
                tr_en_qmx : real := 1.05e-9;
                tf_en_qmx : real := 0.95e-9;
                tsu_d_enmx : real := 0.55e-9);
        port (q : out SDA_std_logic := (0.0, 0.0, false, false, 0.0, 0.0, false,
        false);
        d, en: in SDA_std_logic := (0.0, 0.0, false, false, 0.0, 0.0, false,
        false));
        end DLatch;
architecture only of DLatch is
begin
        p1: process  (en.d0mn, en.d1mn, en.arr0mn, en.arr1mn,
        en.d0mx, en.d1mx, en.arr0mx, en.arr1mx)
                variable i, j, k ,l, m, n: real;
                variable multipl : real;
        begin
            multipl := real(ifo_izl);
            f<=fanout_func(multipl)
            i:= (f*1.0 + (0.03*(gauss_rng)));
            j:= (f*0.9 + (0.03*(gauss_rng)));
            k:= (f*0.45 + (0.03*(gauss_rng)));
            l:= (f*1.05 + (0.03*(gauss_rng)));
            m:= (f*0.5 + (0.03*(gauss_rng)));
            n:= (f*0.55 + (0.03*(gauss_rng)));

            q.arr1mn <= true;
            q.arr0mn <= true;
            q.d1mn <= en.d1mn + i + k;
            q.d0mn <= en.d1mn + j + k;
            q.arr1mx <= true;
            q.arr0mx <= true;
            q.d1mx <= en.d1mx + l + n;
            q.d0mx <= en.d1mx + m + n;
        end process;
end only;
```

Fig. 6. VHDL model of the D-latch.

and implementation in the entire circuit. There are two processes in the circuit description, for determining all maximal (process p1, shown in Fig. 7a) and all minimal delay types (process p2, described in Fig. 7b). For example, the rising transition at the output Q happens if both falling transition at input R and rising transition at input S have arrived. The maximal delay of the rising transition at output Q is then calculated as a maximum between the maximal delays of the falling transition at R and rising transition S, which arrived at those inputs, plus the maximal of new, randomly generated maximal delays of the rising transition delays between S–Q and R–Q ports. The same conditions must be established for achieving the falling transition at the output NQ, while the maximal delays of the falling edge at this output are calculated using other values. The falling transition at the output Q happens if both rising transition at input R and falling transition at input S have arrived. The maximal delay of the falling transition at output Q is then calculated as a maximum between the maximal delays of the rising transition at R and falling transition S, which arrived at those inputs, plus the maximal of new, randomly generated maximal delays of the falling transition delays between S–Q and R–Q ports. Similar process stands for determining the minimal delay types.

```
entity RSLatch is
    generic (ifo_izl_1: integer:= 1;
        ifo_izl_2: integer:= 1;
        tr_rq_mn : real := 1.0e-9;
        tf_rq_mn : real := 0.9e-9;
        tr_rq_mx : real := 1.05e-9;
        tf_rq_mx : real := 0.95e-9;
        tr_rnq_mn : real := 1.0e-9;
        tf_rnq_mn : real := 0.9e-9;
        tr_rnq_mx : real := 1.05e-9;
        tf_rnq_mx : real := 0.95e-9;
        tr_sq_mn : real := 1.0e-9;
        tf_sq_mn : real := 0.9e-9;
        tr_sq_mx : real := 1.05e-9;
        tf_sq_mx : real := 0.95e-9;
        tr_snq_mn : real := 1.0e-9;
        tf_snq_mn : real := 0.9e-9;
        tr_snq_mx : real := 1.05e-9;
        tf_snq_mx : real := 0.95e-9);
    port (q, nq : out SDA_std_logic := (0.0, 0.0, false, false, 0.0, 0.0, false, false);
        r, s : in SDA_std_logic := (0.0, 0.0, false, false, 0.0, 0.0, false, false));
end RSLatch;
architecture only of RSLatch is
begin
    p1: process (r.d0mx, r.d1mx, r.arr0mx, r.arr1mx, s.d0mx, s.d1mx, s.arr0mx, s.arr1mx)
        variable i, j ,k, l, m, n, o, p : real;
        variable multipl1, mulitipl2 : real;
    begin
        multipl1 := real(ifo_izl1);
        multipl2 := real(ifo_izl2);
        f1<=fanout_func(multipl1);
        f2<=fanout_func(multipl2);

        i:= (f1* tr_rq_mx + (0.03*(gauss_rng)));
        j:= (f1* tr_sq_mx + (0.03*(gauss_rng)));
        k:= (f2* tf_rnq_mx + (0.03*(gauss_rng)));
        l:= (f2* tf_snq_mx + (0.03*(gauss_rng)));
        if (r.arr0mx and s.arr1mx) then
            q.d1mx<=max(r.d0mx,s.d1mx)+max(i, j);
            q.arr1mx <= true;
            nq.d0mx<=max(r.d0mx,s.d1mx)+max(k, l);
            nq.arr0mx <= true;
        end if;

        m:= (f1* tf_rq_mx + (0.03*(gauss_rng)));
        n:= (f1* tf_sq_mx + (0.03*(gauss_rng)));
        o:= (f2* tr_rnq_mx + (0.03*(gauss_rng)));
        p:= (f2* tr_snq_mx + (0.03*(gauss_rng)));
        if (r.arr1mx and s.arr0mx) then
            q.d0mx<=max(r.d1mx,s.d0mx)+max(m, n);
            q.arr0mx <= true;
            nq.d1mx<=max(r.d1mx,s.d0mx)+max(o, p);
            nq.arr1mx <= true;
        end if;
    end process;
```

Fig. 7. VHDL model of the RS-Latch.

```
p2: process (r.d0mn, r.d1mn, r.arr0mn, r.arr1mn, s.d0mn, s.d1mn, s.arr0mn, s.arr1mn)
    variable i, j ,k, l, m, n, o, p : real;
    variable multipl1, mulitipl2 : real;
begin
    multipl1 := real(ifo_izl1);
    multipl2 := real(ifo_izl2);
    f1<=fanout_func(multipl1);
    f2<=fanout_func(multipl2);

    i:= (f1* tr_rq_mn + (0.03*(gauss_rng)));
    j:= (f1* tr_sq_mn + (0.03*(gauss_rng)));
    k:= (f2* tf_rnq_mn + (0.03*(gauss_rng)));
    l:= (f2* tf_snq_mn + (0.03*(gauss_rng)));
    if (r.arr0mn and s.arr1mn) then
        q.d1mn<=min(r.d0mn,s.d1mn)+min(i, j);
        q.arr1mn <= true;
        nq.d0mn<=min(r.d0mn,s.d1mn)+min(k, l);
        nq.arr0mn <= true;
    end if;

    m:= (f1* tf_rq_mn + (0.03*(gauss_rng)));
    n:= (f1* tf_sq_mn + (0.03*(gauss_rng)));
    o:= (f2* tr_rnq_mn + (0.03*(gauss_rng)));
    p:= (f2* tr_snq_mn + (0.03*(gauss_rng)));
    if (r.arr1mn and s.arr0mn) then
        q.d0mn<=min(r.d1mn,s.d0mn)+min(m, n);
        q.arr0mn <= true;
        nq.d1mn<=min(r.d1mn,s.d0mn)+min(o, p);
        nq.arr1mn <= true;
    end if;
end process;
end only;
```

Fig. 7. (continued)

To simulate the circuit few hundred times, a specific VHDL testbench is necessary. This is shown in Fig. 9. Now, for each particular input of the circuit (instantiated few hundred times), for the responses to the logic analysis, and for initialization and for the simulation itself, a specific matrices are formed. The process that performs the timing analysis – log_timing1, is used for determining the minimal delay of the rising edges of the output signals for one asynchronous encoder circuit with five outputs. The results of the analysis are written to a text file.

5. Examples

After completing Monte-Carlo sampling and simulation steps, a huge amount of data can be expected. Since each gate model consists of four parallel processes, for each of the signal transitions, which gives the equivalent to four parallel simulations during each run of the analysis. For example, when the circuit was run through 600 analyses, then the equivalent of $4 \cdot 600 = 2400$ simulations are performed per output. For a circuit that has a small number of outputs, the resulting statistical data can be presented in the form of a histogram.
Fig. 10 shows the results (histogram) obtained for C-element, described as a logic gate. This is the easiest form of statistical representation of the simulation results. It is adequate only for circuits with a small number of outputs.
Considering particular delays of the gate outputs the, the corresponding yield can be easily determined by counting the number of the delays that fall into the acceptable range, and

```
p1: process (in1.d0mn, in1.d1mn, in1.arr0mn, in1.arr1mn,
             in2.d0mn, in2.d1mn, in2.arr0mn, in2.arr1mn)
             variable r,p: real;
             variable multipl : real;
begin
             multipl := real(ifo_izl);
             f<=fanout_func(multipl)
             r:= ((f*1.0) + (0.03*(gauss_rng)));
             p:= ((f*0.9 + (0.03*(gauss_rng)));
             if (in1.arr0mn and in2.arr0mn ) then
                         out1.d0mn   <= min(in1.d0mn, in2.d0mn) + r;
                         out1.arr0mn <= true;
             end if;
             if (in1.arr1mn and in2.arr1mn) then
                         out1.d1mn   <= min(in1.d1mn, in2.d1mn) + p;
                         out1.arr1mn<= true;
             end if;
end process p1;

p2: process (in1.d0mx, in1.d1mx, in1.arr0mx, in1.arr1mx,
             in2.d0mx, in2.d1mx, in2.arr0mx, in2.arr1mx)
             variable r,p: real;
             variable multipl : real;
begin
             multipl := real(ifo_izl);
             r:= (multipl*0.95 + (0.03*(gauss_rng)));
             p:= ((multipl*1.05) + (0.03*(gauss_rng)));
             if (in1.arr0mx and in2.arr0mx) then
                         out1.d0mx   <= max(in1.d0mx, in2.d0mx) + r;
                         out1.arr0mx <= true;
             end if;
             if (in1.arr1mx and in2.arr1mx) then
                         out1.d1mx   <= max(in1.d1mx, in2.d1mx) + p;
                         out1.arr1mx<= true;
             end if;
end process p2;
```

Fig. 8. VHDL timing processes for the two input C-element.

```
initialization: for J in 1 to 600 generate
  encoder_inst: encoder port map (inp => inputs(J), outp => outputs(J));
  inputs(J) <= (others => (0.0, 0.0, true, true, 0.0, 0.0, true, true));
end generate;
log_timing1: process
             use std.textio.all;
             file log: text open write_mode is "encoder_mn_r.statdel";
             variable line_1: line;
             variable I, J:integer;
             variable izlaz: SDA_std_logic_vector(0 to 4);
             variable delay_mn_r : real;
begin
             wait for 1 ps;
             for J in 1 to 600 loop
                         for I in 0 to 4 loop
                                     izlaz := outputs(J);
                                     delay_mn_r:= izlaz(I).d1mn;
                                     write (line_1, delay_mn_r, left, 15);
                         end loop;
                         writeline (log, line_1);
             end loop;
             wait;
end process log_timing1;
```

Fig. 9. Testbench process for writing simulation results for minimal delay of the rising edge of all encoder output signals into a file

dividing it by a total number of simulated circuits. Nevertheless, since the yield is contributed not only with the speed of the circuit but also based on the defects present during the chip manufacturing. Many of those defects are timing defects. Therefore, this methodology can be indirectly used to estimate the yield loss caused by the number of chips that are not operating with needed performance, or are having timing defects due to the fabrication process tolerances.
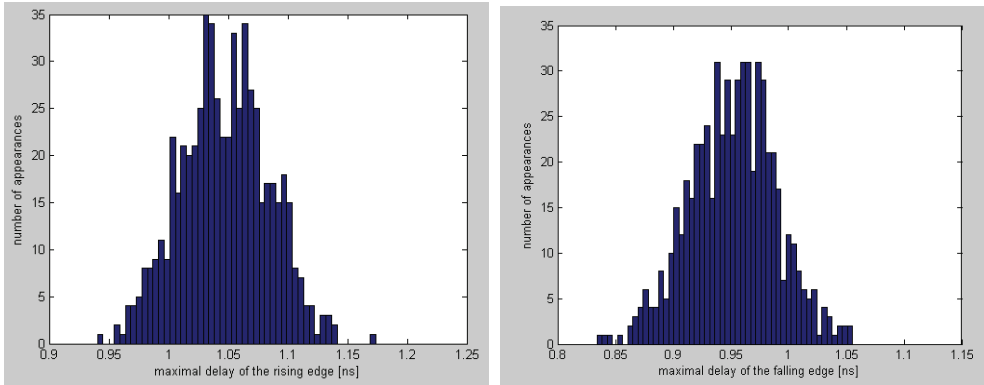
Fig. 10. Histograms of the C-element maximal delays.

Table 3 shows the simulation results of the C-element described at the structural level of abstraction shown in Fig. 2. The first column of the table shows the output number, the second shows the delay type for that output, and the third column gives the topological level of the particular delay type. The next two columns give the worst-case delay estimation results excluding randomness of the delay value, without and with the fanouts of each gate. In this case all fanouts are equal to one, giving the same values in these two columns. The last column shows the results of the statistical analysis of the results. It gives the mean value and the deviation value of the particular delay type.

output	delay type	topol. level	min/ max	fan- out	statistical	
					mean	dev.
	mnr	2	1.9ns	1.9ns	1.854	0.455
	mxr	2	2.0ns	2.0ns	2.041	0.418
1.	mnf	2	1.9ns	1.9ns	1.859	0.441
	mxf	2	2.0ns	2.0ns	2.033	0.439

Table 3. C gate – structural.

Table 4 shows the simulation results of the four stage asynchronous binary counter consisting of 4 T latches. Table 5 gives the timing analysis results for a generalized C-element (Myers 2001), shown in Fig. 11. A simple asynchronous address comparator unit (Myers & Alain 1993) is shown in Fig. 12, and its simulation results are presented in Table 6.

output	delay type	topol. level	min/ max	fan- out	statistical	
					mean	dev.
1.	mnr	4	3.7ns	3.7ns	3.704	0.705
	mxr	4	3.9ns	3.9ns	3.898	0.071
	mnf	4	3.6ns	3.6ns	3.599	0.681
	mxf	4	3.8ns	3.8ns	3.799	0.687

Table 4. T counter.
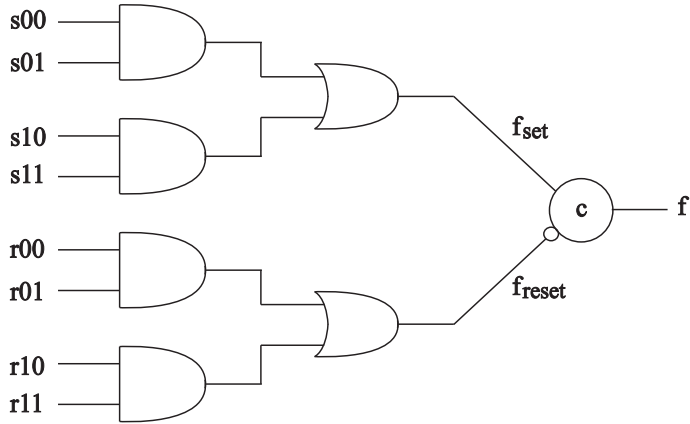
Fig. 11. Generalized C-element.

output	delay type	topol. level	min/ max	fan- out	statistical	
					mean	dev.
1.	mnr	3	2.7ns	2.7ns	2.682	0.058
	mxr	4	4.2ns	4.2ns	4.217	0.071
	mnf	3	3ns	3ns	2.981	0.058
	mxf	4	3.8ns	3.8ns	3.816	0.070

Table 5. Generalized C-element.
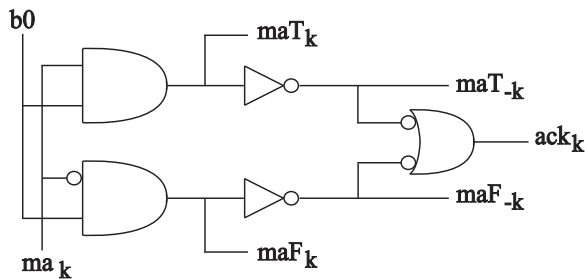
Fig. 12. Address comparator.

out.	delay type	topol. level	min/ max	fan- out	statistical	
					mean	dev.
1.	mnr	1	0.9ns	0.9ns	0.900	0.035
	mxr	1	0.95ns	0.95ns	0.954	0.036
	mnf	1	1.0ns	1ns	1.000	0.036
	mxf	1	1.05ns	1.05ns	1.051	0.035
2.	mnr	2	2.0ns	2ns	1.999	0.050
	mxr	2	2.1ns	2.1ns	2.101	0.050
	mnf	2	1.8ns	1.8ns	1.801	0.053
	mxf	2	1.9ns	1.9ns	1.905	0.049
3.	mnr	3	2.8ns	2.8ns	2.773	0.055
	mxr	4	4.0ns	4ns	4.000	0.072
	mnf	3	2.9ns	2.9ns	2.898	0.060
	mxf	4	4.0ns	4ns	4.000	0.072
4.	mnr	2	2.0ns	2ns	1.999	0.052
	mxr	3	3.05ns	3.05ns	3.051	0.060
	mnf	2	1.8ns	1.8ns	1.802	0.051
	mxf	3	2.95ns	2.95ns	2.954	0.064
5.	mnr	1	0.9ns	0.9ns	0.900	0.036
	mxr	2	2.0ns	2ns	2.002	0.053
	mnf	1	1.0ns	1ns	1.006	0.035
	mxf	2	2.0ns	2ns	2.002	0.049

Table 6. Address comparator.

Finally, Fig. 13 shows one complex asynchronous encoder circuit described in (Kondratyev & Lwin 2002), while table 7 gives its timing analysis results.
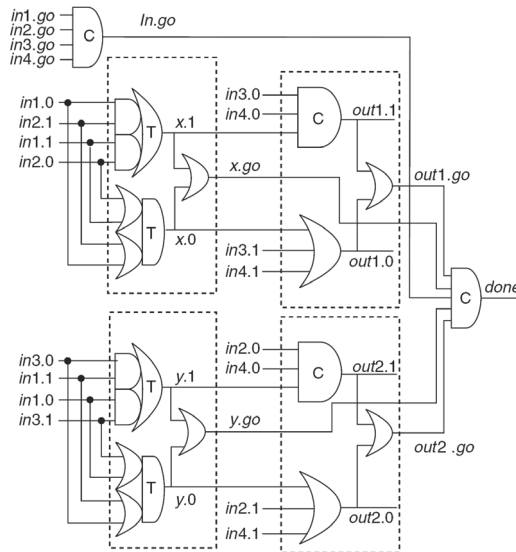
Fig. 13. Encoder circuit

Table 8 gives the simulation run times and the corresponding allocated memory for all these circuits. These results are for 600 timing simulations per circuit, achieved on an AMD Athlon processor at 1.14 GHz with 1 GB RAM.

out.	delay type	topol. level	min/ max	fan- out	statistical mean	dev.
	mnr	1	0.90ns	0.9ns	0.902	0.036
	mxr	3	2.85ns	3.8ns	3.823	0.055
1.	mnf	1	1.00ns	1.0ns	0.999	0.036
	mxf	3	3.15ns	4.2ns	4.218	0.060
	mnr	1	0.9ns	0.9ns	0.900	0.038
2.	mxr	3	2.95ns	3.9ns	3.917	0.060
	mnf	1	1.00ns	1.0ns	0.998	0.035
	mxf	3	3.05ns	4.1ns	4.121	0.067
	mnr	2	1.80ns	1.8ns	1.802	0.049
3.	mxr	5	4.95ns	5.9ns	5.957	0.066
	mnf	2	2.00ns	2.0ns	1.999	0.049
	mxf	5	5.15ns	6.2ns	6.263	0.066
	mnr	1	0.90ns	0.9ns	0.901	0.036
4.	mxr	3	2.85ns	3.8ns	3.819	0.058
	mnf	1	1.00ns	1.0ns	1.001	0.037
	mxf	3	3.15ns	4.2ns	4.222	0.055
	mnr	1	0.90ns	0.9ns	0.897	0.037
5.	mxr	3	2.95ns	3.9ns	3.920	0.060
	mnf	1	1.00ns	1.0ns	0.999	0.036
	mxf	3	3.05ns	4.1ns	4.120	0.069

Table 7. Encoder.

circuit	CPU time [s]	allocated memory [kB]
C element	6.3	19.697
T counter	7.7	20.277
Addr. comp	8.2	25.734
Gen C elem	10.5	40.443
Encoder	28.5	92.583

Table 8. Simulation run times and allocated memory.

We have chosen to compare our method with Monte-Carlo analysis based on the application of standard logic simulations. The merits for comparison are estimation times, complexities of the approach (number of simulator runs or simulations), required resources. Considering the quality of the results, it is very difficult to compare these two approaches, since classical Monte-Carlo analysis requires a huge amount of time for data processing, and obtaining it in a form that can be comparable with our method. The Table 9 gives the results of comparison.

circuit	allocated memory [kB]	Estimation time [s] (CPU time)	Number of simulations
C element	444	1	16
T counter	375	2	2
Addr. comp	446	3	4
Gen C elem	452	4	255
Encoder	480	436	1048576

Table 9. Allocated memory, simulation run times and number of simulations for Monte-Carlo analysis.

Beside the columns that show simulation run times and the corresponding allocated memory for this approach, the last column in the table shows how many Monte-Carlo standard simulations are performed. These simulations are performed for all possible that for complex circuit the number of possible input combinations can be high enough to enable statistical processing of the results. But, with a simple circuits, where the number of inputs is not large enough (if all possible combinations of inputs signals is less than few hundreds), statistical results processing is impossible. The major improvement achieved with the proposed method is that all possible input vector combinations are analysed with only one run of the simulator, that is, with only one timing analysis cycle. Only one run of the simulator is needed for obtaining all possible worst-case delays.

The proposed method was previously verified on a set of combinational ISCAS benchmark circuit (Sokolović & Litovski 2005; Lin & Davoodi 2008; iscas89.html --), with up to 9000 gates, and no bottleneck resulted.

6. Conclusion

A new concept for asynchronous circuit delay analysis was presented in this paper. The method is based on a specific signal and gate modelling. Signals in the circuit do not carry the information about the logic value of the signal, but the information about the delay and about the arrival of the information instead. Specific attributes of the signals are introduced in order to implement that property. The method generates and exploits information about the fanout of each gate implemented in a complex digital system. It has the ability to deal with sequential elements containing loops and feedbacks. It was implemented in VHDL and verified on some particular asynchronous circuits. This way of delay modelling allows for the delay to become a random variable so that the delay estimation method was incorporated into Monte-Carlo analysis to produce statistical worst-case delay distributions. The obtained delay distributions of the output signals can be used for any kind of timing analysis which will consider the simulated circuit as a whole. This also includes the for example incremental timing analysis. Having in mind the number and the complexity of the gate processes, and the number of simulations required to achieve high accuracy, the simulation times for each gate, as well as the allocated memory show high efficiency of the proposed method, comparing to very intensive and complex Monte-Carlo analysis.

7. References

Agarwal, A. et al. (2003). Statistical delay computation considering spatial correlations. *Proc. of the ASP-DAC 2003*, pp. 271 – 276, ISBN: 0-7803-7659-5, Kitakyushu, Japan, January 2003.

Agarwal, A. et al. (2005). Circuit Optimization using Statistical Static Timing Analysis. *Proc. of the 42nd Annual Conf. on Design Automation*, pp. 321-324, ISBN: 1-59593-058-2, San Diego, California, 2005.

Andrikos, N. (2007). A fully-automated desynchronization flow for synchronous circuits. *Proceedings of the 44th annual conference on Design automation*, pp. 982-985, San Diego, California, Juns 2007, ISSN: 978-1-59593-627-1.

Cheng, T. & Agrawal, V. (1989). *Unified Methods for VLSI Simulation and Test Generation*, Kluwer Academic Publishers, Boston/Dordrecht/London, 1989, ISBN: 978-0-7923-9025-1.

Cortadella, J. et al. (1999). Synthesis of asynchronous control circuits with automatically generated relative timing assumptions. *Proceedings of the 1999 IEEE/ACM international conference on CAD*, pp. 324-331, ISBN:0-7803-5832-5, 1999, IEEE Press Piscataway, NJ, San Jose, California.

Cortadella, J. et al. (2006). Desynchronization: Synthesis of Asynchronous Circuits From Synchronous Specifications. *IEEE Transactions on CAD of Integrated Circuits and Systems*, Vol. 25, No. 10, October 2006, pp. 1904-1921, ISSN: 0278-0070. http://courses.ece.uiuc.edu/ece543/iscas89.html

Davis, A. & Nowick, S. (1997). *An Introduction to Asynchronous Circuits Design*, Technical Report UUCS-97-013. Septemper 1997, Computer Science Department, University of Utah.

Kondratyev, A. & Lwin, K. (2002). Design of Asynchronous Circuits Using Synchronous CAD Tools. *IEEE Design & Test*, Vol. 19, Issue 4, July 2002, pp. 107 – 117, ISSN: 0740-7475.

Lewis, M. & Brackenbury, L. (2001). CADRE: A Low-power, Low-EMI DSP Architecture for Digital Mobile Phones. *VLSI Design, special issue on low power system design*, Vol. 3, Issue 12, September 2001, pp. 333-348.

Lin, X. & Davoodi, A. (2008). Robust Estimation of Timing Yield with Partial Statistical Information on Process Variations. *Proc. of A Quality Electronic Design, ISQED 2008*, pp. 156 – 161, ISBN: 978-0-7695-3117-5, San Jose, California, USA, March 2008.

Litovski, V. & Zwolinski, M. (1997). *VLSI circuit simulation and optimization*. Chapman and Hall, London, ISBN: 0412638606, 1997.

Maksimović, D. (2000). *Logic Simulation – Estimation of the Worst-case characteristics of the Designed Digital Circuits*. PhD thesis, Faculty of Electronic Engineering, University of Niš, Serbia, June 2000.

Maksimović, D. & Litovski, V. (1999). Tuning Logic Simulators for Timing Analysis. *Electronic Letters*, Vol. 35, No. 10, May 1999, pp. 800-802, ISSN: 0013-5194.

Maksimović, D. & Litovski, V. (2002). Logic Simulation Methods for Longest Path Delay Estimation. *IEE Proc. Computers and Digital Technique*, Vol. 149, No. 2, March 2002, pp. 53-59, ISSN: 1350-2387.

Mak, T.M. et al. (2004). New Challenges in Delay Testing of Nanometer, Multigigahertz Designs. *Design & Test of Computers*, Vol. 21, Issue 3, May-June 2004, pp. 241-248, ISSN: 0740-7475.

Martin, A. & Nystrom, M. (2006). Asynchronous Techniques for System-on-Chip Design. *Proceedings of the IEEE*, pp. 1089 – 1120, Vol. 94, Issue 6, June 2006, ISSN: 0018-9219.

Myers, C. (2001). *Asynchronous Circuit Design*. University of Utah, John Wiley & Sons, Inc. May, 2001, ISBN: 0-471-41543-X.

Myers, C. & Alain, M. (1993). *The Design of Asynchronous Memory Management Unit. Technical Report CS-TR-93-30*, California Institute of Technology, April 1993.

Sokolovic, M. & Litovski, V. (2005). Using VHDL Simulator to Estimate Logic Path Delays in Combinational and Embedded Sequential Circuits, *Proceedings of IEEE Region 8 EUROCON 2005*. Conference, pp. 547-550, ISBN: 1-4244-0049-X, November 2005, Belgrade.

Sokolovic, M., Litovski, V. & Zwolinski, M. (2009) New concepts of worst-case delay and yield estimation in asynchronous VLSI circuits. *Microelectronics Reliability*, Vol. 49, Issue 2, pp. 186-198. ISSN 0026-2714.

Sparso, J. (2006). *Asynchronous Circuit Design – A Tutorial.* April 2006, Technical University of Denmark.

Spence, R. & Soin, R. (1988). *Tolerance design of Electronic circuits*. Addison-Wesley Publ. Comp. Wokingham, England, ISBN: 978-1-86094-040-8, 1988.

Zwolinski, M. (2004). *Digital System Design with VHDL*, Prentice Hall, London, UK, 2004, ISBN 0-13-039985-X.

Neuron Network Applied to Video Encoder

Branko Markoski[1], Jovan Šetrajčić[2], Jasna Mihailović[1], Branko Petrevski[3],
Miroslava Petrevski[4], Borislav Obradović[5], Zoran Milošević[5],
Zdravko Ivanković[6], Dobrivoje Martinov[1] and Dušanka Tesanović[7]

[1] *University of Novi Sad, Technical Faculty "Mihajlo Pupin" Zrenjanin,*
[2]*Faculty of Sciences University of Novi Sad*
[3]*University of Novi Pazar,*
[4]*University of Belgrade,*
[5]*University of Novi Sad,*
[6]*Faculty of Technical Sciences , University of Novi Sad,*
[7]*Oncology Institute of Vojvodina, Sremska Kamenica,*
Serbia

1. Introduction

There are a number of problems in science and technology that demand separating useful information from certain content. For many of those problems, standard techniques, as signal processing technique, shape recognition, system control theory, artificial intelligence etc., have shown as inadequate. Neural networks are a way to solve these problems in a way they are solved in human brain. Same as the human brain, neural networks are able to learn from given data, and afterwards, when they meet the same or similar data they may give the same or approximate result.

There are several types of transfer functions: sigmoid, logistic sigmoid, linear, semilinear, threshold, Gauss' function. Figure 1 shows the graph for one of most used transfer functions:
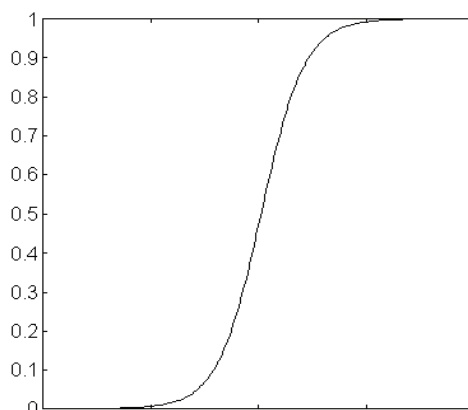

Fig. 1. Logistic sigmoid function

Multilayer neural network with signal propagation forward is very often used architecture (*Bourlard, H at all, 2002*). In it, signals are propagating only forward, and neurons are organized in layers. Most important properties of multilayered networks with propagation forward are given in these two theorems:

1. Multilayered network with single hidden layer may uniformly approximate any real continual function with arbitrary precision at the final real axis.
2. Multilayered network with two hidden layers may uniformly approximate any real continuous function at the final real axis.

Input layer receives data from environment. Hidden layer receives data from previous layer (in this case, outputs from input layer) and gives output depending from sum of input weights. For more complex problems, sometimes it is necessary to have more than one hidden layer. Output layer computes neural network outputs from sum of weights and transfer function.

H.263 is an international standard for video stream compression, widely used in telecommunication systems (ITU, 1995). There are several additions by ITU-T recommendation h.263, aimed at broadening of supported picture formats and video stream compression quality (ITU, 1996).

Enhancement of h.263 standard, presented in this paper, is related to application of artificial neural network (ANN) instead of standard DCT code, for sequences full of quick motion details.

In section 2, a short description of h.263 standard is given. Section 3 describes training code for neural network used. Section 4 describes a way in which ANN is applied as an addition to existing h.263 standard. In section 5, results of experiments showing effects of this approach at quality and compression level of test sequence are presented.

2. H 263 video encoder

Compression of a video signal is the key component in modern telecommunication services, as videotelephony and video conferences, in modern digital TV systems with normal and high resolution, and in numerous multimedia services. The reason is that – without compression – digital video signal consists of huge amount of data. Another problem in multimedia systems is a speed of reading and transferring data from compact disc to computer memory, and in fastest systems, it is up to 4 Mb/s. Having in mind that coding of a video signal is a topic of research for more than two decades, a large number of algorithms had been developed, implemented and tested on existing communication channels. In order to enable connection of equipment form different manufacturers, several international companies defined standards for compression and transfer of video signal. Best known are H.261 and H.263 for transfer of videoconferences and videophony, as well as MPEG standards (MPEG-1, MPEG-2 and MPEG-4) intended for standardization of multimedia systems and digital television (Schäfer, R., T. Sikora, 1995). Three-dimensional (3D) compression of a video signal is a generalization of two-dimensional video signal compression principle. Most frequent way to realize 3D compression of a video signal is the 3D transformation encoding based on DCT. For application of this method, video signal is divided in blocks with dimensions $M \times N \times P$, where M, N and P, respectively, are the horizontal, vertical and time dimensions of a block (Boncelet C. 2005). On every block 3D DCT is applied, and obtained DCT coefficients are being quantumized. As in 2D DCT, only coefficients with very small index values have significant values (Roese, J.A., at all 1997). In

H.261 standard, two picture formats are defined (Markoski, B. & Đ. Babić, 2007). Therefore, for transmission of both formats of video signal by ISDN channels, it is necessary to achieve considerable level of compression (typically about 100 times). Since QCIF format is mostly intended for videophony applications, where mostly only a face of the other person is visible, frame frequency is usually decreased to 10 frames/s. H.261 standard defines algorithms for eliminating redundancy, quantumization algorithms, structure of coders and decoders, as well as data structure (Rijkse, K, 1995). It is interesting that standard does not demand using a certain algorithm for movement estimation, but it is important only to determine and transmit block movement vectors. A mechanism of regulation of bit-stream is also not demanded, but it is determined by choosing the way of processing and a way of deciding whether a block is being transmitted or not. In practice, implementation known as Referent model 8 (COST211bis/SIM89/37, 1989) is used most frequently, and it was used in standards testing.

H.263 standard is intended for standardization of picture transmission by standard telephone commutated lines wit bit-stream under 64 Kb/s, which was not covered by any standard (Rijkse, K., 1995, Girod, B, at all, 1996). It was produced by modifications of existing H.261 standard. Due to very tight deadlines in preparing the standard, original text of standard (Rijkse, K., 1995), defines only most necessary improvements of H.261 standard, but a possibility for further improvements is left open.

The basic difference between H.261 and H.263 standards is in target bit-stream (A.Amer, E. Dubouis, 2005). H.261 was supposed to be used for picture transmission over 64 Kb/s, while H.263 was supposed to be used under 64 Kb/s, most often in 22 Kb/s. In order to realize this goal, four small improvements were done to algorithms prescribed by H.261 standard. Although no one of those, per se, contributes much to total performances, all four together improve performances considerably (LeGall, D.J 1992).

H.263 recommendation is defined by International telecommunications society - telecommunication standard section (ITU-T, 1996). This recommendation standardizes a video stream compression process, defining syntax of compressed data format Compression is necessary in order to translate a conventional video stream into a shape available to computer applications under present limitations. H.263 uses compression code basically similar to JPEG (Joint Photographic Experts Group) and to MPEG (Motion Picture Experts Group codes) (ITU-T, 1995). Video stream is being compressed by a transformation sequence of every single picture.

H.263 video stream is organized in several layers, as shown in Figure 2. The highest layer, picture layer, defines basic properties of the video stream as picture size and coding system. Next layer is a group of blocks layer, enabling unique interpretation of spatially close blocks. Two lowest layers are macroblock layer and block layer, representing code interpretation of a picture. Every picture within video sequence is coded in one of three possible ways of coding: intra (I), inter (P) or bidirection (PB) coding. I-pictures are coded similar as in JPEG standard. P-pictures are envisaged on the basis of previously coded picture, and PB-pictures are envisaged on the basis of blocks from previous and next picture. Coding of every picture consists from its partition into macroblocks and special coding for every one of those. Every macroblock presents a 16x16-pixel zone and is a basic unit for motion compensation. Macroblock consists from 6 blocks: 4 luminent and 2 chrominent blocks. These blocks (8x8 pixels) are basic units for DCT (Discrete cosine transform).

Motion compensation is being done in order to remove time sameness between adjacent pictures in a video sequence. In this way, instead of complete picture, only information on

detected changes and a way of their movement (move vectors) are transmitted. To avoid error accumulation, together with move vectors an error signal is coded, which is a difference between reconstructed and actual picture. The DCT transformation is being done to thus obtained error of move estimation. DCT transformation is being done on 8x8-pixel blocks, resulting in 64 transformation coefficients. The block energy is, after transformation, concentrated in few coefficients, corresponding to low-frequency part of range. Therefore, quantization of these coefficients is possible with relatively small error. Most of DCT coefficients are equalized with zero, which lowers information quantity needed for picture reconstruction.
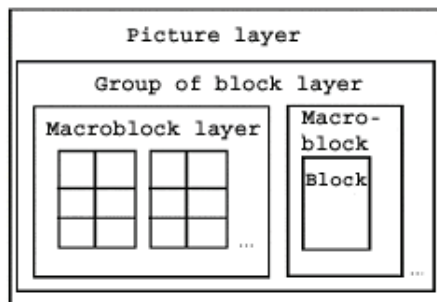
Fig. 2. Structure of H.263 video stream

At the end of coding process, obtained information is statistically coded (Huffman and run-length coding) and written in format defined by h.263 syntax of video stream.

3. Neural network

The discipline we know today as neural networks originated as a result of fusing several quite different ways of research: signal processing, neurobiology and physics (Haykin S, 1994). Neural networks are a typical example of an interdisciplinary discipline (L. Faulsett 1995). On the one hand, this is an attempt to understand workings of a human brain, and on the other to apply the newly acquired knowledge in processing complex information (Lippmann, R. P. 1987). There are other progressive, non-algorithmic systems, as learning algorithms, genetic algorithms, adaptive memory, associative memory, fuzzy logic. General opinion is that neural networks are presently most mature and most applicable technology (Barsterretxea, att all, 2002).

Conventional computers work on logic basis, deterministically, sequentially or wit a very low level of parallelism. Software written for such computers must be almost perfect in order to work appropriately. This requires long and costly designing and testing process.

Neural networks belong to parallel asynchronous distributed processing category. The network is tolerant on damages or falling out of function for a relatively low number of neurons. The network is also tolerant to presence of noise in input signal. Every memory element is delocalized - situated in network as a whole and it is impossible to identify in which part it is stored. Classic addressing is nonexistent, since memory is approached using contents, and not the address (S.P. Teeuwsen, at all. 2003).

Basic component of neural network is a neuron, as shown in figure 3:
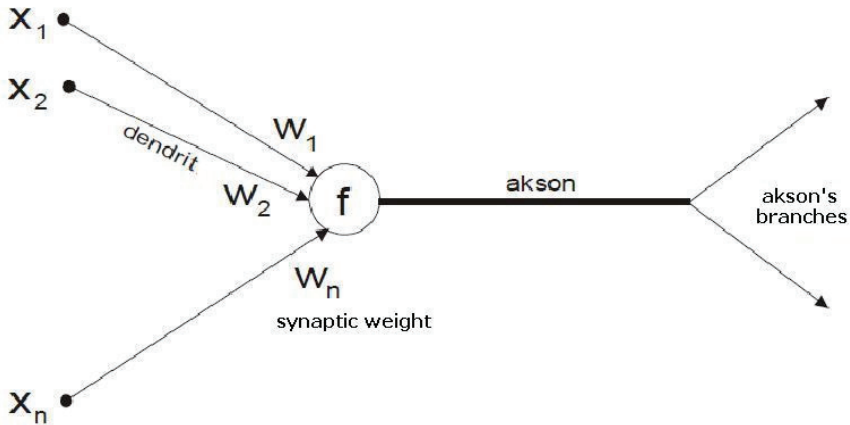
Fig. 3. Basic component of neural network

Dendrites are inputs into neuron. Natural neurons have even hundreds of inputs. Point where dendrites are touching the neuron is called a synapse. Synapse is characterized by effectiveness, called synaptic weight. Neuron output is formed in a following way: signals on dendrites are multiplied by corresponding synaptic weights, results are added and if they exceed threshold level on the result is applied a transfer function of neuron, which is marked f on a figure. Only limitation of transfer function is that it must be limited and non-decreasing. Neuron output is routed to axon, which by its branches transfers result to dendrites. In this way, output from one layer of network is transferred to the next one.

In neural networks, three types of transfer functions are presently being used:

- jumping
- logical with threshold
- sigmoid

All three types are shown in figure 4:
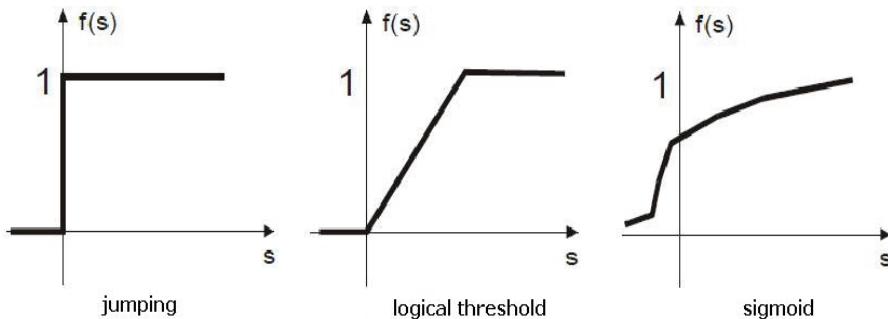
Fig. 4. Three types of transfer functions

The neural network has unique multiprocessing architecture and without much modification, it surpasses one or even two processors of von Neumann architecture characterized by serial of sequential information processing (S.P. Teeuwsen at all, 2003). It has ability to explain every functional dependence and to expose a nature of such

dependence with no need to external incentives, demands for building a model or its change. In short, neural network may be considered as a black box capable of predicting output pattern or a signal after recognizing given input pattern. Once trained, it may recognize similarities when a new input signal is given, which results in predicted output signal. There are two categories of neural networks: artificial and biological ones. Artificial neural networks are in structure, function and in information processing similar to biological ones. In computer sciences, neural network is an intertwined network of elements that processes data. One of more important characteristics of neural networks is their capability to learn from limited set of examples . The neural network is a system comprised of several simple processors (units, neurons), and every one of them gas its local memory where it stores processed data. These units are connected by communication channels (connections). Data exchanged by these channels are usually numerical ones. Units are processing only their local data and inputs obtained directly through connection. Limitations of local operators may be removed during training. A large number of neural networks created as models of biological neural networks. Historically speaking, inspiration for development of neural networks was in desire to construct an artificial system capable of refined, maybe even "intelligent" computations in a way similar to that in human brain. Potentially, neural networks are offering us a possibility to understand functioning of human brain. Artificial neural networks are a collection of mathematical models that simulate some of observed capabilities in biological neural systems and has similarities to adaptable biological learning. They are made of large number of interconnected neurons (processing elements) which are, similarly to biological neurons, connected by their connections comprising of permeability (weight) coefficients, whose role is similar to synapses. Most of neural networks have some kind of rule for "training", which adjusts coefficients of inter-neural connections based on input data (Cao J, at all 2003). Large potential of neural networks lays in possibility of parallel data processing, to compute components independent from each other. Neural networks are systems made of several simple elements (neurons) that process data parallely.

There are numerous problems in science and engineering that demand extracting useful information from certain content. For many of those problems, standard techniques as signal processing, shape recognition, system control, artificial intelligence and so on, are not adequate. Neural networks are an attempt to solve these problems in a similar way as in human brain. Like human brain, neural networks are able to learn from given data; later, when they encounter the same or similar data, they are able to give correct or approximate result.

Artificial neuron, based on sum input and transfer function, computes output values. The following figure shows an artificial neuron:
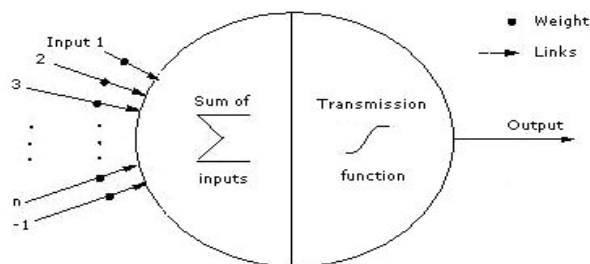

Fig. 5. Artificial neuron

The neural network model consists of:
- neural transfer function
- network topology, i.e. a way of interconnecting between neurons,
- learning laws

According to topology, networks are differing by a number of neural layers. Usually each layer receives inputs from previous one, and sends its outputs to the next layer. The first neural layer is called input layer, the last one is output layer and other layers are called hidden layers. Due to a way of interconnecting between neurons, networks may be divided to recursive and non-recursive ones. In recursive neural networks, higher layers return information to lower ones, while in non-recursive ones there is a signal flow only from lower to higher layers.

Neural networks learn from examples. Certainly there must be many examples, often even tens of thousands. Essence of a learning process is that it causes corrections in synaptic weights. When new input data cause no more changes in these coefficients, it is considered that a network is trained to solve a problem. Training may be done in several ways: controlled training, training by grading and self-organization.

No matter which learning algorithm is used, processes are in essence very similar, consisting from following steps:
1. A set of input data is presented to a network.
2. Network processes information and remembers result (this is a step forward).
3. The error value is calculated by subtracting obtained result from the expected one.
4. For every node a new synaptic weight is calculated (this is a step back).
5. Synaptic weights are changed, or old ones are left and new ones are remembered.
6. On network inputs, a new set of input data is brought to network inputs and steps 1-5 are repeated. When all examples are processed, synaptic weights values are updated and if an error is under some expected value the network is considered trained.

We will consider two training modes: controlled training and self-organization training.

The back-propagation algorithm is the most popular algorithm for controlled training. The basic idea is as follows: random pair of input and output results is chosen. Input set of signals is sent to the network by bringing one signal at each input neuron. These signals are propagating further through the network, in hidden layers, and after some time a results show on output. How has this happened?

For every neuron an input value is calculated, in a way we previously explained; signals are multiplied by synaptic weights of corresponding dendrites, they are added and a neuron's transfer function is being applied to obtained value. The signal is propagated further through the network in a same way, until it reaches output dendrites. Then a transformation is done once again and output values are obtained. The next step is to compare signals obtained on output axon branches to expected values for given test example. Error value is calculated for every output branch. If all errors are equal to zero, there is no need for further training – network is able to perform expected task. However, in most cases error will be different from zero. Then a modification of synaptic weights of certain nodes is called for.

Self-organized training is a process where a network finds statistical regularities in a set of input data and automatically develops different behavior regimes depending on input. For this type of learning, the Kohonen algorithm is used most often.

The network has only two neural layers: input and output one. Output layer is also called a competitive layer (reason will be explained later). Every input neuron is connected to every

neuron in output layer. Neurons in output layer are organized in two-dimensional matrix (Zurada, J. M.1992).

Multilayer neural network with signal propagation forward is one of often used architectures. Within it, signals are propagating only ahead, and neurons are organized in layers. Most important properties of multilayer networks with signal propagation forward are given as following theorems:

1. Multilayer network with a single hidden layer may uniformly approximate any real continual function on the finite real axis, with arbitrary precision.
2. Multilayer network with two hidden layers may uniformly approximate any real continual function of several arguments, with arbitrary precision.

Input layer receives data from environment. Hidden layer receives outputs of a previous layer (in this case, outputs of input layer) and, depending on sum of input weights, gives output. For more complex problems, sometimes is necessary more than one hidden layer. Output layer computes, on the basis of weight sum and transfer function, outputs from neural network.

The following figure shows a neural network with one hidden layer.
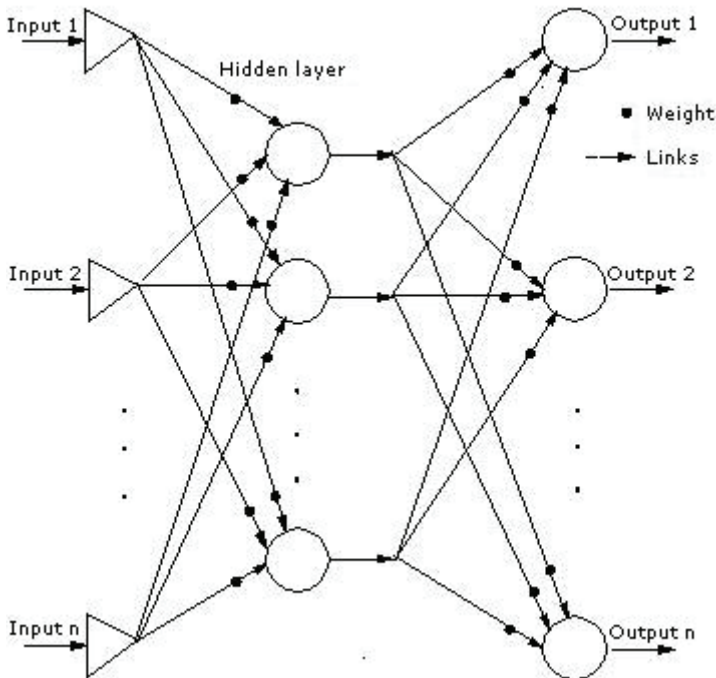
Fig. 6. Neural network with one hidden layer and with signal propagation forward

In this work, we used Kohonen neural network, which is a self-organizing map of properties, belonging to a class of artificial neural networks with unsupervised training (Kukolj D., Petrov M., 2000). This type of neural network may be observed as topologically organized neural map with strong associations to some parts of biological central nervous system. The notion of topological map understands neurons that are spatially organized in

maps that guard, in a certain way, the topology of input space. Kohonen neural network is intended for following tasks:

- Quantumization of input space
- Reduction of output space dimension
- Preservation of topology present within structure of input space.

Kohonen neural network is able to classify input samples-vectors, without need to recognize signals for error. Therefore, it belongs to group of artificial neural networks with unsupervised learning. In actual use of Kohonen network in algorithm for obstacle avoidance, network is not trained but enhancement neurons are given values calculated in advance. Regarding clusterization, if a network may not classify input vector to any output cluster, than it gives data regarding how much the input vector is similar to every of clusters defined in advance. Therefore, this paper uses Fuzzy Kohonen neural clusterization network (FKCN).

Enhancement of h.263 code properties is attained by generating a prototype codebook, characterized by highly changeable differences in picture blocks. Generating codebook is attained by training of self-organizing neural network (Haykin, 1994; Lippmann, 1987; Zurada, 1992). After realization of original training concept (Kukolj and Petrov, 2000), a single-layer neural network is formed. Every node of output ANN layers represents a prototype within codebook. Coordinates of every *and* node within network is represented by difficulty synaptic coefficients w_i. After initialization, the code proceeds in two iterative phases.

First, closest node for every sample is found, using Euclidean distance, and node coordinates are computed as arithmetic means of coordinates for samples clustered around every node. The node balancing procedure is continued by confirmation of following condition:

$$\sum_{i=1}^{K} \left| w_i - w_i' \right| \le T_{SKG} , \tag{1}$$

where T_{ASE} is equal to a certain part of present value of average square error (ASE). Variables w_i and w_i' are synaptic vectors of node *and* in present and previous code iteration. If above condition is not met, this step is repeating, otherwise the procedure is proceeding further.

In a second step, so-called dead nodes are considered, i.e. nodes that have no assigned samples. If there are no dead nodes, T_{ASE} has very low positive value. If dead nodes are existing, value q for pre-defined number of nodes ($q<<K$), with maximum ASE value, is found. Then dead node is moved near one randomly chosen node from q nodes with maximum ASE values. Now new coordinates of the node are as follows:

$$w_i^{new} = w_{\max}^q + \delta , \quad i = 1,..., K , \tag{2}$$

where $w_{max}{}^q$ is location of chosen node between q nodes with highest ASE, $w_i{}^{new}$ is new node location, and $\delta = [\delta_1, \delta_2, ...,\delta_n]^T$ are small random numbers. The process of deriving new coordinates for dead nodes (2) is repeated for all of those nodes. If maximal number of iteration is achieved, or if in previous and present iteration number of dead nodes is equal to zero, code ends. Otherwise it returns to first stage.

4. Application of ANN in video stream coding

The basic way of removing spatial sameness during coding in h.263 code is using of transformation (DCT) coding (Kukolj at all, 2006). Instead of being transferred in original shape after DTC coding, data are presented as the coefficient matrix. Advantage of this transformation is that obtained coefficients could be quantized, which increases the number of coefficients with zero value. This enables removal of excess bits using entropy coding on the bit repeating basis (run-length).

This approach is efficient in cases when a block is poor in details, so the energy is localized in a few first coefficients of DCT transformation. But, when a picture is rich in details, the energy is equally distributed to other coefficients as well, so after quantization we do not obtain consecutive zero coefficients. In these cases, coding of those blocks uses much more bits, since bit-repetition coding could not be efficiently used. Basic way of compression factor control in this case is increase of quantization step, which brings to loss of small details in reconstructed block (block is blurred) with highly expressed block-effect on reconstructed picture (Cloete, Zurada, 2000).

Enclosed improvement of h.263 code is based on detection of these blocks and their replacement by corresponding ANN node. Basic criterion for critical blocks detection is the length of generated bits, using the standard h.263 code.

As training set for ANN we used a set of blocks, which are, during the standard h.263 process, represented with more than 10 bits. Boundary level of code length, N=10 bits, have been chosen with purpose to obtain codebook with 2^N=1024 prototypes.

In order to obtain training set, video sequences from "Matrix" movie were used, as well as standard CIF test video sequences "Mobile and calendar" (Hagan , at all 2002). A training set from about 100,000 samples was obtained for ANN training. As a training result, training set was transformed into 1024 codebook prototypes with least average square error regarding the training set.

The modified code is identical with standard way of h.263 compression of video stream until the stage of move vector compensation. Every block is coded by the standard method (using DCT transformation and coding on the basis of bit repeating), and than decision on application of ANN instead of standard approach is made. Two conditions must be fulfilled in order to use the network.

1. **Condition of code length**: whether standard approach gives the code longer of 10 bits as the representation of observed block. This is the primary condition, providing that ANN is used only in cases when standard code does not give satisfying compression level.

2. **Condition of activation threshold**: whether average square error, obtained using neural network, is within boundaries:

$$SKG_{ANM} \leq k \cdot SKG_{DCT} \tag{3}$$

where:
ASE_{INN} - average square error obtained using ANN;
ASE_{DCT} - average square error obtained using the standard method
k - activation threshold for the network (1.0 - 1.8).

On the basis of these conditions, choice between standard coding method and ANN application is being made.
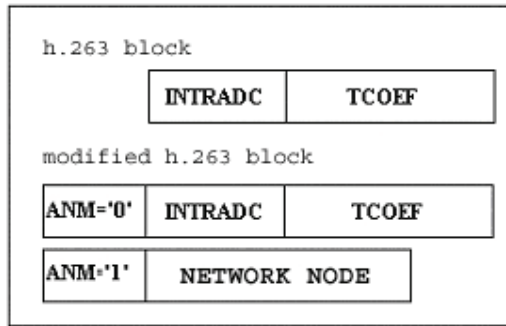
Fig. 7. Changes in h.263 stream format

Format of coded video stream is taken from h.263 syntax (ITU-T, 1996). Data organization in levels has been kept, as well as a way of representation for block moves vector. A modification of syntax of block level was done, introducing additional field (1 bit length) in header of block level (Fig. 3), in order to note which coding method was used in certain blocks.

5. Results of testing

Testing of the described modified h.263 code was done on dynamic video sequence from the "Matrix" movie (525 pictures, 640x304 points). Basic measured parameters were the size of coded video stream and error within coding process. Error is expressed as peak signal to noise ratio (PSNR):

$$PSNR = 10 \cdot \log \frac{255 \cdot 255}{SKG_l} \tag{4}$$

where ASE_l is average square error of reconstructed picture in comparison to the original one.

During the testing, quantization step used in standard DCT coding process and activation threshold of neural network (expressed as coefficient k in formula (4)) were varied as parameters.

The standard h.263 was used as a reference for comparison of obtained results.

Two series of tests were done. In first group of tests, quantization step has been varied, while activation threshold was constant (k=1.0). In second group of tests, activation threshold has been varied, with constant value for quantization step (1.0).

Figure 8 shows the size of obtained coded stream for both methods. It could be seen that compression level obtained using ANN is higher than one obtained using standard h.263 code. For higher quantum values, comparable sizes of stream are obtained, since in this case condition of code length for ANN use was not met, so the coding is being done almost without ANN.

Figure 9. shows the size of error within coded video stream for both methods. It could be seen that, for same values of used quantum, ANN has insignificantly higher error than the standard h.263 approach.
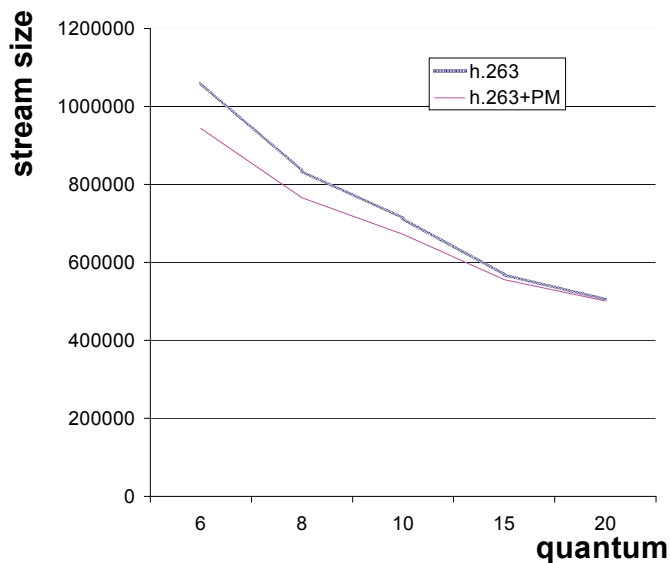
Fig. 8. Dependence of stream size from quantum
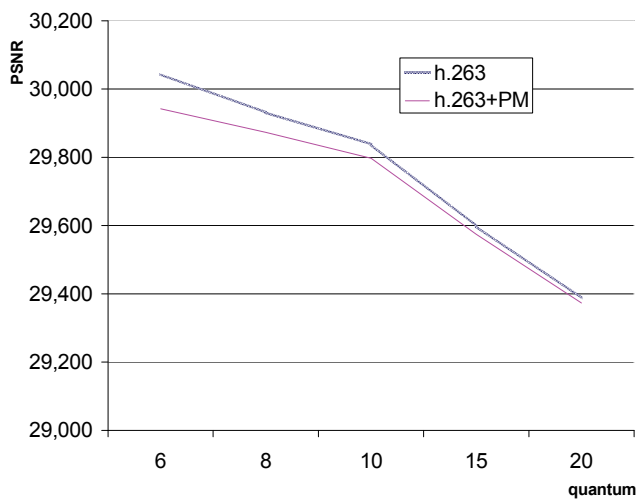
Fig. 9. Dependence of PSNR from quantum

Figures 10. and 11. show results obtained by varying activation threshold of neural network between 1.0 and 1.8. Due to clearness, results are shown for the first 60 pictures from the test sequence. Sudden peaks correspond to changes of camera angle (frame).
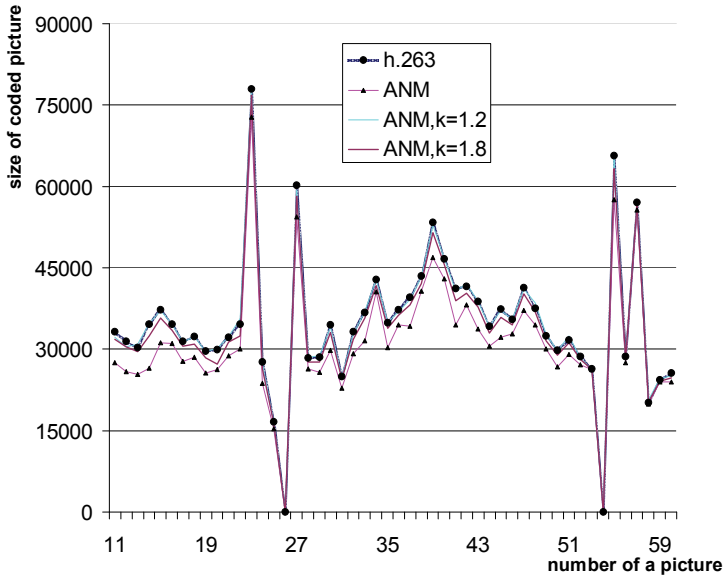
Fig. 10. Dependence of compression from the ANN activation threshold
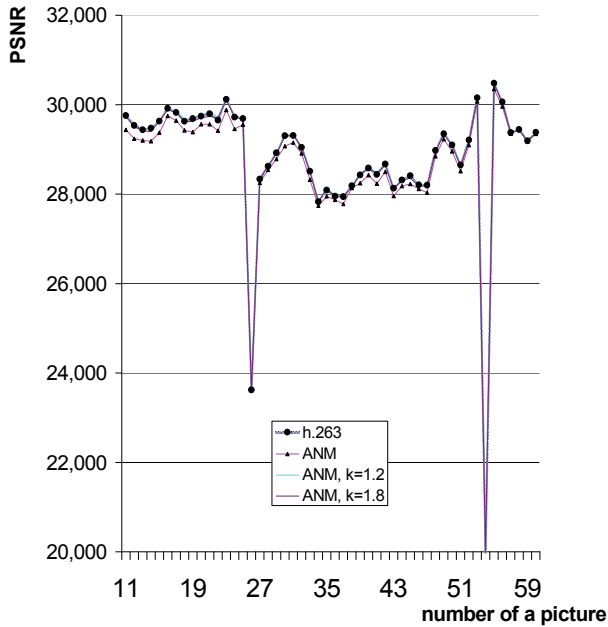
Fig. 11. Dependence of PSNR from the ANN activation threshold

Obtained results show that with increase of neural network activation threshold, compression level decreases and quality of video stream increases. Further increase of activation threshold (above k=1.8), effect of ANN on coding becomes minor.

6. Conclusion

The paper deals with h.263 recommendation for the video stream compression. Basic purpose of the modification is stream compression enhancement with insignificant losses in picture quality. Enhancement of the video stream compression is achieved by artificial neural network. Conditions for its use are described as condition of code length and condition of activation threshold. These conditions were tested for every block within picture, so the coding of the block was done by standard approach or by use of neural network. Results of testing have shown that by this method the higher compression was achieved with insignificantly higher error in comparison to the standard h.263 code.

7. References

Amer, A. and E. Dubois (2005). "Fast and Reliable structure-oriented Video Noise compression standards", *Proc. SPIE, Vol. CR60: Standards and Common Interfaces for estimation", IEEE Transactions on Circuits and systems for Video technology, Generic coding of* moving pictures and associated audio information: video, Laboratories, ISSN: 1051-8215 The Netherlands, May 1989, Video Inform. Syst., Philadelphia, USA, Oct. 1996,

Barsterretxea K, J. M. Tarela, Campo I.D., Digital design of sigmoid approximator for artifical neural networks, Electronics letters, Vol m38., *ISSN*:0013-5194 No.1. January 2002.

Boncelet C. (2005). *Handbook of Image and video procesing*, 2th edit, Elvesier Academic Press. *ISBN* 0121197921.

Bourlard, H., T. Adali, S. Bengio, J. Larsen, and S. Douglas, *Proceedings of the Twelfth IEEE Workshop on Neural Networks for Signal Processing*, ISSN:0018-9464 IEEE Press, 2002

Bronstein, I. N., K. A. Semeddjajew, G. Mosiol, and H. Muhlig (2005). *Taschenbuck der mathematik*, 6th edit, *ISBN*: 978-3-540-72121-5 Verlah Harri Deutch.

Cao J., Wang J., Liao X.F. Novel stability criteria of delayed celluar neural Networks 13 (2), *ISSN*: 0022-0000, 2002.

Cloete Ian, Jacek M. Zurada, "Knowledge-Based Neurocomputing", ISBN: 0-262-03274-0The MIT Press, 2000.

COST211bis/SIM89/37, *Description of Reference Model* 8 (*RM*8), PTT Research

Di Ferdinando, R. Calabretta and D. Parisi, "Evolving modular architectures for neural networks", in R. French.and J. P. Sougné (eds.) Connectionist models of learning, development and evolution, pp. 253-262, *ISSN*:1370-4621 Springer-Verlag: London, 2001.

Faulsett L., Fundamentals of neural networks—architectures, algorithms, and applications (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1994).

Fogel, D. B., C.J. Robinson, "Computational Intelligence", ISBN: 0-471-27454-2 John Wiley & Sons, IEEE Press,

Girod, B., E. Steinbach, and N. Färber, "Comparison of the H.263 and H.261 video

Hagan, H. Demuth, M. Beale, ISBN-10: 0971732108, ISBN-13: 978-0971732100"Neural Network Design", 2002,

Haykin, S. (1994). *Neural Networks*, *ISSN*:1069-2509New York, MacMillan Publ. Co.

Hertz, J., A. Krogh, R. G. Palmer, Introduction to the Theory of Neural Computation, *ISBN* 0-201-51560-1 Addison-Wesley, 1991.

ITU-T Recommendation (1995): H.262, *ISO/IEC 13818-2:1995, Information technology*

ITU-T Recommendation (1996): H.223, *Video coding for low bit rate communication.*

Kukolj, D. and M. Petrov (2000). "Unlabeled data clustering using a heuristic self-organizing neural network", *ISSN*: 1045-9227 *IEEE Transactions on Neural Networks*, 2000.

Kukolj, Dragan, Branislav Atlagić, Milovan Petrov, Unlabeled data clustering using a re-organizing neural network, Cybernetics and Systems, An Int. Journal, Vol. 37, No. 7, 2006, pp. 779-790.

LeGall, D.J., "The MPEG video compression algorithm", ISSN:*1110-8657 Signal Processing: Image Communication*, Vol. 4, No. 2, pp. 129-140, April 1992.

Lippmann, R. P. (1987). "An Introduction to Computing with Neural Nets", ISSN:0164-1212 *EEE ASSP Magazine*, April 1987, pp. 4-22.

Mandic D., Chambers J., Recurrent Neural Networks for prediction: Learning Algorithms, Architecture and stability, ISSN:*0045-7906* John Wiley & Sons, New York, 2002,

Markoski, B. and Đ. Babić (2007). "Polynominal-based filters in Bandpass Interpolation and Sampling rate conversion", *ISSN*: 1790-5022 *WSEAS Transactions on signal procesing*.

Nauck, D. and R. Kruse. Designing Neuro-Fuzzy Systems Through Backpropagation. W. Pedrycz ed., Fuzzy Modelling: ISSN:*01278274* Paradigms and Practice. Kluwer, Amsterdam, Netherlands 1995.

Nauck, D., C. Borgelt, F. Klawonn, R. Kruse, *ISBN*:0-89791-658-1 *Neuro – Fuzzy – Systeme,* Wiesbaden, 2003.

Nürnberger, A. Radetzky und R. Kruse. A problem specific recurrent neural network for the description and simulation of dynamic spring models. ISBN *0-7803-4859-1* In Proc. IEEE International Joint Conference on Neural Networks 1998 (IJCNN '98), 572-576. Anchorage, Alaska, Mai 1998.

Rijkse, K., "ITU standardisation of very low bitrate video coding algorithms", ISSN:*0923-5965 Signal Processing: Image Communication*, Vol. 7, pp. 553-565, 1995

Roese, J.A., W.K. Pratt, G.S. Robinson, "Interframe cosine transform image coding", *ISSN*:*0-387-08391-X IEEE Trans. Commun.*, Vol. 25, No. 11, pp. 1329-1339, Nov. 1977.

Schäfer, R., and T. Sikora, "Digital video coding standards and their role in video communications", ISSN*: 1445-1336Proc. IEEE*, Vol. 83, No. 6, pp. 907-924, June 1995.

Teeuwsen, S. P., I. Erlich, & M.A. El-Sharkawi, Neural network based classification method for small-signal stability assessment, ISSN: 0885-8950 Proc. IEEE Power Technology Conf.,Bologna, Italy, 2003, 1–6.

Teeuwsen, S. P., I. Erlich, & M.A. El-Sharkawi, Small-signal stability assessment based on advanced neural network methods, ISSN: 0885-8950 Proc. IEEE PES Meeting, Toronto, CA, 2003, 1–8.

Zurada, J. M. (1992). *Introduction to Artificial Neural Systems*, St. Paul, ISBN:*0-13-611435-0*. West Publishing Co.

Single Photon Eigen-Problem
with Complex Internal Dynamics

Nenad V. Delić[1], Jovan P. Šetrajčić[1,8], Dragoljub Lj. Mirjanić[2,8],
Zdravko Ivanković[3], Dobrivoje Martinov[4], Snežana Jokić[4],
Ivana Petrevska–Đukić[5], Dušanka Tešanović[6] and Svetlana Pelemiš[7]

[1] *Department of Physics, Faculty of Sciences, University of Novi Sad,*
[2] *Faculty of Medicine, University of Banja Luka,*
[3] *Faculty of Technical Sciences, University of Novi Sad,*
[4] *Technical Faculty Zrenjanin, University of Novi Sad,*
[5] *UniCredit Bank Srbija, a.d. Novi Sad,*
[6] *Oncology Institute of Vojvodina, Sremska Kamenica,*
[7] *Faculty of Technology Zvornik, University of East Sarajevo,*
[8] *Academy of Sciences and Arts in Banja Luka,*
[1,3,4,5,6] *Vojvodina – Serbia*
[2,7,8] *Republic of Srpska, BiH*

1. Introduction

Linearized single photon Hamiltonian is used for the analysis of its features in coordinate systems of various geometries. As it could have been expected, based on the general theory of relativity, it turned out that space geometry and physical features are closely interrelated. In Cartesian's coordinates single photons are spatial plane waves, while in cylindrical coordinates they are one-dimensional plane waves the amplitudes of which falls in planes normal to the direction of propagation. The most general information on single photon characteristics has been obtained by the analysis in spherical coordinates. The analysis in this system has shown that single photon spin essentially influences its behavior and that the wave functions of single photon can be normalized for zero orbital momentum, only.

A free photon Hamiltonian is linearized in the second part of this paper using Pauli's matrices. Based on the correspondence of Pauli's matrices kinematics and the kinematics of spin operators, it has been proved that a free photon integral of motion is a sum of orbital momentum and spin momentum for a half one spin. Linearized Hamiltonian represents a bilinear form of products of spin and momentum operators. Unitary transformation of this form results in an equivalent Hamiltonian, which has been analyzed by the method of Green's functions. The evaluated Green's function has given possibility for interpretation of photon reflection as a transformation of photon to anti-photon with energy change equal to double energy of photon and for spin change equal to Dirac's constant. Since photon is relativistic quantum object the exact determining of its characteristics is impossible. It is the reason for series of experimental works in which photon orbital momentum, which is not

integral of motion, was investigated. The exposed theory was compared to the mentioned experiments and in some elements the satisfactory agreement was found.

2. Eigen-problem of single photon Hamiltonian

In the first part of this work the eigen-problem of single photon Hamiltonian was formulated and solutions were proposed. Based on the general theory of relativity, it turned out that space geometry and physical features are closely interrelated. Because of that the analyses was provided in Cartesian's, cylindrical and spherical coordinate systems.

2.1 Introduction

Classical expression for free photon energy is:

$$E = c\sqrt{p_x^2 + p_y^2 + p_z^2} \,, \tag{1.1}$$

where c is the light velocity in vacuum and p_x, p_y and p_z are the components of photon momentum. If instead of classical momentum components we use quantum-mechanical operators $p_\nu \rightarrow \hat{p}_\nu = -i\hbar \dfrac{\partial}{\partial x_\nu}$; $\nu = (x,y,z)$, where $\hbar \equiv \dfrac{h}{2\pi} = 1{,}05456 \cdot 10^{-34}$ Js is Dirac's constant, we obtain quantum-mechanical single photon Hamiltonian:

$$\hat{H} = \pm c\sqrt{\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2} \,. \tag{2.2}$$

This Hamiltonian is not a linear operator that contradicts the principle of superposition (Gottifried, 2003; Kadin, 2005). Klein and Gordon (Sapaznjikov, 1983) skirted this problem solving the eigen-problem of square of Hamiltonian (2.2):

$$\hat{H}^2 \varphi = E^2 \varphi \,, \tag{2.3}$$

since the square of Hamiltonian is a linear operator. This approach has given satisfactory description of single photon behaving. Up to now it is considered that this approach gives real picture of photon. Here will be demonstrated that Kline–Gordon picture of photon is incomplete.

Here we shall try to examine single photon behavior by means of linearized Hamiltonian (2.2). Linearization procedure is analogous to the procedure that was used by Dirac's in the analysis of relativistic electron Hamiltonian (Dirac, 1958). We shall take that

$$\hat{p}_x^2 + \hat{p}_y^2 + \hat{p}_z^2 = \left(\hat{\alpha}\,\hat{p}_x + \hat{\beta}\,\hat{p}_y + \hat{\chi}\,\hat{p}_z \right)^2 \,, \tag{2.4}$$

i.e. we shall transform the sum of squares into the square of the sum using $\hat{\alpha}, \hat{\beta}$ and $\hat{\chi}$ matrices. In accordance with (2.4) these matrices fulfill the following relations:

$$\begin{aligned}
&\hat{\alpha}^2 = \hat{\beta}^2 = \hat{\chi}^2 = 1; \\
&\hat{\alpha}\,\hat{\beta} + \hat{\beta}\,\hat{\alpha} = \hat{\alpha}\,\hat{\chi} + \hat{\chi}\,\hat{\alpha} = \hat{\beta}\,\hat{\chi} + \hat{\chi}\,\hat{\beta} = 0.
\end{aligned} \tag{2.5}$$

It is easy to show (Tošić, et al., 2008; Delić, et al., 2008) that (2.5) conditions are fulfilled by Pauli's matrices

$$\hat{\alpha} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \qquad \hat{\beta} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \qquad \hat{\chi} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.6}$$

Combining (2.6), (2.4) and (2.2), we obtain linearized photon Hamiltonian which completely reproduces the quantum nature of light (Holbrow, et al., 2001; Torn, et al., 2004) in the form:

$$\hat{H} = \pm c \begin{pmatrix} \hat{p}_z & \hat{p}_x - i\hat{p}_y \\ \hat{p}_x + i\hat{p}_y & -\hat{p}_z \end{pmatrix} = \pm \frac{\hbar c}{i} \begin{pmatrix} \dfrac{\partial}{\partial z} & \dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y} & -\dfrac{\partial}{\partial z} \end{pmatrix}. \tag{2.7}$$

Since linearized Hamiltonian is a 2×2 matrix, photon eigen-states must be columns and rows which two components. Operators of other physical quantities must be represented in the form of diagonal 2×2 matrices.

At the end of this presentation, it is important to underline the orbital momentum operator $\begin{pmatrix} \hat{\vec{L}} & 0 \\ 0 & \hat{\vec{L}} \end{pmatrix}$; $\hat{\vec{L}} = \hat{\vec{r}} \times \hat{\vec{p}}$ does not commute with Hamiltonian (2.7). It means that it is not integral of motion as in Klein-Gordon theory (Davidov, 1963). It can be shown that integral of motion represents total momentum $\begin{pmatrix} \hat{\vec{J}} & 0 \\ 0 & \hat{\vec{J}} \end{pmatrix}$, where $\hat{\vec{J}}$ is the sum of orbital momentum $\hat{\vec{L}}$ and rotation momentum $\hat{\vec{S}}$ which corresponds to 1/2 spin.

In further the eigen-problem of linearized single photon Hamiltonian will be analyzed in Cartesian's, cylindrical and spherical coordinates.

2.2 Photons in Cartesian's picture

The eigen-problem of single photon Hamiltonian in Cartesian coordinates (we shall take it with plus sign) has the following form:

$$\frac{\hbar c}{i} \begin{pmatrix} \dfrac{\partial}{\partial z} & \dfrac{\partial}{\partial x} - i\dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial x} + i\dfrac{\partial}{\partial y} & -\dfrac{\partial}{\partial z} \end{pmatrix} \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = E \begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}, \tag{2.8}$$

wherefrom we obtain the following system of equations from:

$$\left(\frac{\partial}{\partial z} - ik \right)\Psi_1 + \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y} \right)\Psi_2 = 0 \, ; \tag{2.9a}$$

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y} \right)\Psi_1 - \left(\frac{\partial}{\partial z} + ik \right)\Psi_2 = 0 \, , \tag{2.9b}$$

where $k = \dfrac{E}{\hbar c}$. It follows from (2.9a) that:

$$\Psi_1 = \left(\frac{\partial}{\partial z} - ik \right)^{-1} \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y} \right)\Psi_2 . \tag{2.10}$$

Since the operators $\dfrac{\partial}{\partial z} \pm ik$ and $\dfrac{\partial}{\partial x} \pm i\dfrac{\partial}{\partial y}$ commute, through (2.10) we come to the following

relation:

$$\left(\frac{\partial}{\partial z} + ik\right)\left(\frac{\partial}{\partial z} - ik\right)\Psi_1 + \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\Psi_1 = 0 \cdot \tag{2.11}$$

In the same manner, from (2.9b) and (2.10), we come to the relation:

$$\left(\frac{\partial}{\partial z} - ik\right)\left(\frac{\partial}{\partial z} + ik\right)\Psi_2 + \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\Psi_2 = 0 \cdot \tag{2.12}$$

The two last relations are of identical form and can be substituted by one unique relation:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} + k^2\right)\Psi(x, y, z) = 0 \cdot \tag{2.13}$$

If we take in (2.13) that $k^2 = k_x^2 + k_y^2 + k_z^2$ and $\Psi(x, y, z) = A(x)B(y)C(z)$, we come to the following equation:

$$\frac{1}{A}\frac{d^2 A}{dx^2} + k_x^2 + \frac{1}{B}\frac{d^2 B}{dy^2} + k_y^2 + \frac{1}{C}\frac{d^2 C}{dz^2} + k_z^2 = 0, \tag{2.14}$$

which is fulfilled if:

$$\frac{d^2 A}{dx^2} + k_x^2 A = 0; \quad \frac{d^2 B}{dy^2} + k_y^2 B = 0; \quad \frac{d^2 C}{dz^2} + k_z^2 C = 0. \tag{2.15}$$

Equations (2.15) can be easily solved and each of them has two linearly independent particular integrals:

$$\begin{aligned} A_1 &= a_1 e^{ixk_x}; \quad A_2 = a_2 e^{-ixk_x}; \\ B_1 &= b_1 e^{iyk_y}; \quad B_2 = b_2 e^{-iyk_y}; \\ C_1 &= c_1 e^{izk_z}; \quad C_2 = c_2 e^{-izk_z}. \end{aligned} \tag{2.16}$$

Based on these expressions, we conclude that eigen-vector of single photon $\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix}$ has the

following form:

$$\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = \begin{pmatrix} D\,e^{i\vec{k}\vec{r}} \\ D\,e^{-i\vec{k}\vec{r}} \end{pmatrix}. \tag{2.17}$$

Since $\vec{k}$ is a continuous variable, the normalization of (2.17) mast be done to δ–function, wherefrom follows:

$$D^2 \int d^3\vec{r}\left(e^{-i\vec{k}\vec{r}}\; e^{i\vec{k}\vec{r}}\right)\begin{pmatrix} e^{i\vec{k}\vec{r}} \\ e^{-i\vec{k}\vec{r}} \end{pmatrix} = \delta(\vec{k} - \vec{k}'). \tag{2.18}$$

Solving these integrals, we come to: $2\,D^2\,(2\pi)^3 = 1$, wherefrom we get the normalized single photon eigen-vector as:

$$\begin{pmatrix} \Psi_1 \\ \Psi_2 \end{pmatrix} = \frac{1}{4\sqrt{\pi^3}} \begin{pmatrix} e^{i\vec{k}\vec{r}} \\ e^{-i\vec{k}\vec{r}} \end{pmatrix}. \tag{2.19}$$

As it can be seen from (2.19), the components of single photon eigen-vector are progressive plane wave $\sim e^{i\vec{k}\vec{r}}$ and the regressive one $\sim e^{-i\vec{k}\vec{r}}$. Since we consider a free single photon, the obtained conclusion is physically acceptable.

2.3 Photons in cylindrical picture

In this section of first part of the paper we are going to analyze the same problem in cylindrical coordinates. Since solving of partial equation of $(\Delta + k^2)\Psi = 0$ type in cylindrical coordinates requires more general approach than that which was used in Cartesian's coordinates, it is necessary to examine single photon eigen-problem in cylindrical system.
 In order to examine this problem, we shall start from the equation (2.13) in which Laplacian $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \equiv \Delta$ will be given in cylindrical coordinates (ρ,φ,z) where $\rho \ni [0,\infty]$, $\varphi \ni [0,2\pi]$ and $z \ni [-\infty,+\infty]$. The Laplacian in cylindrical coordinates has the following form:

$$\Delta = \frac{\partial^2}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2}{\partial \varphi^2} + \frac{\partial^2}{\partial z^2}$$

and therefore (2.13) with $\Psi(x,y,z) \Rightarrow \Phi(\rho,\varphi,z)$, reduces to:

$$\frac{\partial^2\Phi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial\Phi}{\partial \rho} + \frac{1}{\rho^2}\frac{\partial^2\Phi}{\partial \varphi^2} + \frac{\partial^2\Phi}{\partial z^2} + k^2\Phi = 0. \tag{2.20}$$

The square of wave vector k will be separated into two parts $k_x^2 + k_y^2 + k_z^2 = q^2 + k_z^2$. On the basis of this the equation (2.20) can be written as follows:

$$\frac{\partial^2\Phi}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial\Phi}{\partial \rho} + q^2\Phi + \frac{1}{\rho^2}\frac{\partial^2\Phi}{\partial \varphi^2} = -\frac{\partial^2\Phi}{\partial z^2} - k_z^2\Phi. \tag{2.21}$$

By the substitution:

$$\Phi(\rho,\varphi,z) = F(\rho,\varphi)\,G(z), \tag{2.22}$$

the equation (2.21) reduces to:

$$\frac{1}{F}\left(\frac{\partial^2 F}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial F}{\partial \rho} + q^2 F + \frac{1}{\rho^2}\frac{\partial^2 F}{\partial \varphi^2} \right) = -\frac{1}{G}\left(\frac{d^2 G}{\partial z^2} + k_z^2 G \right). \tag{2.23}$$

This equation is fulfilled if:

$$\frac{\partial^2 F}{\partial \rho^2} + \frac{1}{\rho}\frac{\partial F}{\partial \rho} + q^2 F + \frac{1}{\rho^2}\frac{\partial^2 F}{\partial \varphi^2} = 0; \tag{2.24a}$$

$$\frac{d^2 G}{\partial z^2} + k_z^2 G = 0.$$ (2.24b)

Now we separate the variables by substitution:

$$F(\rho, \varphi) = X(\rho) S(\varphi),$$ (2.25)

after which, the (2.24a) goes over to:

$$\frac{1}{X}\left(\rho^2 \frac{\partial^2 X}{\partial \rho^2} + \rho \frac{\partial X}{\partial \rho} + q^2 \rho^2 X\right) = -\frac{1}{S}\frac{d^2 S}{\partial \varphi^2} \equiv m^2.$$ (2.26)

Introduction of the variables separation constant m^2 represents generalization with respect to approach used in previous section. Since the function $S(\varphi)$ must be single-sign $S(\varphi) = S(\varphi + 2\pi)$ we must that m is integer, i.e. $m = 0, \pm 1, \pm 2, ...$
Relation (2.26) is separated into two differential equations:

$$\frac{d^2 S}{d\varphi^2} + m^2 S = 0;$$ (2.27a)

$$\frac{d^2 X}{d\rho^2} + \frac{1}{\rho}\frac{dX}{d\rho} + (q^2 - \frac{m^2}{\rho^2})X = 0.$$ (2.27b)

The equation (2.24b) has two particular integrals:

$$G_1 = g_1 e^{izk_z}; \quad G_2 = g_2 e^{-izk_z},$$ (2.28)

while the solution of the equation (2.27a) is:

$$S_m(\varphi) = s_0 e^{im\varphi}.$$ (2.29)

By the substitution of argument $\rho = b\xi$, the equation (2.27b) reduces to

$$\frac{d^2 X}{d\xi^2} + \frac{1}{\xi}\frac{dX}{d\xi} + \left(q^2 b^2 - \frac{m^2}{\xi^2}\right)X = 0,$$ (2.30)

and taking that $b = \frac{1}{q}$, we translate (2.30) into Bessel's equation with integer index m:

$$\frac{d^2 X}{d\xi^2} + \frac{1}{\xi}\frac{dX}{d\xi} + \left(1 - \frac{m^2}{\xi^2}\right)X = 0.$$ (2.31)

It means that the solution of (2.27b) is the m–order Bessel's function: J_m, i.e.

$$X(\rho) = a_0 J_m(q\rho).$$ (2.32)

Taking into account (2.28), (2.29) and (2.32), we obtain the components of single photon eigen-vector:

$$\Phi_1(\rho, \varphi, z) = D_1 J_m(q\rho)e^{izk_z}e^{im\varphi}; \quad \Phi_2(\rho, \varphi, z) = D_2 J_m(q\rho)e^{-izk_z}e^{im\varphi}.$$ (2.33a)

Since q and k_z are continuous variables, while m is a discrete one the normalization of eigen-vector must be done partially to δ–functions and partially to Kronecker's symbol. It means that normalization condition is the following:

$$\left(\left|D_1\right|^2 + \left|D_2\right|^2\right)\int_0^{2\pi}d\varphi\, e^{\pm i(m-m')\varphi}\int_{-\infty}^{+\infty}dz\, e^{\pm i(k_z-k_z')z_z}\int_0^{\infty}\rho\,d\rho\, J_m(q'\rho)J_m(q\rho) = q^{-1}\delta_{nm}\delta(k_z - k_z')\delta(q - q') .$$

Using formula for normalization of Bessel functions with integer index (Korn & Korn, 1961):

$$\int_0^{\infty}dx\,x\,J_m(k'x)J_m(kx) = \frac{1}{k}\delta(k - k') ,$$

the normalization condition reduces into: $\left|D_1\right|^2 + \left|D_2\right|^2 = \frac{1}{4\pi^2}$. It means that normalized single photon eigen-vector in cylindrical coordinates is given by:

$$\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} = \begin{pmatrix} D_1 J_m(q\rho)e^{im\rho}e^{izk_z} \\ D_2 J_m(q\rho)e^{im\rho}e^{-izk_z} \end{pmatrix}. \tag{2.34}$$

The first component Φ_1 corresponds to photon (velocity $+c$), while second component Φ_2 corresponds to anti-photon (velocity $-c$). From this formula we conclude that single photon eigen-vector components are progressive and regressive plane waves along z-axis. In the (x,y) planes components change periodically with polar angle φ and decrease by the rule $\rho^{-1/2}$ with distance between the axis and envelope of cylinder. The last is concluded on the basis of asymptotic behaving of Bessel's functions (Korn & Korn, 1961): $J_m(\rho) \approx \dfrac{\sin\rho}{\sqrt{\rho}}$, when $\rho \to \infty$. We have seen during the analysis of a photon in Cartesian's coordinates that only zero values of parameters of variables separation are physically imposed. In cylindrical coordinates, due to physical reasons again, one parameter of variable separation had zero value, while the other has to be a square of integer. The last is necessary since the solution must be single-sign function.

2.4 Photon in spherical picture

The analysis of single photon eigen-problem in spherical coordinates, as it well be shown later, requires introduction of two variable separation parameters. We start from the equation (2.13), where the Laplace's operator will be written down in spherical coordinates (r,θ,φ), where $r \in [0,\infty]$, $\theta \in [0,\pi]$ and $\varphi \in [0,2\pi]$. In these coordinates it is of the form:

$$\Delta = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} . \tag{2.35}$$

It means that (2.13), with $\Psi(x,y,z) \to \Omega(r,\theta,\varphi)$, becomes:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\Omega}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Omega}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\Omega}{\partial\varphi^2} + k^2\Omega = 0 . \tag{2.36}$$

In the first stage of variables separation, we shall take that:

$$\Omega(r,\theta,\varphi) = R(r)Q(\theta,\varphi) \,, \tag{2.37}$$

after which substitution into (2.36), it goes over to:

$$\frac{1}{R}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) + k^2 r^2 R\right] = -\frac{1}{Q}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Q}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Q}{\partial\varphi^2}\right] = \Lambda^2 \,, \tag{2.38}$$

where Λ^2 is the variable separation parameter. Double equality in (2.38) gives two equations:

$$\frac{d^2 R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left(k^2 - \frac{\Lambda^2}{r^2}\right)R = 0 \,; \tag{2.39a}$$

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Q}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Q}{\partial\varphi^2} + \Lambda^2 Q = 0 \,. \tag{2.39b}$$

It should be noted that equation (2.39b) represents eigen-problem of $\dfrac{\hat{L}^2}{\hbar^2}$ operator. It means that Λ^2 determines orbital quantum numbers. In this equation we shall take that:

$$Q(\theta,\varphi) = T(\theta)S(\varphi) \,, \tag{2.40}$$

after this substitution, which goes over to:

$$\frac{1}{B}\left[\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial T}{\partial\theta}\right) + T\,\Lambda^2\sin^2\theta\right] = -\frac{1}{S}\frac{\partial^2 S}{\partial\varphi^2} = m^2 \,. \tag{2.41}$$

In this double equality the variable separation parameter m must be integer since the solution $S(\varphi)$ must be single-signed function. The same requirement appeared in the previous section where single photon vas analyzed in cylindrical coordinates. The equation (2.41) gives two second order differential equations:

$$\frac{d^2 S}{d\varphi^2} + m^2 S = 0 \,; \tag{2.42a}$$

$$\frac{d^2 T}{d\theta^2} + \cot\theta\frac{dT}{d\theta} + \left(\Lambda^2 - \frac{m^2}{\sin^2\theta}\right)T = 0 \,. \tag{2.42b}$$

When the solution (2.42a) is:

$$S_m(\varphi) = s_0 e^{im\varphi} \,; \quad m = 0,\pm1,\pm2, \dots, \tag{2.43}$$

the equation (2.42b) is associated Legendre's equation (Gottifried, 2003; Davidov, 1963). The complete procedure of solving of this equation cannot be found in literature. Instead of the general solving procedure of the equation (2.42b) is solved for $m = 0$. Its solutions are Legendre's polynomials (Korn & Korn, 1961; Janke, et al., 1960). Differentiating these polynomials m-th times it was possible to conclude that solution (2.42b) can be expressed through m-th Legendre's polynomials derivations.

In order to avoid such an artificial solving of the equation (2.42b), we shall expose, briefly, its solving by means of potential series. This solving procedure may be comprehended as methodological contribution of this part of the paper. At the first stage, we translate the equation (2.42b) into algebraic form by means of substitution of argument $\cos\theta = \zeta$:

$$\left(1-\zeta^2\right)\frac{d^2B}{d\zeta^2} - 2\zeta\frac{dB}{d\zeta} + \left(\zeta^2 - \frac{m^2}{1-\zeta^2}\right)B = 0; \quad \zeta\in[-1,+1]. \tag{2.44}$$

The term $\frac{m^2}{1-\zeta^2}$ in (2.44) does not allow the solving of this equation by means of potential

series. Consequently this term must be eliminated from the equation. The strategy of elimination is the following: by the substitution of $T = U\cdot V$, where U is an arbitrary function, the equation (2.44) reduces to the same form but with arbitrary constant in linear function with is multiplied by first derivative of V function. This arbitrary coefficient will be taken in the form $-2(2s+1)$ where s is arbitrary. Arbitrary constant s will be determined in a way which eliminates the term $\frac{m^2}{1-\zeta^2}$ from equation for V function. By the described strategy the

(2.44) becomes:

$$\left(1-\zeta^2\right)\frac{d^2V}{d\zeta^2} - 2(2s+1)\zeta\frac{dV}{d\zeta} + \left(\Lambda^2 - 2s - 4s^2\right)V = 0\cdot \tag{2.45}$$

This equation is suitable for solving by means of potential series. Arbitrary function U is given by $U = \left(1-\zeta^2\right)^s$, where $s = \pm\, m/2$. This means that function T has the form:

$$T = \left(1-\zeta^2\right)^S V. \tag{2.46}$$

Since $\zeta\in[-1,+1]$ the exponent s must not be negative since T would then have singularities in $\zeta = \pm 1$ not allowing the normalization. Fortunate circumstance is that the exponent of the function $1 - \zeta^2$ has $\pm$ sign. This means that for $m > 0$ can be taken $s = +\, m/2 = |m|/2$. If $m < 0$, we take $s = -\, m/2 = |m|/2$. Based on this reasoning the equation (2.45) becomes:

$$\left(1-\zeta^2\right)\frac{d^2V}{d\zeta^2} - 2(|m|+1)\zeta\frac{dV}{d\zeta} + \left[\Lambda^2 - |m|(|m|+1)\right]V = 0\cdot \tag{2.47}$$

The solution of this equation will looked for in the form of potential series:

$$V = \sum_{n=0}^{\infty} v_n\zeta^n, \tag{2.48}$$

after which substitution in (2.47) we obtain recurrent formula for series coefficients:

$$v_{n+2} = -\frac{\Lambda^2 - \left(n+|m|\right)\left(n+|m|+1\right)}{(n+1)(n+1)}v_n; \quad n = 0,1,2, \ldots \tag{2.49}$$

Here arises a dilemma whether to leave the whole series or to cut it and retain a polynomial instead of series. In order to solve this dilemma, we shall analyze a special case of formula (2.49) when $m = \Lambda = 0$. In this case formula (2.49) becomes:

$$v_{n+2} = \frac{n}{n+2} v_n; \quad n = 1,3,5, \ldots, \tag{2.50}$$

wherefrom it turns out that $v_n = \dfrac{v_1}{2n+1}$, and this means that series solution (2.48) becomes:

$$V = \zeta + \frac{\zeta^3}{3} + \frac{\zeta^5}{5} \equiv \int \frac{d\zeta}{1-\zeta^2} = \frac{1}{2} \ln \frac{\zeta-1}{\zeta+1}. \tag{2.51}$$

From this formula is obvious that the series has singularities for $\zeta = \pm 1$. This resolves above mentioned dilemma: the series must be cut and the polynomial obtained in this way must be taken as solution. From the formula (2.49) it is clear that the series will be cut if:

$$\Lambda^2 = l\,(l+1)\,; \quad l = 0,1,2, \ldots \tag{2.52}$$

Now is clear that the series is cut when $l = |m| + n$, wherefrom it follows that the degree of polynomial is $l = |m| - n$ and that quantum number m per module must not exceed l: $|m| \le l$. The obtained polynomials of $l - |m|$ degree are called the associated Legendre's polynomials (Korn & Korn, 1961; Janke, et al., 1960) and by means of them T function is expressed as:

$$T_{l,|m|}(\zeta) = (1 - \zeta^2)^{|m|/2} L_{l-|m|}(\zeta). \tag{2.53}$$

 The product of functions (2.43) and (2.53) normalized per angles gives spherical harmonics (Gottifried, 2003; Davidov, 1963):

$$Y_{l,|m|} = \frac{(-1)^{l+|m|}}{2^l\,l!} \frac{e^{im\varphi}}{\sqrt{2\pi}} \sqrt{\frac{2l+1}{2} \frac{(l-|m|)!}{(l+|m|)!}} \sin^{|m|} \theta \frac{d^{l+|m|}}{d(\cos\theta)^{l+|m|}} (\sin\theta)^{2l}. \tag{2.54}$$

Finally we shall solve the equation (2.39a) in which Λ^2 is substituted by $l\,(l+1)$. It means that it goes over to:

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left(k^2 - \frac{l(l+1)}{r^2}\right) R = 0\,; \quad r \in [0,\infty). \tag{2.55}$$

Substituting the function R with $r^{-1/2} J(r)$ and substituting the argument r by ρ/k, we translate last equation into Bessel's equation (Korn & Korn, 1961; Janke, et al., 1960) with $l+1/2$ index having two linearly independent particular solution $J_{l+1/2}(kr)$ and $J_{-l-1/2}(kr)$. Consequently the solutions of (2.39a) are:

$$R_1 = w_1(kr)^{-1/2} J_{l+1/2}(kr)\,; \ R_2 = w_2(kr)^{-1/2} J_{-l-1/2}(kr). \tag{2.56}$$

It is necessary for further to quote behaving of Bessel's functions with half integer indices. It can be easily shown that:

$$J_{1/2}(kr) = \frac{\sin kr}{\sqrt{kr}}\,; \quad J_{-1/2}(kr) = \frac{\cos kr}{\sqrt{kr}}, \tag{2.57}$$

As well as using recurrent formula for Bessel's functions (Janke, et al., 1960):

$$\frac{d}{dx}J_p = \frac{1}{2}J_{p-1} - \frac{1}{2}J_{p+1}, \tag{2.58}$$

and taking that $p = +1/2$ and $p = -1/2$, we obtain respectively:

$$J_{3/2}(x) = x^{-3/2}\sin x - x^{-1/2}\cos x \,; J_{-3/2}(x) = -x^{-3/2}\cos x - x^{-1/2}\sin x. \tag{2.59}$$

Due to the factor $x^{-3/2}$ functions $J_{\pm 3/2}$ have strong singularities in zero so that they cannot be normalized in the interval $0 \leq r < \infty$. Due to the same reasons neither $J_{\pm 5/2}$, $J_{\pm 7/2}$, etc. cannot be normalized. It can be concluded that only solutions for A which are proportional to $J_{\pm 1/2}$ have chances to be normalized. Those solutions are:

$$R_1 = \frac{W}{\sqrt{r}}\frac{\sin kr}{\sqrt{kr}} \,; \ R_2 = \frac{W}{\sqrt{r}}\frac{\cos kr}{\sqrt{kr}}. \tag{2.60}$$

The very important conclusion of this analysis is: only free photons with zero orbital momentum have chances to be normalized exist. For $l > 0$ photon eigen-vector cannot be normalized.

We shall now examine whether the components of photon eigen-vector proportional to R_1 and R_2 can be normalized. Those components are:

$$\Omega_1 = \frac{W}{\sqrt{r}}Y_{00}(\theta,\varphi)\frac{\sin kr}{\sqrt{kr}} \,; \Omega_2 = \frac{W}{\sqrt{r}}Y_{00}(\theta,\varphi)\frac{\cos kr}{\sqrt{kr}}. \tag{2.61}$$

The normalization condition is the following:

$$W^2\int_0^{2\pi}d\varphi\int_0^{\pi}d\theta\,\sin\theta\left|Y_{0,0}(\theta,\varphi)\right|^2\int_0^{\infty}dr\,r^2\left[J_{1/2}(k'r)J_{1/2}(kr) + J_{-1/2}(k'r)J_{-1/2}(kr)\right] = $$

$$= \frac{W^2}{\sqrt{k'k}}\int_0^{\infty}dr\cos(k-k')r = \frac{1}{k^2}\delta(k-k'). \tag{2.62}$$

It is not difficult to show that: $\int_0^{\infty}dr\cos(k-k')r = 0$, so that the condition (2.62) becomes meaningless. This means that even for $l = 0$ photon eigen-vector cannot be normalized.

The last possibility for normalization free photons eigen-vector is so called box quantization method. In this method the sphere is substituted by cube enveloping it and cyclic boundary conditions are required: $e^{ikr} = e^{ik(r+L)}$, wherefrom it follows that wave vector is quantized:

$$k = \frac{2\pi}{L}n \,; \quad n = 1,2,3, \dots \tag{2.63}$$

Since $k = 2\pi/\lambda$, it gives that:

$$L = n\lambda \,; \quad n = 1,2,3, \dots \tag{2.64}$$

It is seen that the first harmonic of electromagnetic waves has the wave length equal to the cube edge.

Photon energy is determined in the standard way:

$$E = \hbar ck = \frac{h}{2\pi}c\frac{2\pi}{L}n = h\nu_0 n \,; \quad \nu_0 = \frac{c}{L}. \tag{2.65}$$

This expression for energy is in full accordance with Plank's hypothesis (Planck, 1901). In the normalization condition (2.62) the following translations has to be used:

$$\delta(k-k') \to \delta_{nn} \xrightarrow[n=m]{} 1 \; ; \quad \int_0^\infty dr \to \int_0^L dr = L \; ; \quad \frac{\cos(k-k')r}{\sqrt{kk'}} \to \frac{\cos\dfrac{2\pi}{L}(n-m)r}{\dfrac{2\pi}{L}\sqrt{nm}} \xrightarrow[n=m]{} \frac{L}{2\pi n} \; .$$

Combining this and (2.62) we obtain that the normalization constant is $W = \dfrac{1}{\sqrt{2\pi n}}$. On the basis of this the normalized photon eigen-vector is given by:

$$\begin{pmatrix} \Omega_1 \\ \Omega_2 \end{pmatrix} = \frac{1}{\sqrt{2\pi n}} \begin{pmatrix} Y_{00}(\theta,\varphi)\, r^{-1/2}\, J_{1/2}(\dfrac{2\pi}{L}nr) \\ Y_{00}(\theta,\varphi)\, r^{-1/2}\, J_{-1/2}(\dfrac{2\pi}{L}nr) \end{pmatrix} = \frac{n}{2}\frac{1}{(2\pi n)^{3/2}} \begin{pmatrix} \dfrac{\sin\dfrac{2\pi}{L}nr}{\sqrt{\dfrac{2\pi}{L}nr}} \\[4ex] \dfrac{\cos\dfrac{2\pi}{L}nr}{\sqrt{\dfrac{2\pi}{L}nr}} \end{pmatrix}; \quad n = 1,2,3,\ldots \qquad (2.66)$$

As it can be seen the analysis of single photon eigen-problem in spherical coordinates has shown it orbital momentum of photon is equal to zero and that the spin $S = 1/2$ is its unique rotational characteristics (Yao, et al., 2005). Physically it is fully understandable that orbital momentum of a free photon is equal to zero since it moves along the straight line. On straight line photon radius-vector $\vec{r}$ and its momentum $\vec{p} = m_f \dot{\vec{r}}$ are parallel and this gives that $\vec{l} = \vec{r} \times \vec{p} = 0$.

3. Free photon as a system with complex internal dynamics

In the second part of this work the free photon Hamiltonian will be linearized using Pauli's matrices. Based on the correspondence of Pauli matrices kinematics and the kinematics of spin operators, the unitary transformation of this form (equivalent Hamiltonian), will be analyzed by the method of Green's functions. Since photon is relativistic quantum object the exact determining of its characteristics is impossible. It is the reason for series of experimental works in which photon orbital momentum, which is not integral of motion, will be theoretically investigated.

3.1 Introduction

The fact that photon Hamiltonian is not a linear operator has a set of consequences that have not been studied sufficiently so far. The main reason is that photon characteristics have been mainly examined by means of Klein-Gordon's equation (Gottifried, 2003; Davidov, 1963; Messiah, 1970; Davydov, 1976), which represents eigen-problem of photon Hamiltonian square. In this part of our paper we shall linearized photon Hamiltonian and examine some of photon characteristics witch follow from linearized Hamiltonian. The analogy with Dirac's approach to the problem of electrons will be used (Gottifried, 2003; Dirac, 1958).

Firstly will be examined integrals of motion of free photon and will be shown that the photon integral of motion is not orbital momentum. It will be shown that the integral of motion is total momentum being the sun of orbital one and spin momentum.

The evaluated Green's function has given possibility for interpretation of photon reflection as a transformation of photon to anti-photon with energy change equal to double energy of photon and for spin change equal to Dirac's constant (Dirac, 1958; Messiah, 1970). Since photon is relativistic quantum object the exact determining of its characteristics is impossible.

The discussion of obtained results and their comparison to the experimental data will be done at the last part.

3.2 Linearized photon Hamiltonian

We shall not deal with this eigen-problem in further of this paper. Instead of this we shall look for integrals of motion, i.e. those operators that commute with free-photon Hamiltonian (2.7). It is obvious that any function depending on momentum components represents an integral of motion, but this fact is not of physical interest.

It is of particular importance whether orbital momentum:

$$\hat{\vec{L}} = \begin{pmatrix} \hat{\vec{L}} & 0 \\ 0 & \hat{\vec{L}} \end{pmatrix}; \quad \hat{\vec{L}} = \vec{r} \times \hat{\vec{p}} \tag{3.1}$$

is photon integral of motion, since in non-relativistic quantum mechanics operator $\hat{\vec{L}}$ is integral of motion of electron (Messiah, 1970; Davydov, 1976). The components of orbital momentum are given as follows:

$$\hat{L}_x = y\hat{p}_z - z\hat{p}_y ; \quad \hat{L}_y = z\hat{p}_x - x\hat{p}_z ; \quad \hat{L}_z = x\hat{p}_y - y\hat{p}_x . \tag{3.2}$$

If we use commutation relations for components of radius vector and the components of momentum: $[x_i, p_j] = i\hbar \delta_{ij}, \quad i,j \in (x,y,z)$ and look for commutators of (3.2) with Hamiltonian (2.7), we come to the following relations:

$$\left[\hat{L}_x, \hat{H}\right] = \pm i\hbar c\left(\hat{p}_z\hat{\beta} - \hat{p}_y\hat{\chi}\right); \quad \left[\hat{L}_y, \hat{H}\right] = \pm i\hbar c\left(\hat{p}_x\hat{\chi} - \hat{p}_z\hat{\alpha}\right); \quad \left[\hat{L}_z, \hat{H}\right] = \pm i\hbar c\left(\hat{p}_y\hat{\alpha} - \hat{p}_x\hat{\beta}\right), \tag{3.3}$$

based on which it follows that orbital momentum is not a free photon integral of motion.

It should be pointed out that signs in (3.3) are obtained on the basis of obvious symmetry properties $\hat{H}(-\vec{r}) = \hat{H}(\vec{r})$ and $\vec{L}(-\vec{r}) = \vec{L}(\vec{r})$, where $\vec{r}$ is radius-vector.

In order to find some rotation characteristics that commute with a free photon Hamiltonian, we shall first show that commutation relations for matrices $\hat{\alpha}, \hat{\beta}$ and $\hat{\chi}$, given in section 2.1 by expression (2.6), are:

$$\left[\hat{\alpha}, \hat{\beta}\right] = 2i\hat{\chi}; \quad \left[\hat{\chi}, \hat{\alpha}\right] = 2i\hat{\beta}; \quad \left[\hat{\beta}, \hat{\chi}\right] = 2i\hat{\alpha}, \tag{3.4}$$

while commutation relations for spin components (Dirac, 1958; Messiah, 1970):

$$\left[\hat{S}_x, \hat{S}_y\right] = i\hbar\hat{S}_z ; \quad \left[\hat{S}_z, \hat{S}_x\right] = i\hbar\hat{S}_y ; \quad \left[\hat{S}_y, \hat{S}_z\right] = i\hbar\hat{S}_x , \tag{3.5}$$

are very similar to (3.4). Comparing (3.4) to (3.5) we can establish the correspondence between spin operator components and matrices $\hat{\alpha}, \hat{\beta}$ and $\hat{\chi}$:

$$\hat{S}_x = \frac{\hbar}{2} \hat{\alpha} \; ; \quad \hat{S}_y = \frac{\hbar}{2} \hat{\beta} \; ; \quad \hat{S}_z = \frac{\hbar}{2} \hat{\chi} \; . \tag{3.6}$$

Commutators of matrices $\hat{\alpha}, \hat{\beta}$ and $\hat{\chi}$ with Hamiltonian are given by:

$$\left[\hat{\alpha}, \hat{H} \right] = \mp 2ic(\hat{p}_z \hat{\beta} - \hat{p}_y \hat{\chi}) \; ; \quad \left[\hat{\beta}, \hat{H} \right] = \mp 2ic(\hat{p}_x \hat{\chi} - \hat{p}_z \hat{\alpha}) \; ; \quad \left[\hat{\chi}, \hat{H} \right] = \mp 2ic(\hat{p}_y \hat{\alpha} - \hat{p}_x \hat{\beta}) \; . \tag{3.7}$$

We shall now look for a commutator of component $\hat{J}_x$ of total momentum, with photon Hamiltonian i.e. with $\hat{H}(\vec{r})$. Using upper signs in formulas (3.3) and (3.7) we obtain:

$$\left[\hat{J}_x, \hat{H}(\vec{r}) \right] = \left[(\hat{L}_x + \hat{S}_x), \hat{H}(\vec{r}) \right] = \left[(\hat{L}_x + \frac{\hbar}{2} \hat{\alpha}), \hat{H}(\vec{r}) \right] = \left[\hat{L}_x, \hat{H}(\vec{r}) \right] + \frac{\hbar}{2} \left[\hat{\alpha}, \hat{H}(\vec{r}) \right] =$$
$$= i\hbar c \left(\hat{p}_z \hat{\beta} - \hat{p}_y \hat{\chi} \right) + \frac{\hbar}{2} (-2ic) \left(\hat{p}_z \hat{\beta} - \hat{p}_y \hat{\chi} \right) = 0 \; . \tag{3.8}$$

For lower signs in formulas (3.3) and (3.7) [1], we have:

$$\left[\hat{J}_x, \hat{H}(-\vec{r}) \right] = \left[(\hat{L}_x + \hat{S}_x), \hat{H}(-\vec{r}) \right] = \left[(\hat{L}_x + \frac{\hbar}{2} \hat{\alpha}), \hat{H}(-\vec{r}) \right] = \left[\hat{L}_x, \hat{H}(-\vec{r}) \right] + \frac{\hbar}{2} \left[\hat{\alpha}, \hat{H}(-\vec{r}) \right] =$$
$$= i\hbar c \left(\hat{p}_z \hat{\beta} - \hat{p}_y \hat{\chi} \right) + \frac{\hbar}{2} (2ic) \left(\hat{p}_z \hat{\beta} - \hat{p}_y \hat{\chi} \right) = 0 \; . \tag{3.9}$$

It can be proved, in the same manner, that both y and z components of total momentum $\vec{J} = \vec{L} + \vec{S}$ commute with photon Hamiltonian (the expression (2.7) with sign +, i.e. $\hat{H}(\vec{r})$ will be called photon Hamiltonian). The expression (2.7) with sign –, i.e. $\hat{H}(-\vec{r})$ will be called anti-photon Hamiltonian. In the same manner can be proved that y and z components of total momentum $\vec{J} = \vec{L} - \vec{S}$ commute with anti-photon Hamiltonian.

The final conclusion is the following: total momentum $\vec{L} + \vec{S}$ is integral of motion for photon, while total momentum $\vec{L} - \vec{S}$ is integral of motion for anti-photon. Up to now we have the proof that total momentum $\vec{L} + \hat{S}$ is free photon integral of motion, but we do not know what magnitude of photon spin is.

If spin is $S = 1/2$, then the following relation is valid:

$$\left(\hat{S}_x - i\hat{S}_y \right)^2 = 0 \; . \tag{3.10}$$

For spin $S > 1/2$ the exponent in (3.10) is higher than 2, i.e. it must be 3,4, ... etc. In (3.10) we shall go over to matrices $\hat{\alpha}$ and $\hat{\beta}$ through relation (3.6). So we obtain:

$$\left(\hat{S}_x - i\hat{S}_y \right)^2 = \frac{\hbar^2}{4} (\hat{\alpha} - i\hat{\beta})^2 = \frac{\hbar^2}{4} [\hat{\alpha}^2 - \hat{\beta}^2 + 2i(\hat{\alpha}\hat{\beta} + \hat{\beta}\hat{\alpha})] = 0$$

---

[1] this corresponds to negative photon energies, i.e. corresponds to $\hat{H}(-\vec{r})$

(in the last stage of the upper proof the relations (3.5) from section 2.1 were used). Consequently, we can conclude that free photon integral of motion represents a total momentum which is the sum of orbital momentum and spin momentum which corresponds to the case when $S = 1/2$.

In the same way can be concluded that anti-photon integral of motion is the sum of orbital momentum and spin momentum which corresponds to spin $S = -1/2$. It should be noticed that negative spin is rather senseless concept so that $\pm S$ really means $\pm S_z$, where $S_z = \hbar/2$.

In nonrelativistic quantum mechanics (Gottifried, 2003; Davidov, 1963) the conclusion that $\hat{J}$ is integral of motion would mean that energy and total momentum of the quantum object can be measured simultaneously and exactly. Since photon is relativistic object (Berestetskii, et al., 1982) the maximal exactness of measuring of photon momentum is given by $\Delta p \, \Delta t \sim \hbar/c$, and consequently energy and total momentum can be determined with an error of the order $\Delta E \, \Delta t \sim \hbar$. The orbital momentum $\hat{\vec{L}}$, as it follows from (3.3), is not integral of motion, but for relativistic object this fact is not essential, since for relativistic objects absolutely exact determining of physical characteristics is in possible.

Considering the correspondence (3.6), photon Hamiltonian which is given by $\hat{H} = c\left(\hat{\alpha}\hat{p}_x + \hat{\beta}\hat{p}_y + \hat{\chi}\hat{p}_z\right)$ can be expressed by means of spin operators in the following form:

$$\hat{H} = \frac{2c}{\hbar}\left(\hat{S}_x\hat{p}_x + \hat{S}_y\hat{p}_y + \hat{S}_z\hat{p}_z\right). \tag{3.11}$$

The obtained form of photon Hamiltonian, which includes operators of translation moment $\hat{\vec{P}}$ and spin $\hat{\vec{S}}$ suggest that a free photon has wealthy internal dynamics that consists of mutual action of its translation and spin characteristics. This "internal life" will be examined further in the paper.

3.3 Unitary transformation of photon Hamiltonian

Photon Hamiltonian (3.11) represents bilinear form in which photon momentum operators are multiplied by spin operators. Since momentum characterizes translation photon motion, and spin characterizes rotation, it is obvious that the internal dynamic structure of a photon is determined by both its translation and rotation characteristics, and that their interaction – considering the form of Hamiltonian (3.11), leads to hybridization of excitations (Agranovich, 2009). Spin operators in (3.11) correspond to spin $S = 1/2$ and its can then is represented by Pauli's operators in the following manner (Tyablikov, 1967):

$$\hat{S}_x - i\hat{S}_y = \hbar P^+ \; ; \quad \hat{S}_x + i\hat{S}_y = \hbar P \; ; \quad \frac{1}{2} - \hat{S}_z = \hbar P^+ P \,. \tag{3.12}$$

Pauli's operators fulfill commutation relations:

$$\left[P_i, P_j^+\right] = \left[1 - 2P_i^+ P_i\right]\delta_{ij} \; ; \quad \left[P_i, P_j\right] = \left[P_i^+, P_j^+\right] = 0 \; ; \quad P_i^2 = P_i^{+2} = 0 \; ; \quad \left(P^+ P\right)_{e.v} = \begin{cases} 0 \; ; \\ 1 \,. \end{cases} \tag{3.13}$$

After substitution of (3.12) in (3.11) (in this formula sign + is retained), we obtain the following form of Hamiltonian:

$$\hat{H} = c\hat{p}_z + c\left[(\hat{p}_x - i\hat{p}_y)P + (\hat{p}_x + i\hat{p}_y)P^+ - 2\hat{p}_z P^+ P\right].$$ (3.14)

This conversion to Pauli operators has been made because the physical picture of processes is clearer through operator's creation and annihilation of excitation.

Operators of moments are linear in operators of creation and annihilation of photon: $P \sim A + A^+$, so it can easily be concluded that mean value of photon Hamiltonian over states $\frac{1}{n!}(A^+)^n P^+ |0\rangle$ is equal to zero. This means that the method of theory of perturbation would be inappropriate for Hamiltonian (3.14) analysis. This is why we would make unitary transformation of photon Hamiltonian with the goal to bring it into the form more suitable for calculation than the form (3.14), i.e. we shall go to equivalent Hamiltonian given by:

$$\hat{H}_{eq} = e^{\hat{W}} \hat{H} e^{-\hat{W}},$$ (3.15)

where:

$$\hat{W} = i\vec{k}\vec{r} + \rho(P - P^+) + i\lambda P^+ P,$$ (3.16)

and ρ and λ are real parameters.

Equivalent Hamiltonian is found using Weil's identity (Tošić, 1978):

$$e^{\hat{W}} \hat{D} e^{-\hat{W}} = \sum_{n=0}^{\infty} \frac{(1)^n}{n!} \underbrace{\left[\hat{W},\left[\hat{W},...\left[\hat{W},...\left[\hat{W},\hat{D}\right]\right]...\right]\right.}_{n-\text{times}}.$$ (3.17)

It has included the terms of the following type: $P + P^+$, $P - P^+$ and P^+P. Undetermined parameter λ has been determined so that the member proportional to $P - P^+$ disappear from equivalent Hamiltonian. The final result of the described procedure is as follows:

$$\hat{H}_{eq} = E_0 + \hat{H} + \hat{H}_S,$$ (3.18)

where $\hat{H}$ is starting Hamiltonian, and

$$E_0 = \hbar c(k_x \sin 2\rho + k_z \cos 2\rho);$$ (3.19a)

$$\hat{H}_S = -g(P + P^+) + 2aP^+ P,$$ (3.19b)

where are:

$$g = \hbar c\sqrt{k_y^2 + k_x^2 \cos^2 2\rho + k_z^2 \sin^2 2\rho - k_x k_z \sin 4\rho}; \quad a = \hbar c(k_x \sin 2\rho + k_z \cos 2\rho) \cdot$$

We shall further analyze free photon behavior using method of Green's functions (Tyablikov, 1967; Tošić, 1978; Rickayzen, 1980; Mahan, 1990; Šetrajčić, et al., 2008). Hamiltonian E_0 is irrelevant in Green function techniques. Starting Hamiltonian $\hat{H}$, as we have already concluded earlier, has zero mean value over states $\frac{1}{n!}(A^+)^n P^+ |0\rangle$. This is why we shall exclude it from calculations. The analysis of photon internal processes will be made with Hamiltonian $\hat{H}_S$.

3.4 Green's function of free photons

Since Pauli operators figure in $\hat{H}_S$ Hamiltonian without various configuration indices, the analysis of spin processes in a free photon will be made by means of anticommutator Pauli Green function:

$$\Gamma(t) = \left\langle\left\langle P(t)\middle|P^+(0)\right\rangle\right\rangle = \Theta(t)\left\langle P(t)P^+(0) + P^+(0)P(t)\right\rangle, \qquad (3.20)$$

where $\Theta(t)$ is Heaviside's step function (Tyablikov, 1967; Tošić, 1978; Rickayzen, 1980). Correlator of anticommutator Pauli's Green's function contains mean value of anticommutator of Pauli's operator of the same configuration index, and according to (3.13) it is equal to one. This fact simplifies evaluation of mean values by means of spectral intensity of Green function.

Differentiating $\Gamma(t)$ per time and using equation of motion for operator P, we come to the following equation:

$$i\hbar\frac{d\Gamma(t)}{dt} = i\hbar\,\delta(t) + 2a\,\Gamma(t) + 2g\,\Delta(t)\,. \qquad (3.21)$$

The Green's function of type: $\left\langle\left\langle\text{const}\middle|P^+\right\rangle\right\rangle$ are equal to zero. The function Δt is given by:

$$\Delta(t) = \left\langle\left\langle P^+(t)P(t)\middle|P^+(0)\right\rangle\right\rangle. \qquad (3.22)$$

Using the same procedure, for defining function $\Delta(t)$ we obtain the following equation:

$$i\hbar\frac{d\Delta(t)}{dt} = g\,\Gamma(t) - g\,F(t), \qquad (3.23)$$

where:

$$F(t) = \left\langle\left\langle P^+(t)\middle|P^+(0)\right\rangle\right\rangle, \qquad (3.24)$$

with defining following equation:

$$i\hbar\frac{dF(t)}{dt} = -2g\,\Delta(t) - 2a\,F(t)\,. \qquad (3.25)$$

In differential equations (3.21), (3.23) and (3.25), Furrier's transformations time-frequency are then made:

$$f(t) = \int\limits_{-\infty}^{+\infty} dt\,\mathrm{e}^{-i\omega t}\,f(\omega); \quad f \equiv (\Gamma, \Delta, F); \quad \delta(t) = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} dt\,\mathrm{e}^{-i\omega t}, \qquad (3.26)$$

so we obtain the system of algebraic equations:

$$\begin{aligned}
(E - 2a)\,\Gamma(\omega) - 2g\,\Delta(\omega) &= \frac{i\hbar}{2\pi}; \\
\Delta(\omega) &= g\left[\Gamma(\omega) - F(\omega)\right]; \\
E\,F(\omega) &= -2\left[g\,\Delta(\omega) + a\,F(\omega)\right].
\end{aligned} \qquad (3.27)$$

Solving this system of equations, we find that:

$$\Gamma(\omega) = \frac{i\hbar}{2\pi} \frac{E^2 + 2aE - 2g^2}{(E^2 - E_0^2)^2}, \tag{3.28}$$

where:

$$E_0 = 2\sqrt{a^2 + g^2} = 2\hbar ck. \tag{3.29}$$

In order to determine spectral intensity of function Γ, it is necessary to break down the right side of the formula (3.20) into common fractions. So, we obtain the following:

$$\Gamma(\omega) = \frac{i}{2\pi} \left[\frac{2g^2}{E_0^2} \frac{1}{\omega} + \left(\frac{1}{2} - \frac{g^2}{E_0^2} + \frac{a}{E_0} \right) \frac{1}{\omega - \omega_0} + \left(\frac{1}{2} - \frac{g^2}{E_0^2} - \frac{a}{E_0} \right) \frac{1}{\omega + \omega_0} \right], \tag{3.30}$$

where: $\omega = E/\hbar$ and $\omega_0 = E_0/\hbar$. Since function Γ is anticommutator function, its spectral intensity is given by the formula (Tyablikov, 1967; Tošić, 1978; Rickayzen, 1980):

$$I_\Gamma(\omega) = \frac{\Gamma(\omega + i\delta) + \Gamma(\omega - i\delta)}{e^{\frac{\hbar\omega}{k_B T}} + 1}; \quad \delta \to +0, \tag{3.31}$$

and using Dirac's formula:

$$\frac{1}{\omega - \omega_k \pm i\delta} = \text{P.V.} \left\{ \frac{1}{\omega - \omega_k} \right\} \mp i\pi \delta(\omega - \omega_k), \tag{3.32}$$

where P.V. denotes principal value of integral, we find the explicit expression for spectral intensity:

$$I_\Gamma(\omega) = \frac{2g^2}{E_0^2} \frac{\delta(\omega)}{e^{\frac{\hbar\omega}{k_B T}} + 1} + \left(\frac{1}{2} - \frac{g^2}{E_0^2} + \frac{a}{E_0} \right) \frac{\delta(\omega - \omega_0)}{e^{\frac{\hbar\omega}{k_B T}} + 1} + \left(\frac{1}{2} - \frac{g^2}{E_0^2} - \frac{a}{E_0} \right) \frac{\delta(\omega + \omega_0)}{e^{\frac{\hbar\omega}{k_B T}} + 1}. \tag{3.33}$$

Now we can defined the expression for correlation function of a free photon as:

$$\left\langle P^+(0)P(t) \right\rangle \equiv \int_{-\infty}^{+\infty} d\omega\, e^{-i\omega t} I_\Gamma(\omega) = \frac{2g^2}{E_0^2} \frac{1}{2} + \left(\frac{1}{2} - \frac{g^2}{E_0^2} + \frac{a}{E_0} \right) \frac{e^{-i\omega_0 t}}{e^{\frac{\hbar\omega_0}{k_B T}} + 1} + \left(\frac{1}{2} - \frac{g^2}{E_0^2} - \frac{a}{E_0} \right) \frac{e^{i\omega_0 t}}{e^{\frac{\hbar\omega_0}{k_B T}} + 1}. \tag{3.34}$$

Next, we can calculate expression for concentration of spin excitations of a free photon. It is obtained from (3.34), if we take in it that $t = 0$, i.e.

$$\left\langle P^+ P \right\rangle = \frac{1}{2} - \frac{a}{E_0} \tanh \frac{\hbar ck}{k_B T}. \tag{3.35}$$

Combining formulae for a over formula (3.19b), and E_0 from (3.29), and converting to sphere coordinate system, we find that:

$$\frac{a}{E_0} = \frac{1}{2} \left(\sin 2\rho \sin\theta \cos\varphi + \cos 2\rho \cos\theta \right).$$

In accordance with this and formula (3.35), we get the following expression for ordering parameter of spin subsystem in a free photon:

$$\sigma = 1 - 2\langle P^+ P \rangle = (\sin 2\rho \sin \theta \cos \varphi + \cos 2\rho \cos \theta) \tanh \frac{\hbar ck}{k_B T} . \tag{3.36}$$

The set of results of this section requires some explanations. The most interesting results is that energy for spin translation from $\hbar/2$ to $-\hbar/2$ is $2\hbar ck$. This can be explained on the basis of measuring process in which incident photon bean is reflected by measuring devices. The momentum of incident phonon is $\hbar k$ while the momentum reflected phonon is $-\hbar k$. So we obtain the change of photon momentum $\Delta p = \hbar k - (-\hbar k) = 2\hbar k$, and consequently the energy change $\Delta E = 2\hbar ck$. The energy $-\hbar ck$ corresponds to anti-photon, so that we can consider the described process as a transformation of photon to anti-photon. In this process the spin change takes place, also (the Green's function $\Gamma(t) = \langle\langle P(t) | P^+(0) \rangle\rangle$ was calculated). Since photon and anti-photon spins have opposite signs the change the spin is $\Delta S = \hbar/2 - (-\hbar/2) = \hbar$. The value of ΔS is equal the value $\hbar$ and this is eigen-value of spin $s = 1$. This is the reason for behaving of photon as particle with spin $s = 1$.

The polar and azimuthally dependences of ordering parameter comes from the fact that incident bean must not be always orthogonal to the plane of measuring device.

4. Conclusions

1. The analysis of single photon behaving in coordinate systems of various geometries has shown the following:
    -   In Cartesian coordinate system the components of single photon eigen-vector are progressive and regressive plane waves. In other words, the photon behaving is characterized by coordinates and momentum, only.
    -   In cylindrical coordinate system the components of single photon eigen-vector of are progressive and regressive plane waves, but only along z axis. In comparison to the results of analysis in Cartesian coordinates the fact that in $x\ y$ planes photon oscillations are attenuated according to $\rho^{-1/2}$ law represents a generalization. This conclusion has been based on asymptotic behaving of Bessel functions.
    -   The most authentic result is obtained by the analysis in spherical coordinates. Photon states do not depend on angles, and they are damped according to r^{-1} law, where r is radius of the sphere. It has also turned out that orbital momentum of a free photon is equal to zero (this is understandable considering the fact that it moves along a straight line). The important element determining the behavior of a free photon is spin. One components of photon eigen-vector corresponds to spin $S_z = 1/2$ projection, while the other component corresponds to spin $S_z = -1/2$ projection. This was concluded on the basis of the fact that its eigen-vector components are Bessel functions having indices $+1/2$ and $-1/2$.
        The last result shows that linearization of photon Hamiltonian gives more complete picture of single photon than Kline-Gordon's approach.
2. Concluding the exposed analysis we shall try to connect the results obtained in series of experimental investigation of photon orbital momentum (Beth, 1936; Leach, et al., 2002; Allen, et al., 1992; Allen, 1966; He, et al., 1995; Friese, et al., 1996; Markoski, et al., 2008; van Enk & Nienhuis, 2007; Santamato, et al., 1988; O'Neil, et al., 2002; Volke-Sepulveda, et al., 2002). We shall not describe all quoted experiments. Instead of it we shall describe the essential idea: the orbital momentum of photon was determines from the changes of

torque of rotating particles. These changes where lied in some interval, so that the values of orbital momentum have had determined dispersion. As it vas said at the end of first section, such result is expectable for relativistic objects, in this case for photons. The azimuthally dependence of measured results is also predicted by the theory exposed in last Section.

Ending this analysis it should by noticed out that on the bases of given analysis the photon reflection can be considered as a transformation of photons to anti-photons.

5. Acknowledgements

6. References

Agranovich, V.M. (2009). *Excitations in Organic Solids*, University Press, ISBN 13 9780199234417, Oxford

Allen, P.J.; (1966). A Radiation Torque Experiment, *Amer.J.Phys.* Vol.34, No.12 (Dec.1966), pp.1185-1192, ISSN 0002-9505

Allen, L.; Beijersbergen, M.W.; Spreeuw, R.J.C. & Woerdman, J.P. (1992). Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes *Phys.Rev.*A, Vol.45, No.11 (Jun 1992), pp.8185-8189

Berestetskii, V.B.; Lifshitz, E.M & Pitaevskii, L.P. (1982). *Quantum Electrodynamics,* Pergamon, ISBN 0080265049, Oxford

Beth, R.A. (1936). Mechanical Detection and Measurement of the Angular Momentum of Light. *Phys.Rev.*Vol.50, No.2 (Jul 1936), pp.115-125

Davydov, A.S. (1963). *Quantum Mechanics,* Nauka, Moscow 1963 (in Russian)

Davydov, A.S. (1976). *Quantum Mechanics*, Pergamon, ISBN 0080204384, London

Delić, N.V.; Pelemiš, S.S. & Šetrajčić, J.P. (2008). About Eigen-Problem of Single Photon Hamiltonian, *Proceedings of 26th International Conference on Microelectronics*, pp. 129-130, ISBN 978-1-4244-1881-7, Niš, May 2008, Eds of the IEEE, Danvers

Dirac, P.A.M. (1958). *Principles of Quantum Mechanics,* 4 Ed., *University Press*, ISBN 0198520115, Oxford

Friese, J.M.E.; Enger, J.; Rubinsztein-Dunlop, H. & Heckenberg, N.R. (1996). Optical Angular-Momentum Transfer to Trapped Absorbing Particles, *Phys.Rev.* A,Vol.54, No.2 (Aug.1996), pp.1593 -1596.

Gottifried, K. (2003). *Quantum Mechanics: Fundamentals*, Springer Verlag, ISBN 9780387955766, Massachuset

He, H.; J.Friese, M.E.; Heckenberg, N.R. & Rubinsztein-Dunlop, H. (1995). Direct Observation of Transfer of Angular Momentum to Absorptive Particles from a Laser Beam with a Phase Singularity, *Phys.Rev.Lett.* Vol.75, No.5 (Jul 1995), pp.826-829

Holbrow, C.H., Galvez E. & Parks, M.E. (2001). Photon Quantum Mechanics and Beam Splitters, *Am.J.Phys.*, Vol.70, No.3 (Nov. 2001), pp.260-265

Janke, E.; Emde, F. & Losch, F. (1960). *Tafeln Hoherer Funktionen*, p.176, Teubner, ISBN Buchnummer des Verkäufers 956816, Stuttgart

Kadin, A.M. (2005). *Quantum Mechanics without Complex Numbers: A Simple Model for the Electron Wavefunction including Spin*, ArXiv Quantum Physics, http://arxiv.org/abs/quant-ph/0502139

Korn, G.A. & Korn, T.M. (1961). *Mathematical Handbook for Scientists and Engineers,* Mc Graw-Hill, ISBN 0486411478, London

Leach, J.; Padgett, M.J.; Barnett, S.M.; Franke-Amold, S. & Courtial, J. (2002). Measuring the Orbital Angular Momentum of a Single Photon *Phys.Rev.Lett.* Vol.88, No.25 (Jun 2002), pp.297-300

Mahan, G. (1990). *Many Particle Physics*, Plenum Press, ISBN 0-306-43423-7, New York

Markoski B.; Pelemiš, S. & Mihailović, J., Applying Neuron Network to Improve Characteristic of Video Encoder, *Proceedings of 26th International Conference on Microelectronics*, pp.131-143, ISBN 978-1-4244-1881-7, Niš, May 2008, Eds of the IEEE, Danvers

Messiah, A. (1970). *Quantum Mechanics,* North-Holland, ISBN-10 0486409244, Amsterdam

O'Neil, A.T.; Mac Vicar, I.; Allen, L. & Padgett, M.J. (2002). Intrinsic and Extrinsic Nature of the Orbital Angular Momentum of a Light Beam, *Phys.Rev.Lett.* Vol.88, No.5 (Jan.2002), pp.053601-053605

Planck, M. (1901). Ueber das Gesetz der Energieverteilung im Normalspectrum (On the Law of Distribution of Energy in the Normal Spectrum), *Annalen der Physik (Leipzig),* Vol.309, No.3, pp.553-563

Rickayzen, G. (1980). *Green's Functions and Condensed Matter*, Academic Press, ISBN 1040674597, London

Santamato, E.; Daino, B.; Romagnoli, M.; Settembre, M. & Shen, Y.R. (1988) *Phys.Rev.Lett,* Vol.61, No.1 (Jul 1988), pp.113-116

Sapaznjikov, M. (1983). *Anti-World – a Reality*, Znanie, YU ISBN 86-19-01299-1, Moscow (in Serbian).

Šetrajčić, J.P.; Ilić, D. I.; Markoski, B.; Šetrajčić, A.J.; Vučenović, S.M.; Mirjanić, D.Lj.; Škipina, B. & Pelemiš, S.S. (2008). Adapting and Application of the Green's Functions Method onto Research of the Molecular Ultrathin Film Optical Properties, *Book of Abstracts of 15th Central European Workshop on Quantum Optics*, pp.34-35, ISBN 978-86-82441-23-6, Belgrade, May-June 2008, Institute of Physics, Belgrade

Torn, J.J.; Neel, M.S.; Donato, V.W.; Bergreen, G.S.; Davies, R.E. & Beck, M. (2004). Observing the Quantum Behavior of Light in an Undergraduate Laboratory, *Am.J.Phys.* Vol.72, No.9 (Sept. 2004), pp.1210-1219

Tošić, B.S. (1978). *Statistical Physics*, Faculty of Sciences, Novi Sad (in Serbian)

Tošić, B.S.; Delić, N.V.; Mašković, Lj.D.; Ilić, D.I.; Šetrajčić, J.P. & Jaćimovski S.K. (2008). Brain Photons, *Book of Abstracts of 15th Central European Workshop on Quantum Optics,* pp. (20-21), ISBN 978-86-82441-23-6, Belgrade, May-June 2008, Institute of Physics, Belgrade

Tyablikov, S.V. (1967). *Methods in the Quantum Theory in Magnetism*, Plenum, New York

Yao, W.; Liu, R.B. & Sham, L.J. (2005). Theory of Control of the Spin-Photon Interface for Quantum Networks, *Phys.Rev.Lett.* Vol.95, No.3, (Jul, 2005), pp.030504-030508

van Enk, J.S. & Nienhuis, G. (2007). Photons in Polychromatic Rotating Modes, *Phys.Rev.*A, Vol.76, No.5 (Nov.2007), pp.053825–1-11

Volke-Sepulveda, K.; Garcés-Chávez, V.; Chávez-Cerda, S.; Arlt, J. & Dholakia, K. (2002). Orbital angular momentum of a high-order Bessel light beam *J.Opt.*B: *Quantum Semiclass.Opt. Vol.*4 (April 2002), pp.82-89