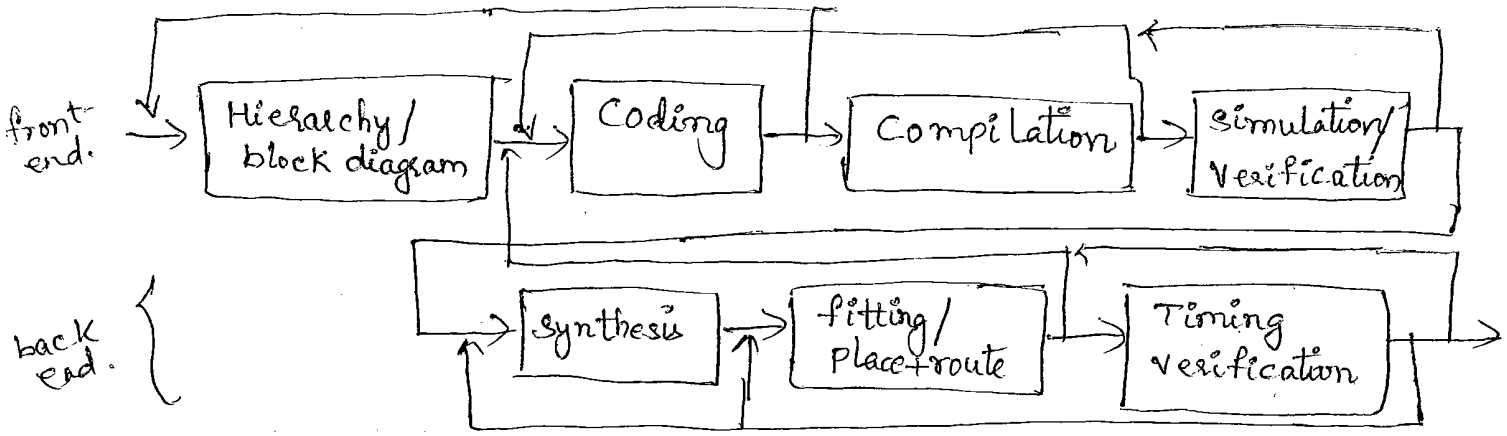
VLSI CIRCUIT DESIGN PROCESS

VLSI DESIGN FLOW



The "front-end" begins with figuring out basic approach & building blocks at the block diagram level.

The next level is writing HDL code for modules, their interfaces, & their internal details. The HDL compiler analyzes code for syntax errors and also checks it for compatibility with other modules on which it relies.

The HDL simulator allows to define and apply I/ps to design and to observe its o/ps without having to build physical circuits. (functional verification).

Back end stage starts with synthesis, thats converts HDL description into a set of primitives or components that can be assembled in the target technology.

This is called netlist that specifies how they should be interconnected.

In fitting step, a fitter maps the synthesized components onto available device resources.

Place & route process lays components and finds ways to connect them. The designer can usually specify additional constraints at this stage, like placement of modules within a chip or the pin assignments of external I/p & O/p pins.

The final step is post-fitting timing verification of the fitted circuit. At this stage actual ckt delays due to wire lengths, electrical loading and other factors can be calculated.

MOS LAYERS

Mos circuits are formed on 4 basic layers:

(i) n-diffusion
(ii) p-diffusion
(iii) Poly Si
(iv) Metal.

→ The thinox mask region includes n-diff, p-diff & transit channels.

→ Poly Si & thinox regions interact so that a transistor is formed where they cross one another.

→ Contacts formed by joining layers.

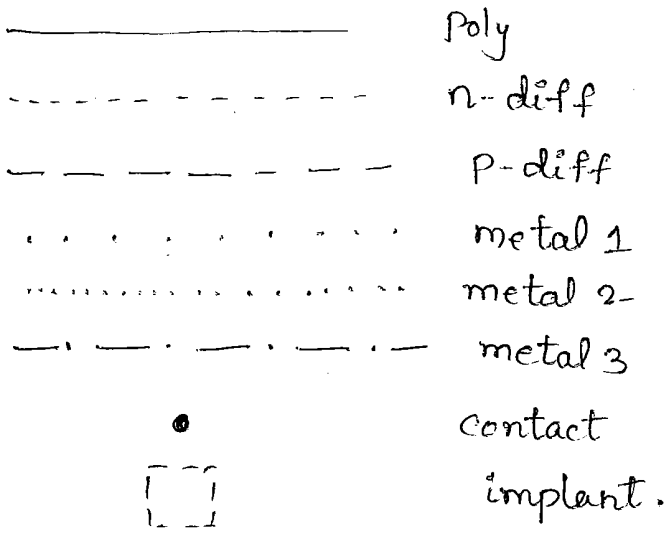
→ some processes includes 2nd metal layer, also 2nd polysi layer.

→ Bipolar transistors can be included in design by addition of extra layers to CMOS process.

STICK DIAGRAMS :

A stick diagram is a cartoon of a chip layout. It represents rectangles with lines which represent wires and component symbols.

Representation



Colors:- (nMOS design style).

Red: Poly
Green: n-diff
Blue: Metal
Black: Contact
Yellow: implant

(CMOS design style)

Red: Poly
Green: n-diff
Yellow: P-diff
Blue: Metal
Black: Contact

→ A Tr is formed wherever poly crosses diffusion.

→ Area and aspect ratio are difficult to estimate from stick diagrams.

→ Faster to design.

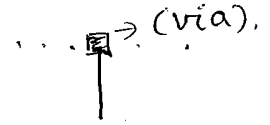
→ Important tool for layouts built from large cells & testing connections b/w cells.

→ A stick diagram is interface b/w symbolic ckt & the actual layout.

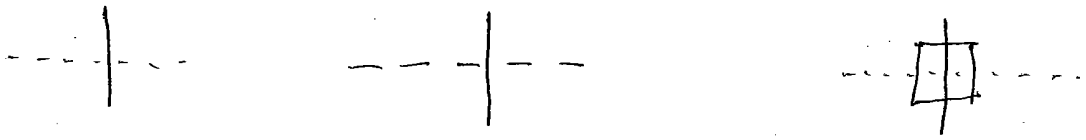
→ Often used to solve routing problems.

→ Rules:- (1) When two or more sticks of same type cross or touch each other represent electrical contact.

(2) When two or more sticks of different type cross or touch each other there is no electrical contact.



(3) When poly crosses diff^n, it represents MOSFET.



nMOS pMOS depletion nMOS.

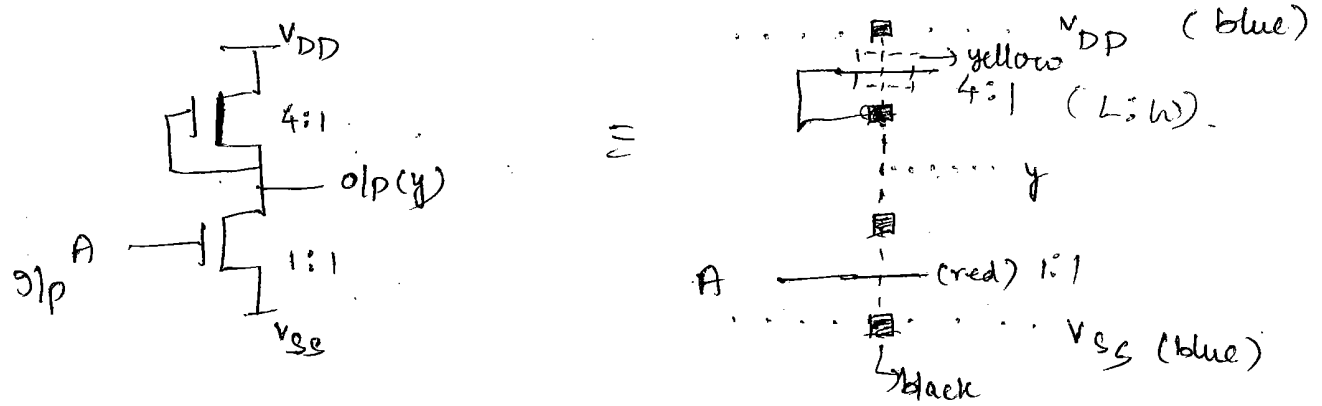
nMOS Design Style :-

A transistor is formed wherever poly crosses n-diff (red over green) and all diffusion wires (interconnections) are n-type (green).

Draw V_{DD} & gnd rails in parallel using metal (blue) allowing enough space b/w them for other circuit elems.

ex: ① nMOS inverter

$y = \bar{A}$

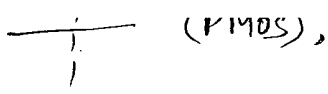


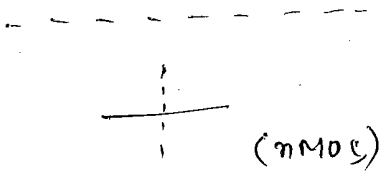
implant in yellow.

CMOS Design Style :-

→ Logical extension of nMOS approach.

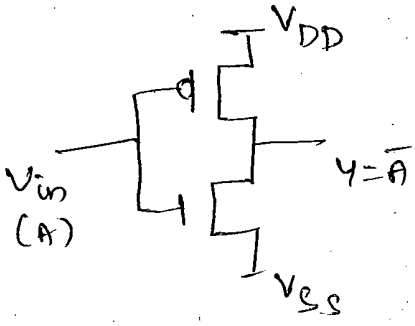
→ The two types of trs used 'n' & 'p' are separated in the stick layout by the demarcation line above which all p-type devices are placed. The n-devices (green) are placed below the demarcation line. and are thus located in the p-well.

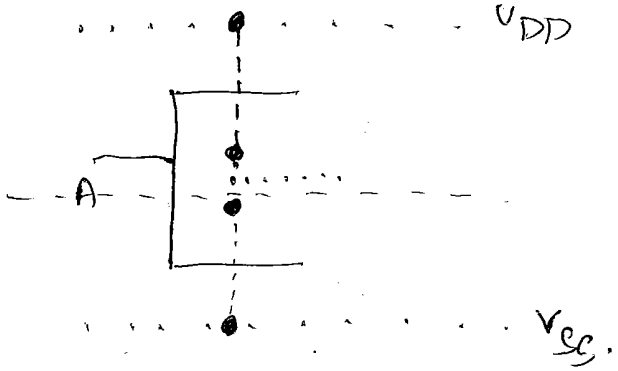




Diffusion paths must not cross the demarcation line and n-diff & p-diff wires must not join. The 'n' & 'p' features are normally joined by metal where a connection is required.

ex:- CMOS inverter.

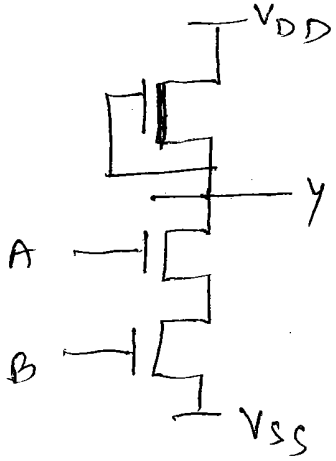


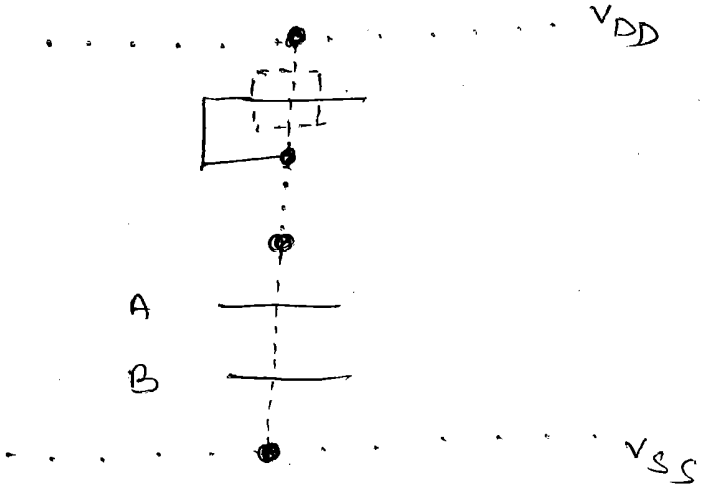


More Examples

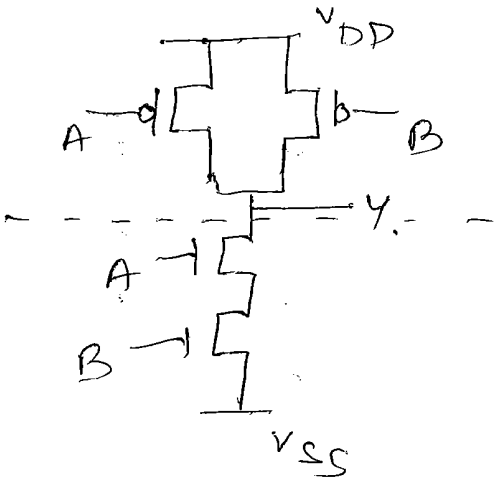
1) $Y = \overline{A \cdot B}$ (Nand Gate)

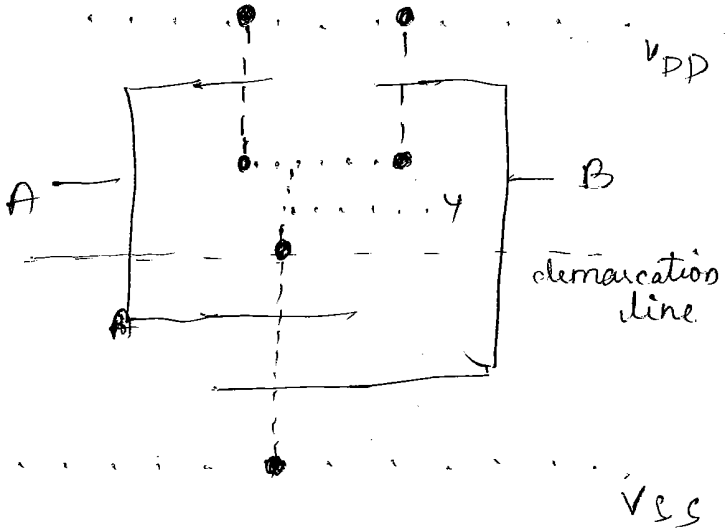
nMOS Logic





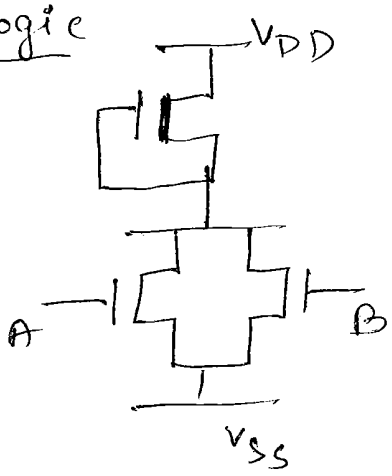
CMOS logic

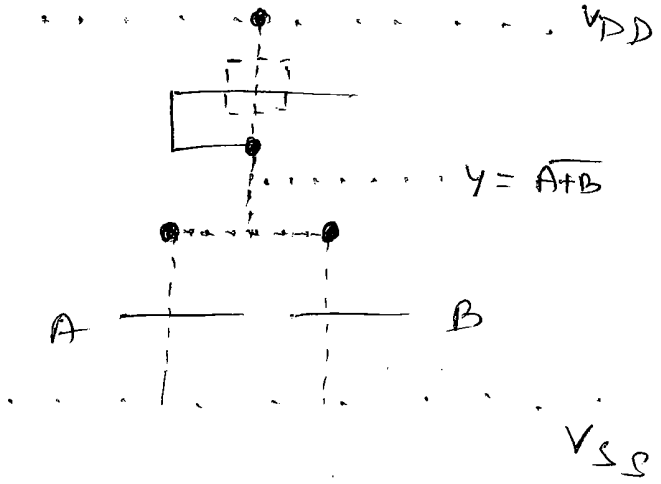




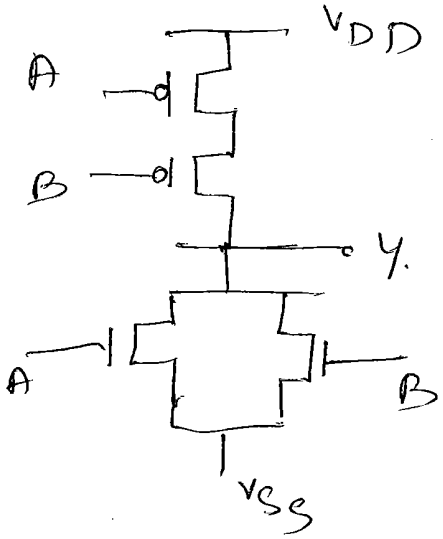
2) $Y = \overline{A+B}$ (NOR).

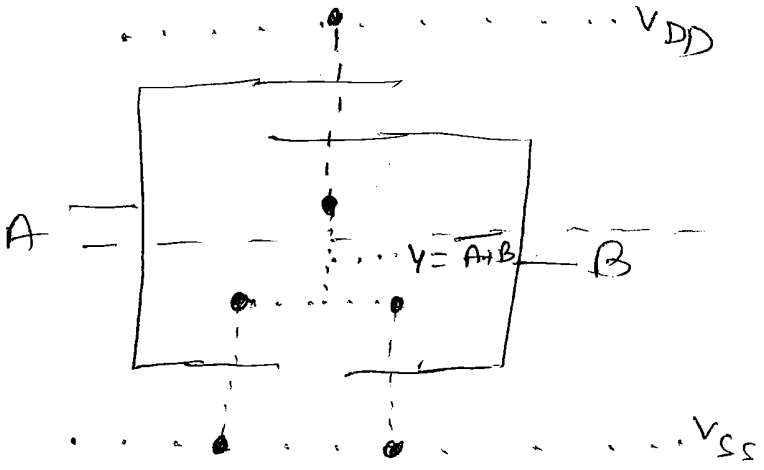
nMOS logic





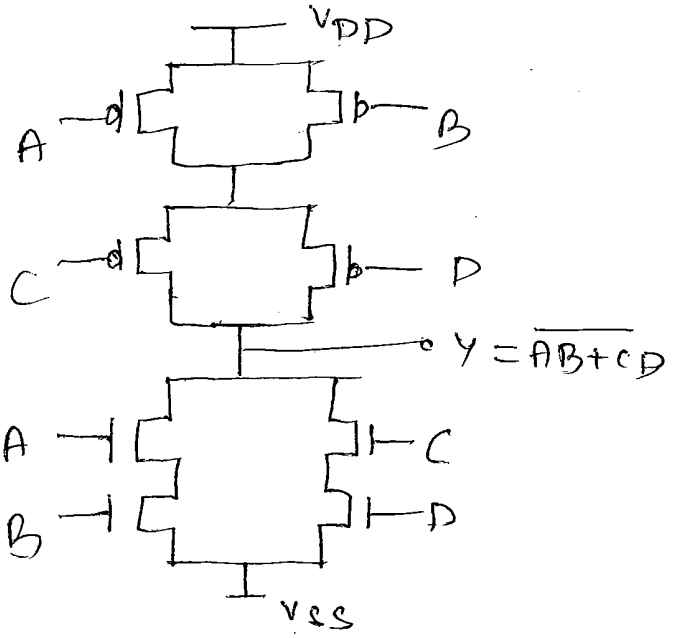
3) CMOS logic

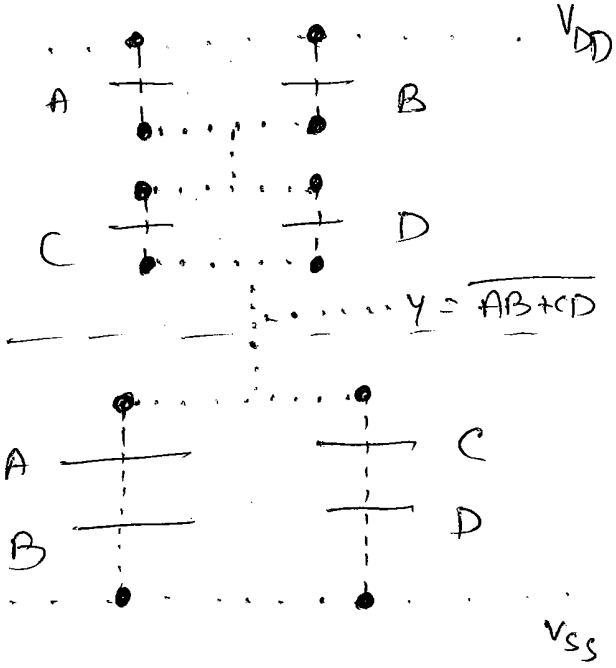




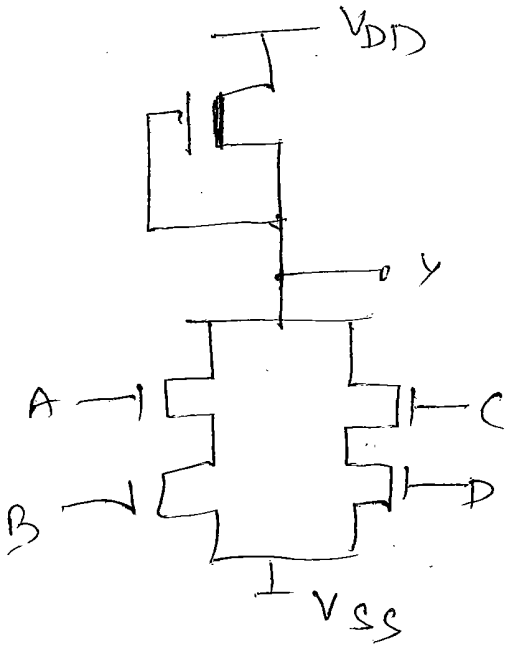
3) $Y = AB + CD$

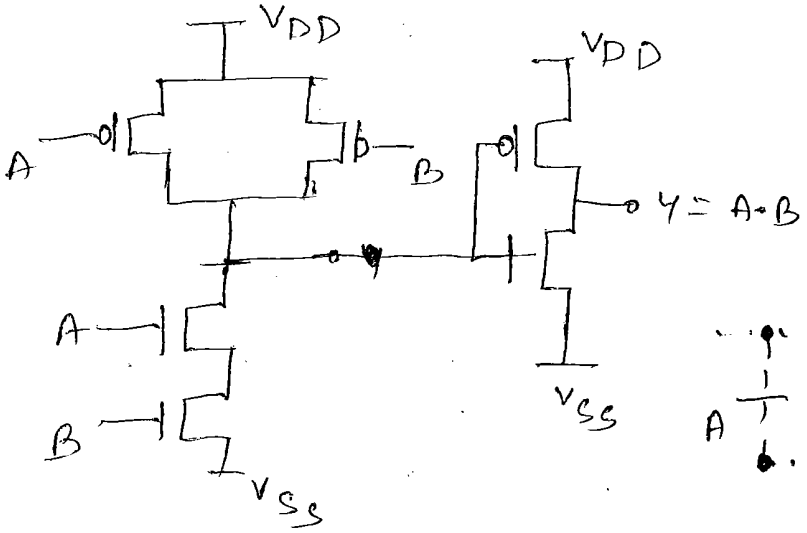
nMOS logic

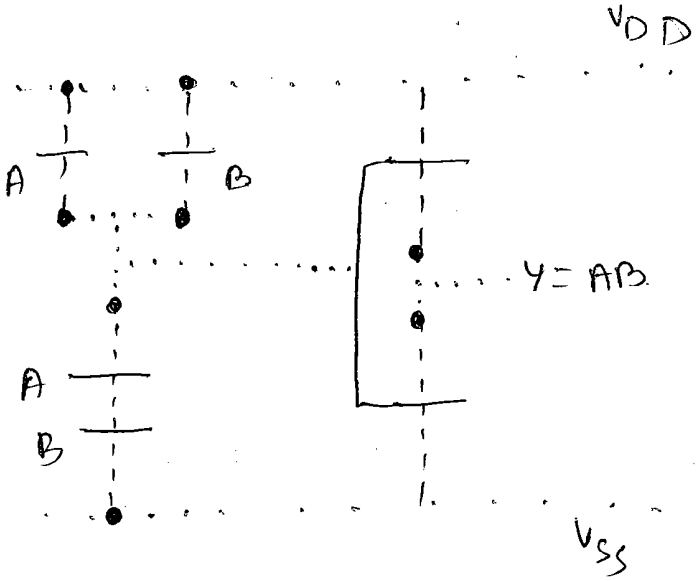




similarly nMOS logic.

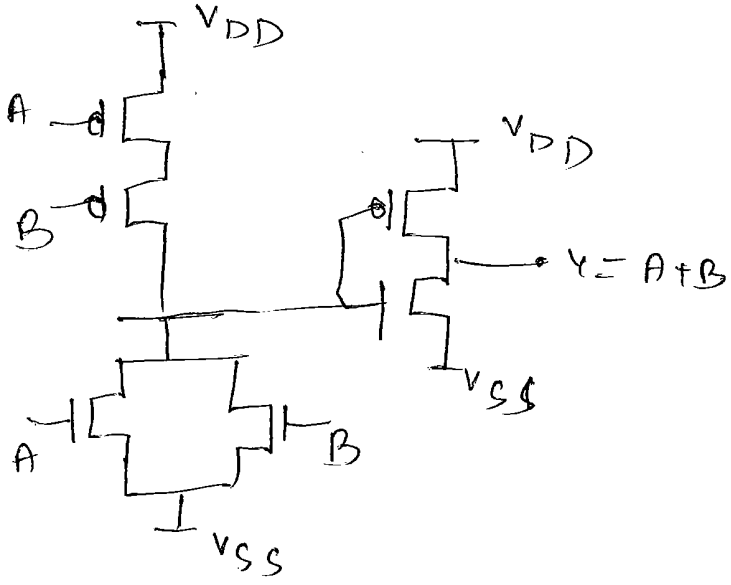


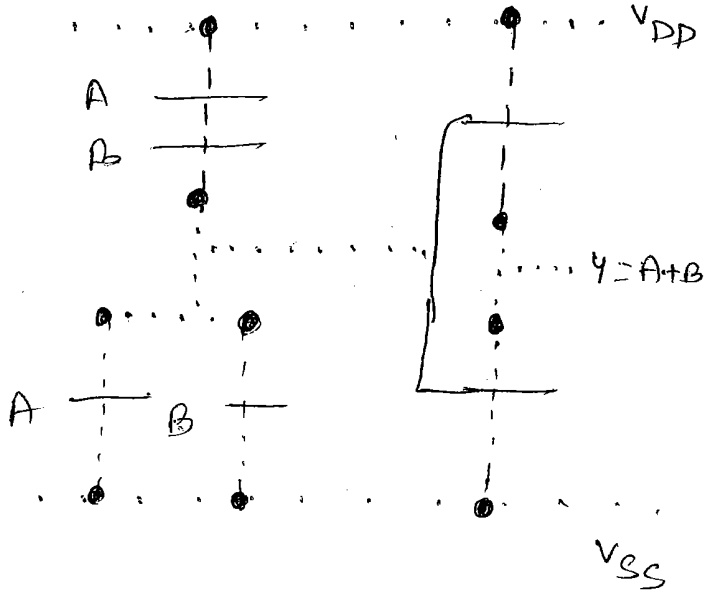




* similarly nMos. logic

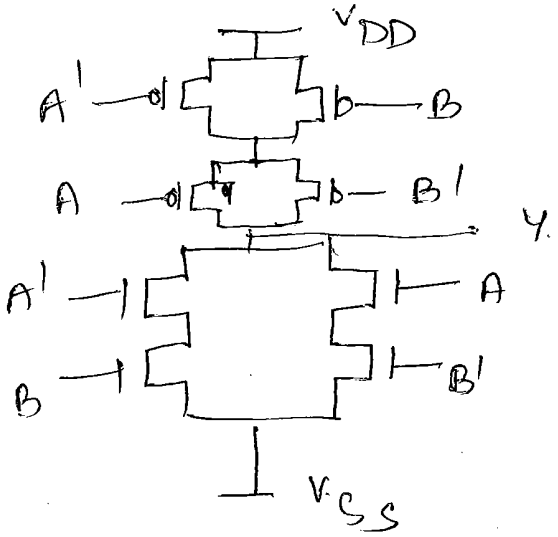
5) $Y = A+B = \overline{\overline{A+B}}$

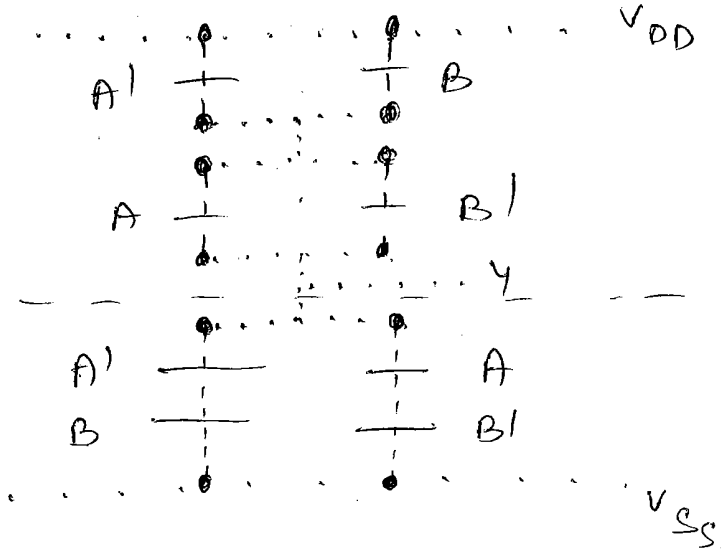




* Do nMos logic.

6) $Y = A'B + AB'$ (xnor)





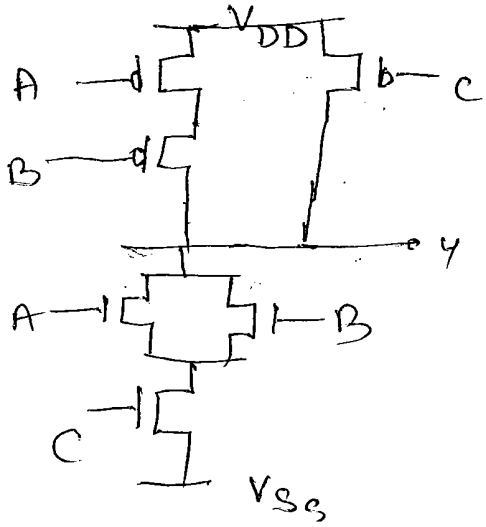
* for nMOS logic: Replace Pull up with nMOS depletion mode transistor with gate connected to source.

7) Similarly $Y = A'B + AB'$ (xor gate).

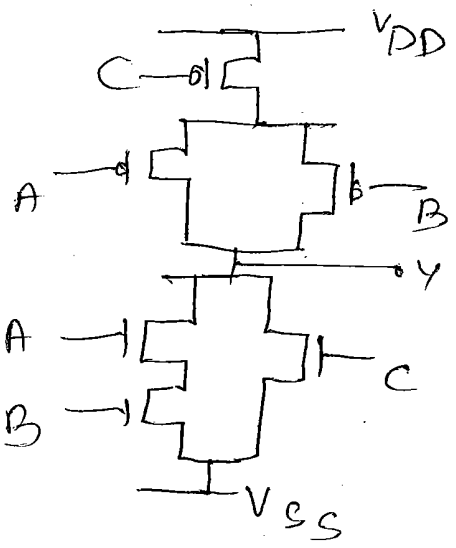
$$Y = \overline{A'B + AB'}$$

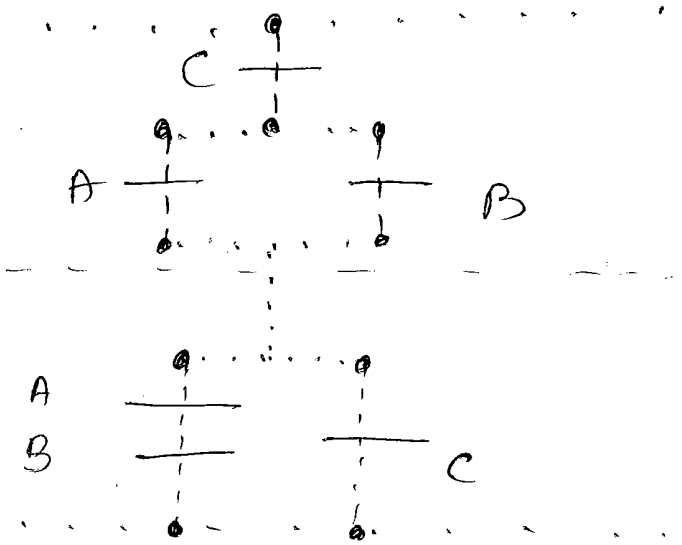
connect inverter to o/p of xnor gate

8) $Y = \overline{(A+B)C}$



9) $Y = AB + C$





DESIGN RULES AND LAYOUTS

→ Design rules allow the translation of circuit design concepts, usually in stick diagrams or symbolic form into actual geometry in silicon.

→ Design rules govern the layout of individual components and interactions – spacings & electrical connections b/w the components.

→ Design rules are specific to a particular semiconductor manufacturing process. It determines low-level props of chip designs (how small individual logic gates are made, how small can the wires be).

→ As small a component size as possible is desired to increase the no. of functions in the chip. But fabrication error arise such as shorting together of wires, or absence of connection b/w wires, faulty transistors etc.

Design rules are used to minimize the occurrence of common fabrication problems and to bring yield of correct chips to acceptable level.

One of fabrication problem is that a wire or any feature being made too wide or too narrow. A too narrow wire may never conduct or may burn off when conducting. A wide wire may short itself with other wires. If poly crosses or cuts diffusion, then it is formation of a new element.

Remedy

1) Introduction of spacing rules
2) Introduction of min - width rules.

Min width rule:

Gives min size for layout element. It also ensures that even with minutest variations, the elem will be of acceptable size.

Spacing rule:-

Gives min distance b/w the edges of layout elems, so that even with minor variations it will not cause the element to overlap nearby layout elems.

Composition rules:-

Ensures that components are well- formed.

Construction rules (vias)

→ Material on both layers to be connected must extend beyond SiO_2 cut and cut must be [illegible]

Scalable Design Rules

→ Design rules can be scaled in terms of 'λ', which is the size of the smallest elem in the layout.

→ When devices shrink, layouts need not be completely redesigned. All features can be measured in integral multiples of 'λ'.

→ By choosing a value for λ, all dimensions set at a scalable layout.

→ Scalable layouts are advantageous as chips become faster as size shrinks.

→ Digital ckt designs scale, b'coz the cap loads that must be driven by logic gates shrink faster than the currents supplied by the Trs.

→ Assuming that the basic physical paramtrs of chip are shrunk by a factor of $1/x$.

Length = $L \rightarrow L/x$; Width = $W \rightarrow W/x$.

Thickness = $D \rightarrow D/x$

Supply V_t = $V_{DD} - V_{SS} \rightarrow (V_{DD} - V_{SS})/x$

Doping = $N_d \rightarrow N_d/x$

Tr transconductance:

$$\hat{g_m} = x \cdot g_m$$

Threshold vtg:

$$\hat{V_t} = \frac{V_t}{x}$$

Satⁿ Drain Current:

$$I_{ds} = K \frac{W}{L} [(V_{gs} - V_t)^2] .$$

$$K = \frac{\mu \epsilon_0 \epsilon_{ins}}{D}$$

$$\hat{K} = K \cdot x$$

$$\Rightarrow \frac{\hat{I_{ds}}}{I_{ds}} = \frac{1}{x}$$

&

$$\frac{\hat{C_g}}{C_g} = \frac{1}{x}$$

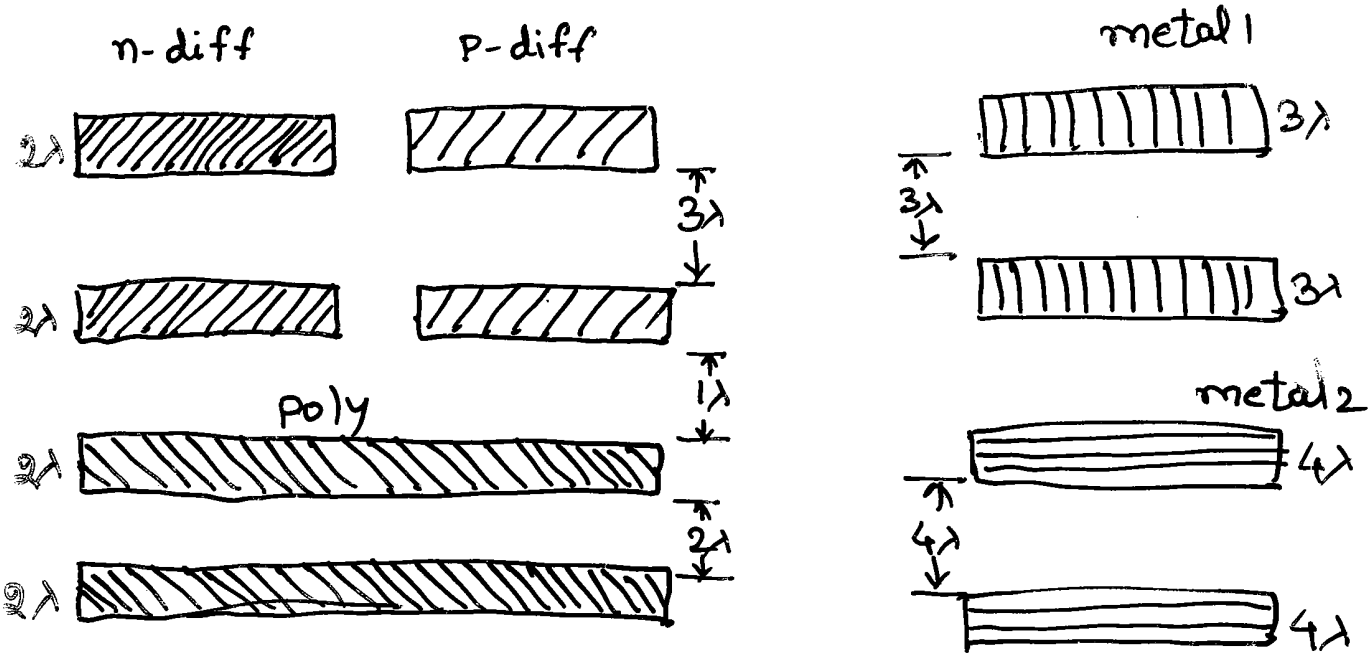
→ $\frac{CV}{I}$ is measure of speed of ckt over scaling.

$$\therefore \frac{\hat{C}\hat{V}/\hat{I}}{CV/I} = \frac{1}{x}$$

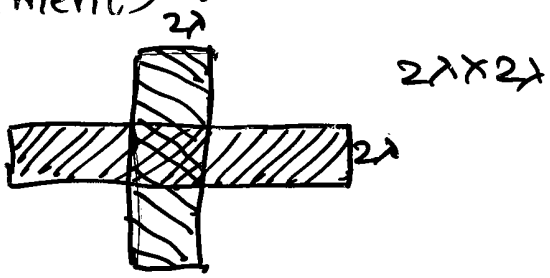
⇒ Scaling is done on λ, thus $\frac{\hat{\lambda}}{\lambda} = \frac{1}{x}$.

(thus speed up by factor x)

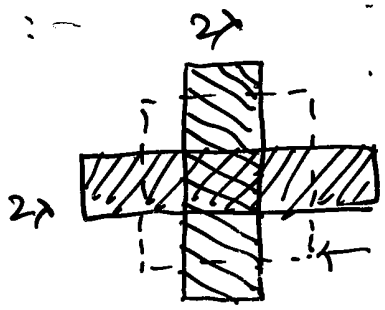
Design rules for wires (nMOS & CMOS)



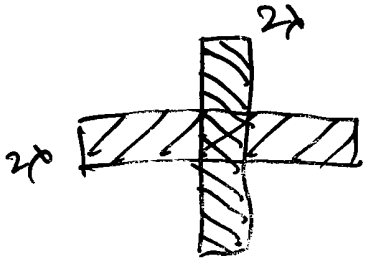
nMOS (enhancement) :-



nMOS (Depletion) :-



pMOS (Enhancement)



Summary:

1) Metal 1 : min-width = 3λ
min-sep^n = 3λ

2) Metal 2 : min-width = 4λ
min-sep^n = 4λ

3) Poly : min-width = 2λ
min poly-poly sep^n = 2λ

4) P & n diffusion: min width = 2λ
min sep^n b/w same diff = 2λ.

5) Tubs: 10λ wide.

Min Separation b/w tub & src/drain = 5λ.

Tub Tie: P-tub tie : 2λ × 2λ cut, 4λ × 4λ metal, 4λ × 4λ p$^+$ diff
n-tub tie.

Construction rules:-

1) Transistors : width = 2λ
Length = 2λ

2) Poly extends 2λ beyond active region.

3) Diffusion extends 2λ.

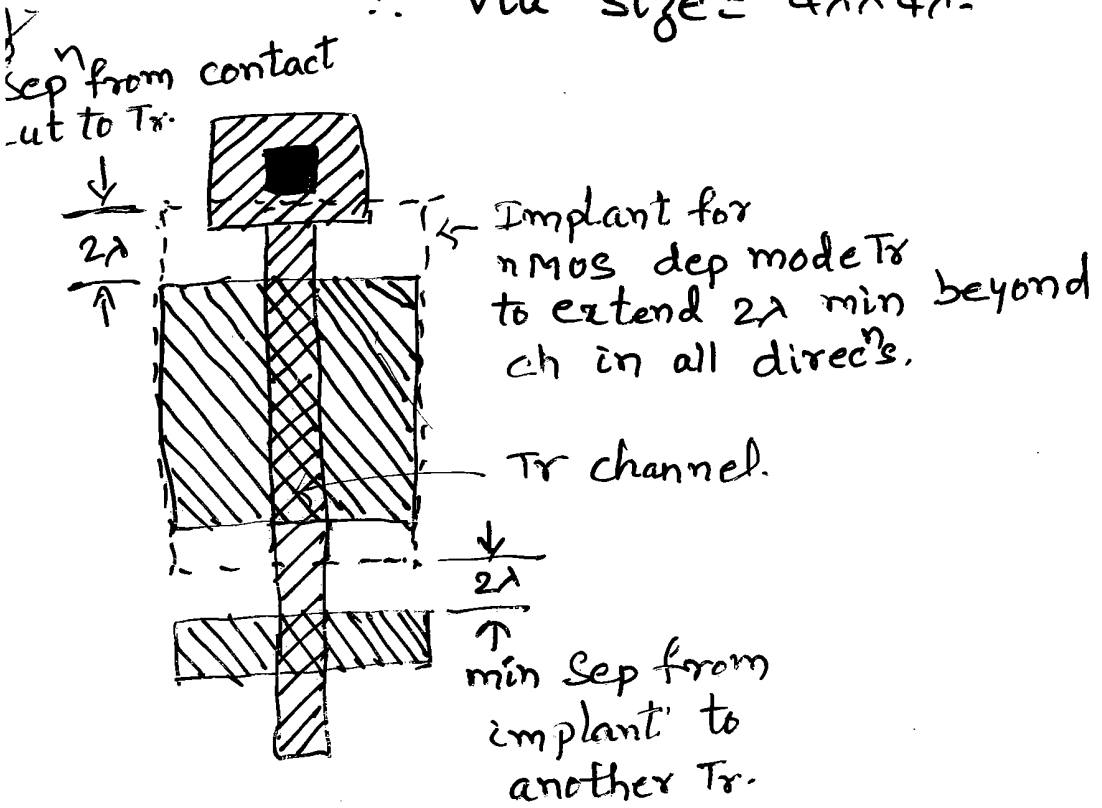
4) Active region must be atleast 1λ from poly-metal via, 2λ from another Tr,

Vias :

→ cuts : $2\lambda \times 2\lambda$

→ Material on both layers extend 1λ in all dirⁿ from cut.

∴ via size = $4\lambda \times 4\lambda$.



Extentions & Separations (Trs.)

Contact Cuts :

3 ways to make contacts b/w poly & diffusion in nMOS ckts :

(i) poly to metal then metal to diffusion
(ii) buried contact (poly to diff).
(iii) butting contact (poly to diff using metal).

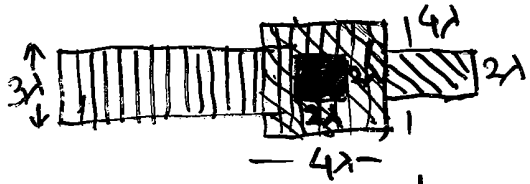
The $2\lambda \times 2\lambda$ contact cut indicates an area in which the oxide is to be removed down to the underlying polysi or diff surface

When deposition of metal layer takes place the metal is deposited thru contact cut areas onto underlying area so that contact is made b/w the layers.

ex:-

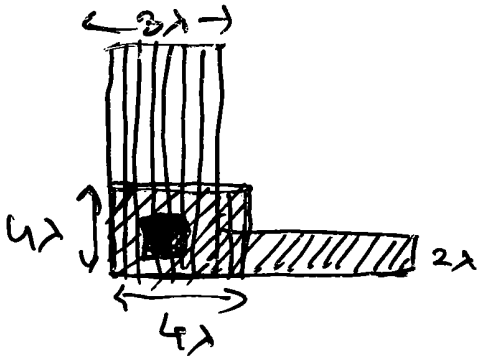
1) Metal1 to polysi or to diffusion.

Metal1 to poly.

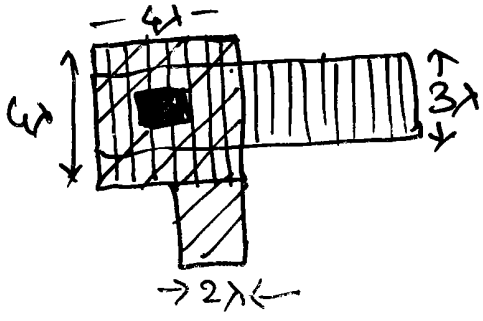


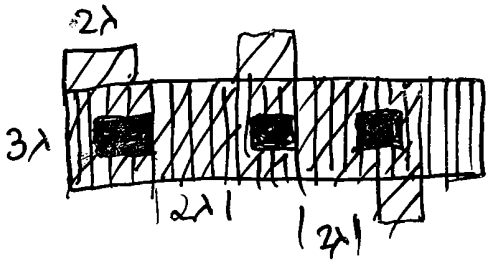
$2\lambda \times 2\lambda$: Cut centered on $4\lambda \times 4\lambda$ superimposed areas of layers to be joined in all cases.

Metal1 to n-diff



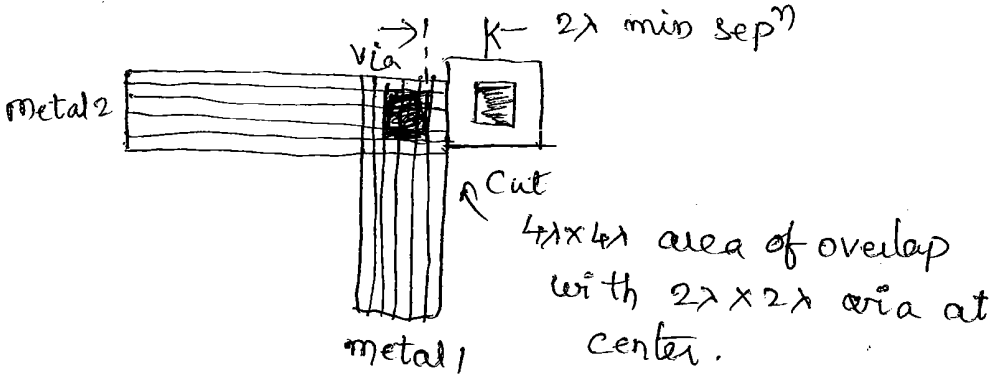
Metal1 to P-diff





Min Sep, Multiple cut

2) Via (Contact from metal2 to metal1)



NOTE: min sepⁿ b/w ~~n~~diffⁿ wire and poly wire = 1λ.

→ Contact cuts are also known as via cuts.

→ 4λ×4λ size.

→ Contact cut types:

(i) n/p diffⁿ to polysi
(ii) poly to metal1
(iii) n/p diffⁿ to metal1
(iv) metal1 to metal2.

Contact b/w polysi and diffusion wires can be done in 3 ways:-

(a) Polysi to the metal and then metal to polysi.

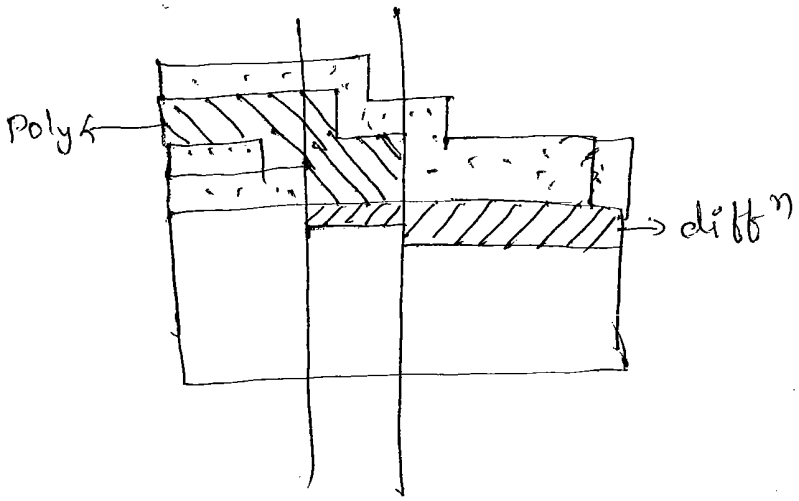
* oxide is removed from 2λ×2λ contact cut down to underlying polysi wire. Then metal is deposited. It flows thru the oxide etched area to polysi area. Then polysi is deposited on the surface, which acts as conduction path.

(b) Buried Contact

Before starting the process, there is oxide layer on Si surface, oxide is etched to expose the underlying Si surface. polysi is deposited on the surface.

In the next step, diffⁿ is carried out on the exposed surface. When diffⁿ takes place impⁿ will diffuse into polysi as well as diffused area within the contact area.

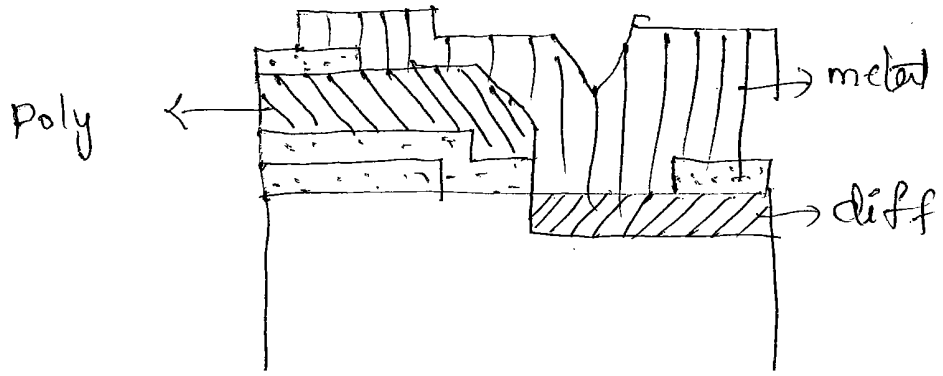
This ensures a satisfactory connecⁿ b/w polysi & diffⁿ. Buried contacts are smaller than butting contacts.



(c) Butting Contact:-

→ a complex process.

→ 2λ × 2λ contact cut is made down to each layers to be joined. Layers are butted together so that two contact cuts become contiguous. The poly & diffⁿ outlines overlap & thin oxide under poly acts as mask in the diffⁿ process. Poly & diffused layers are butted together. The contact b/w two layers is then made by metal overlay.



Double metal MOS process rules

In this process a second metal layer is used so that V_{DD} & V_{SS} (gnd) rails in the system are distributed more flexibly on the chip. vias are used to establish connection b/w metal 2 to other layers thru metal 1.

The first level metal can be used for local distribution of power & for signal lines.

CMOS Lambda-based design rules :-

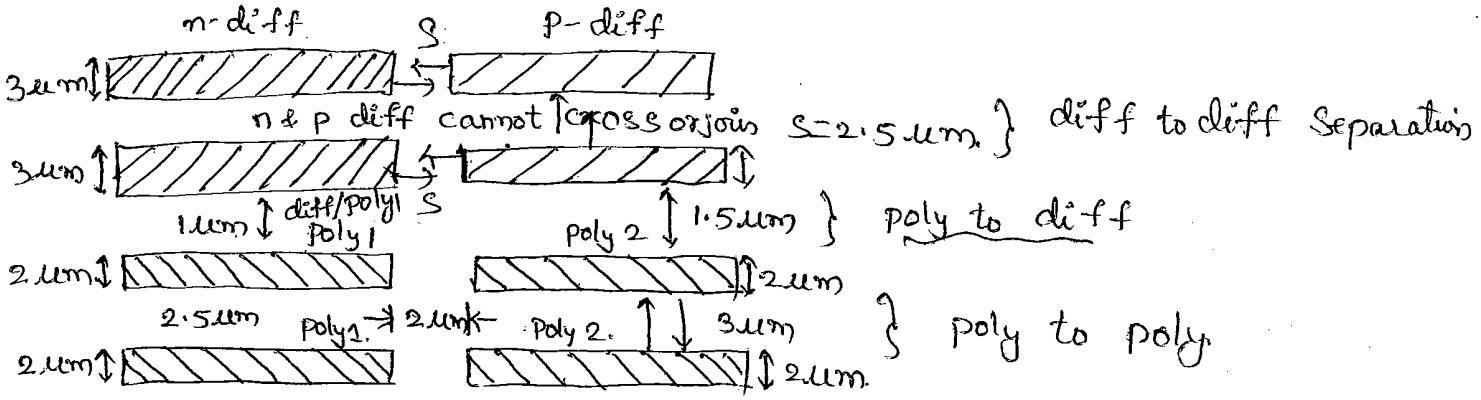
The rules of n-well (PMOS Tr), p-wires & special substrate contacts are added to the existing nMOS rules.

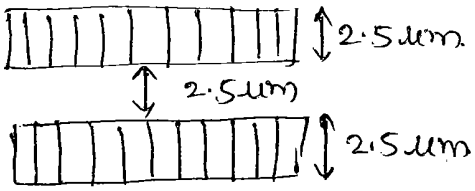
2 μm CMOS DESIGN RULES

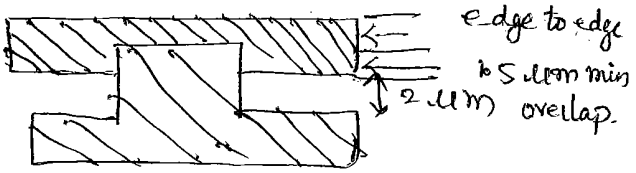
* 2 μm double metal, double poly

→ n-well : brown
Poly 1 : red
Poly 2 : Orange
n-diff : Green
P-diff : Yellow
} CMOS

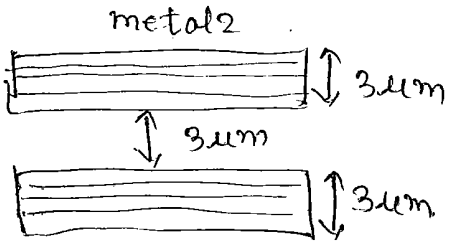
buried n+ subcoll : pale green
p-base : pink
} BiCMOS

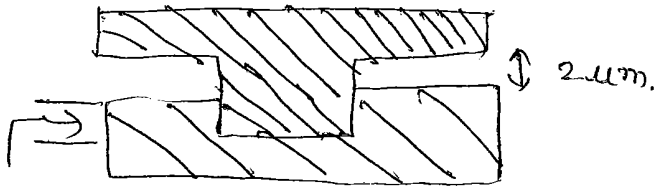






poly 2 overlapping poly 1.





1.5um min. overlap

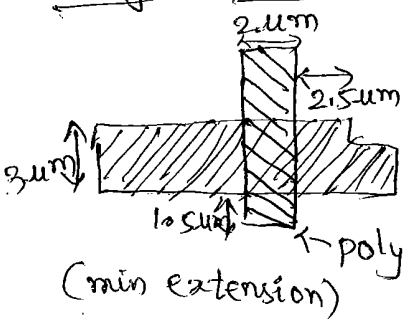
Poly 1 overlapping poly 2.

fig: Design rules for wires (2 um CMOS).

NOTE:-

For p-well CMOS, n-diff can only exist inside & p-diff wires outside p-well. For n-well CMOS, p-diff wires can only exist inside & n-diff wires outside n-well.

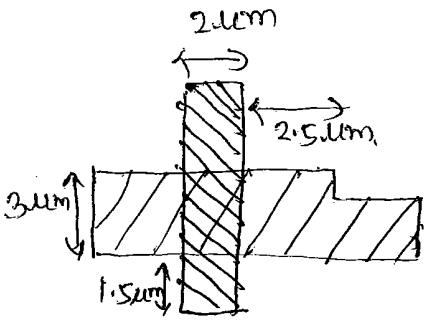
Design rules for transistors:-



min overlap of n-diff beyond gate.

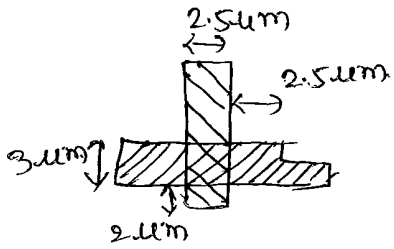
(min extension)

(i) n-type enhancement

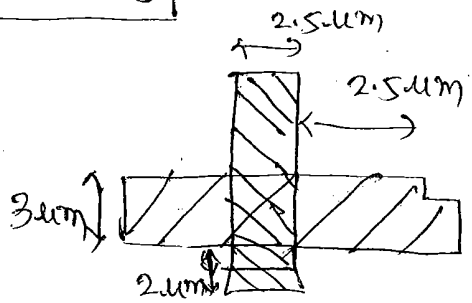


(ii) p-type Tr

fig:- Polysi Transistors.



(i) n-type Tr

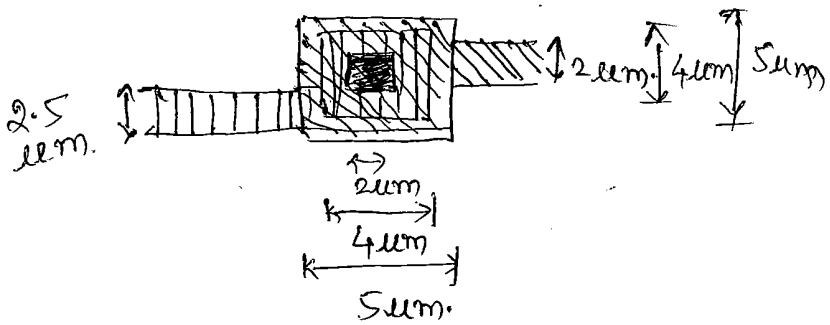


(ii) p-type Tr

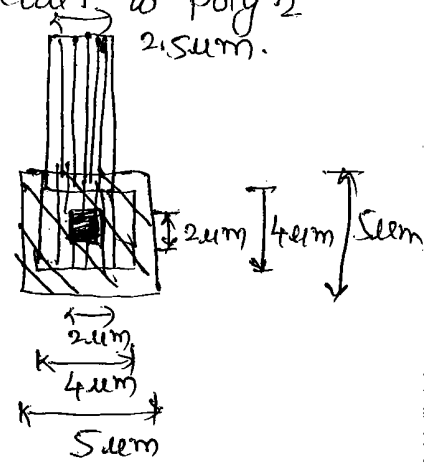
fig:- Polysi 2 Transistors.

Design rules ...

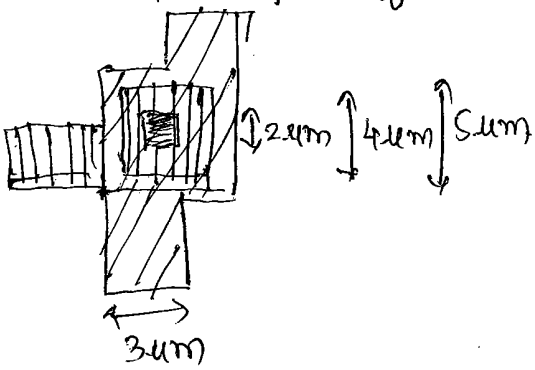
(a) Metal 1 to poly-1



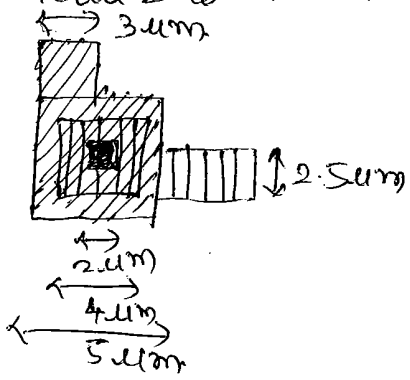
(b) Metal 1 to poly 2



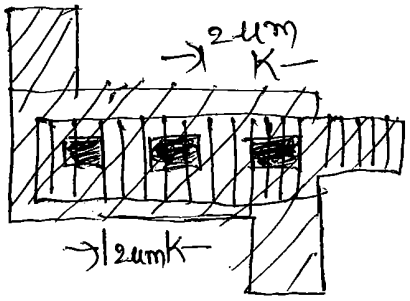
(c) Metal 1 to p diff.



(d) Metal 1 to n diff.

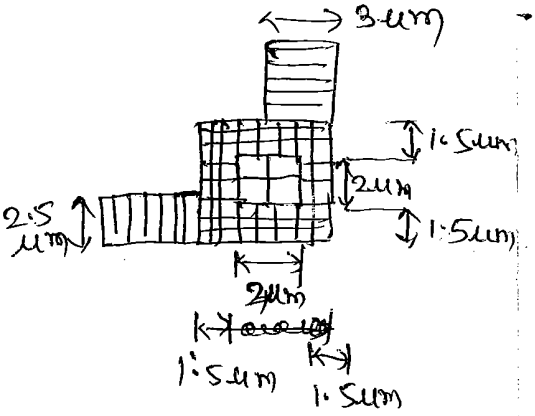


(e) Multiple contact cuts.



min spacing b/w contact cuts = 2µm

(f) Via metal 1/metal 2



Limitations of Scaling

(1) Substrate doping. $d = \sqrt{\frac{2\epsilon_{si}\epsilon_0 V}{qN_B}}$ (where $V = V_a + V_B$)

(2) Limits on miniaturization

(3) limits of interconnect & contact resistance.

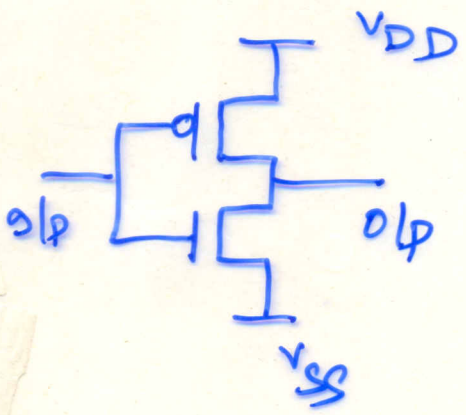
(4) Limits due to subthreshold currents.

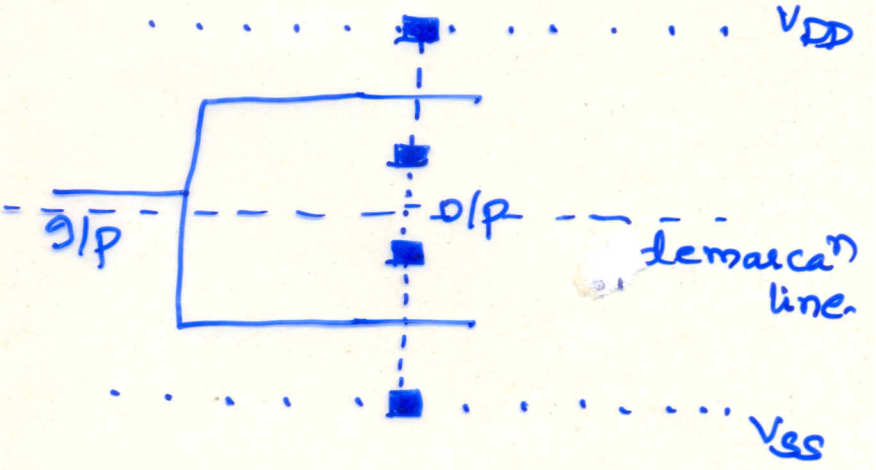
(5) Limits on logic levels & supply vtg due to noise.

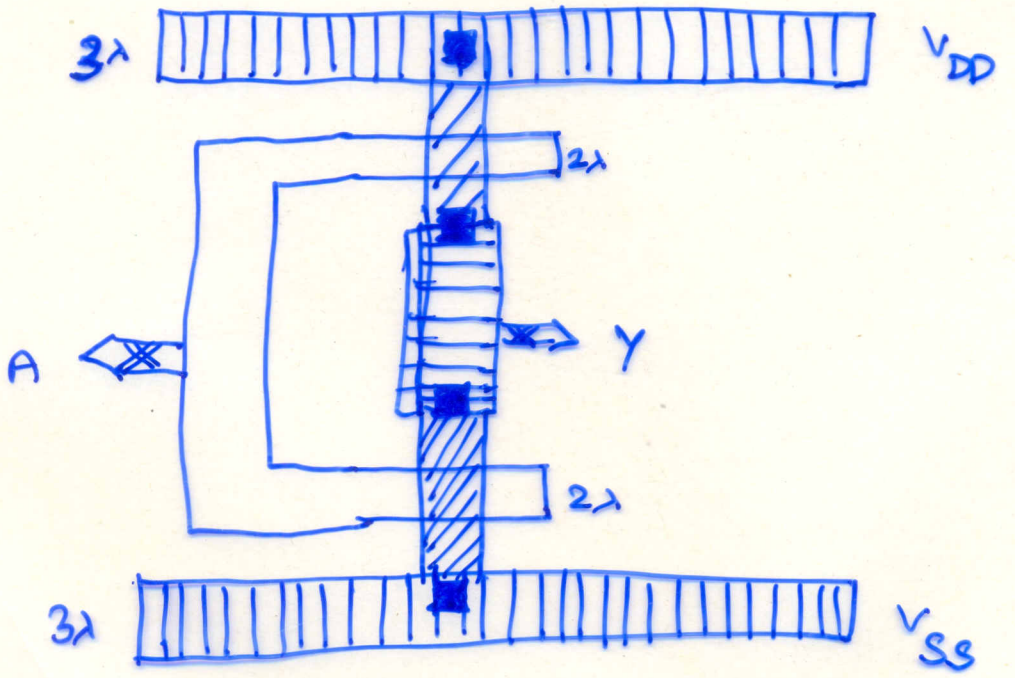
(6) Limits due to current density.
[$J = 1$ to $2 mA/\mu m^2$]

Examples:

1) NOT (INVERTER)





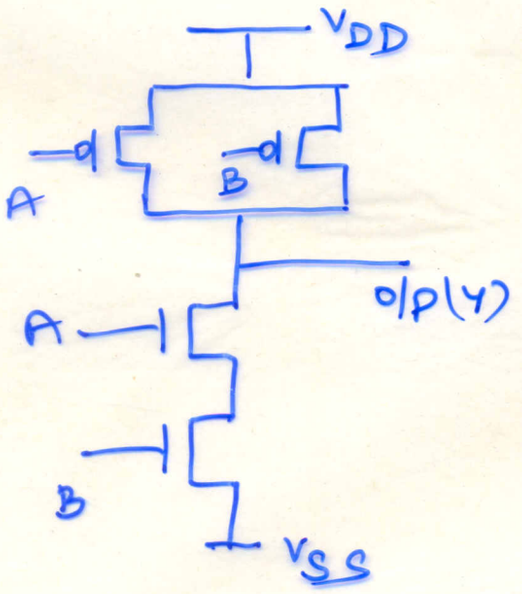


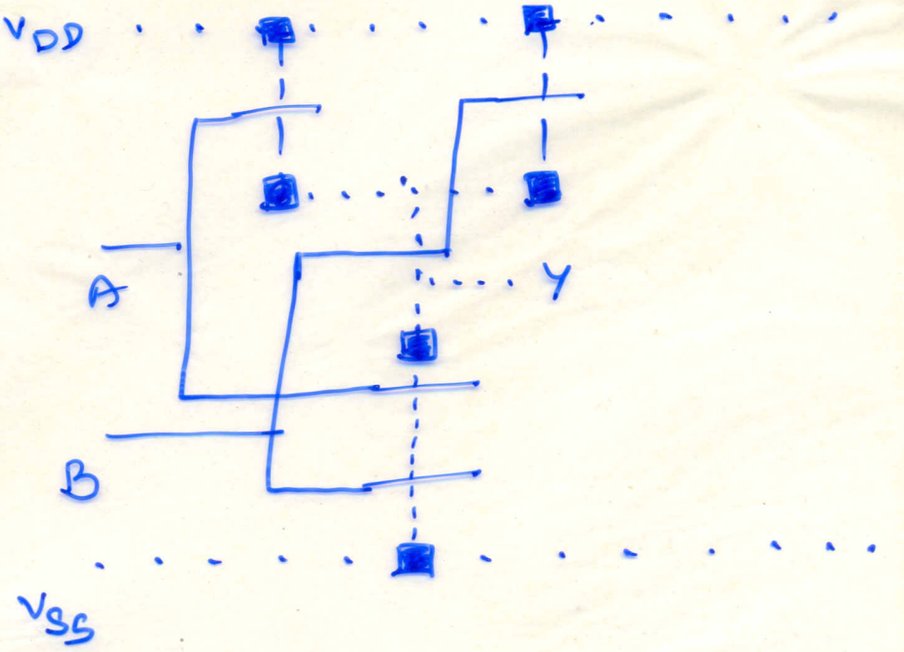
I/p: Poly
O/p: Metal
V_{DD}, V_{SS}: Metal

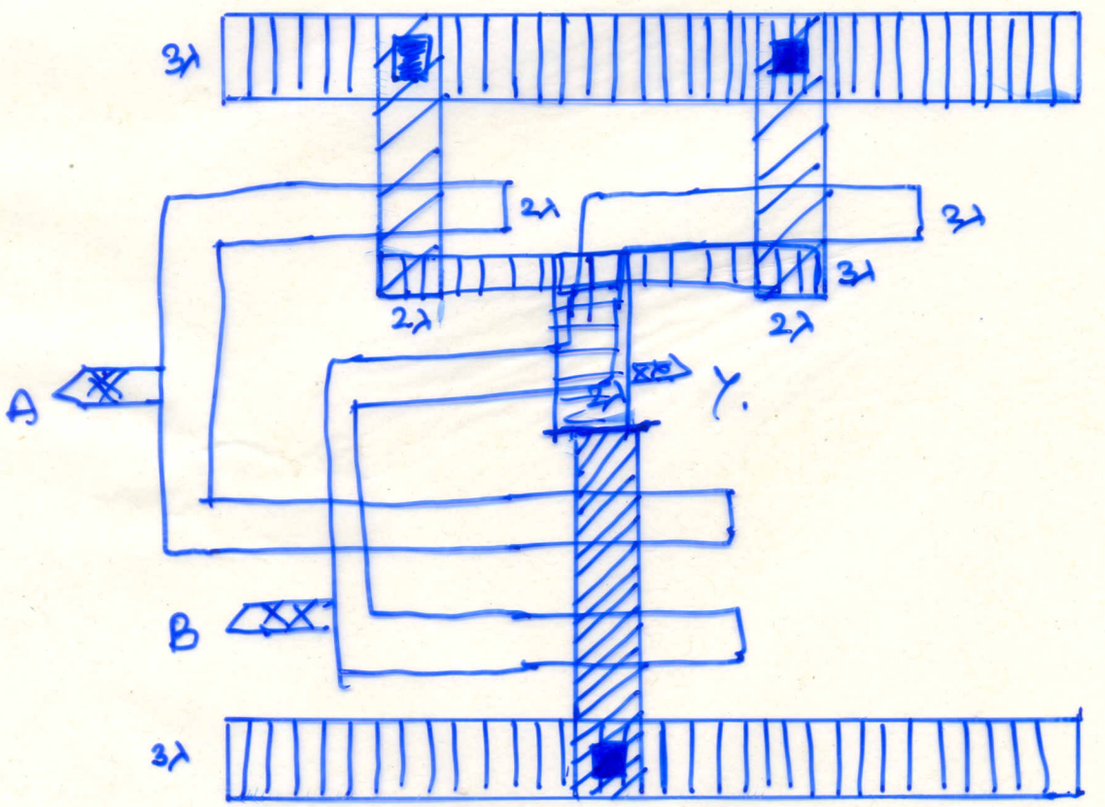
Color Codes:

Metal - Blue
Poly - Red.
n-diff - Green
P-diff - Yellow
via - Black.

2) $Y = \overline{A \cdot B}$





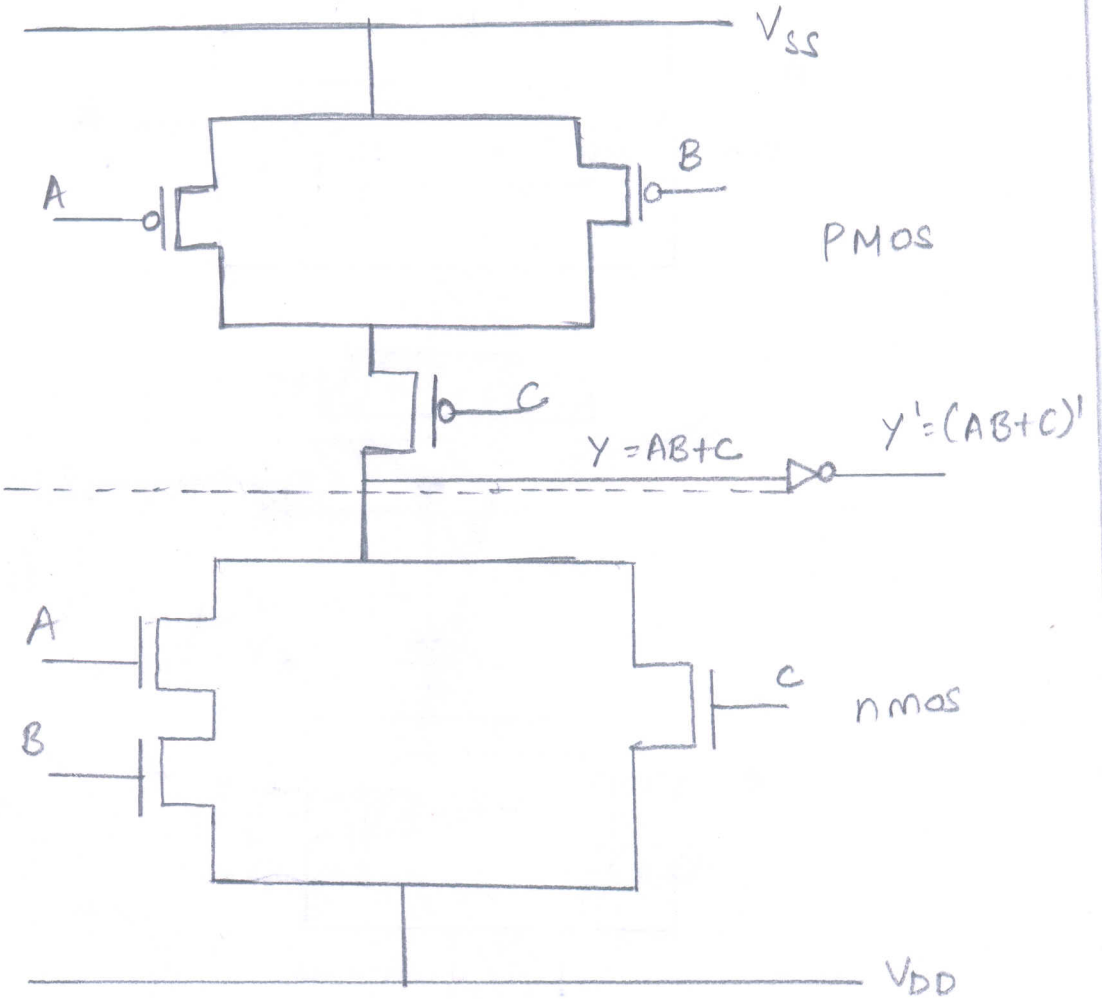


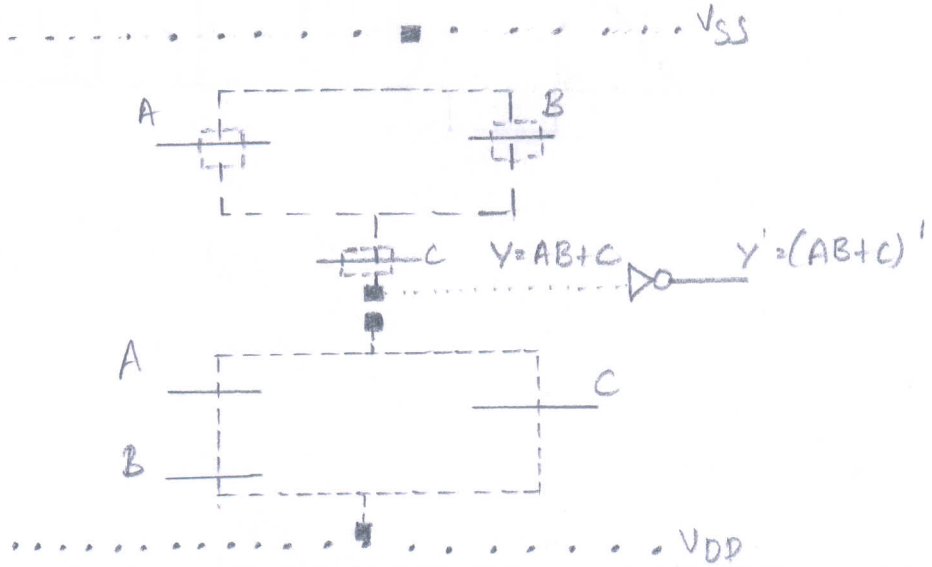
NOTE :- Via types :

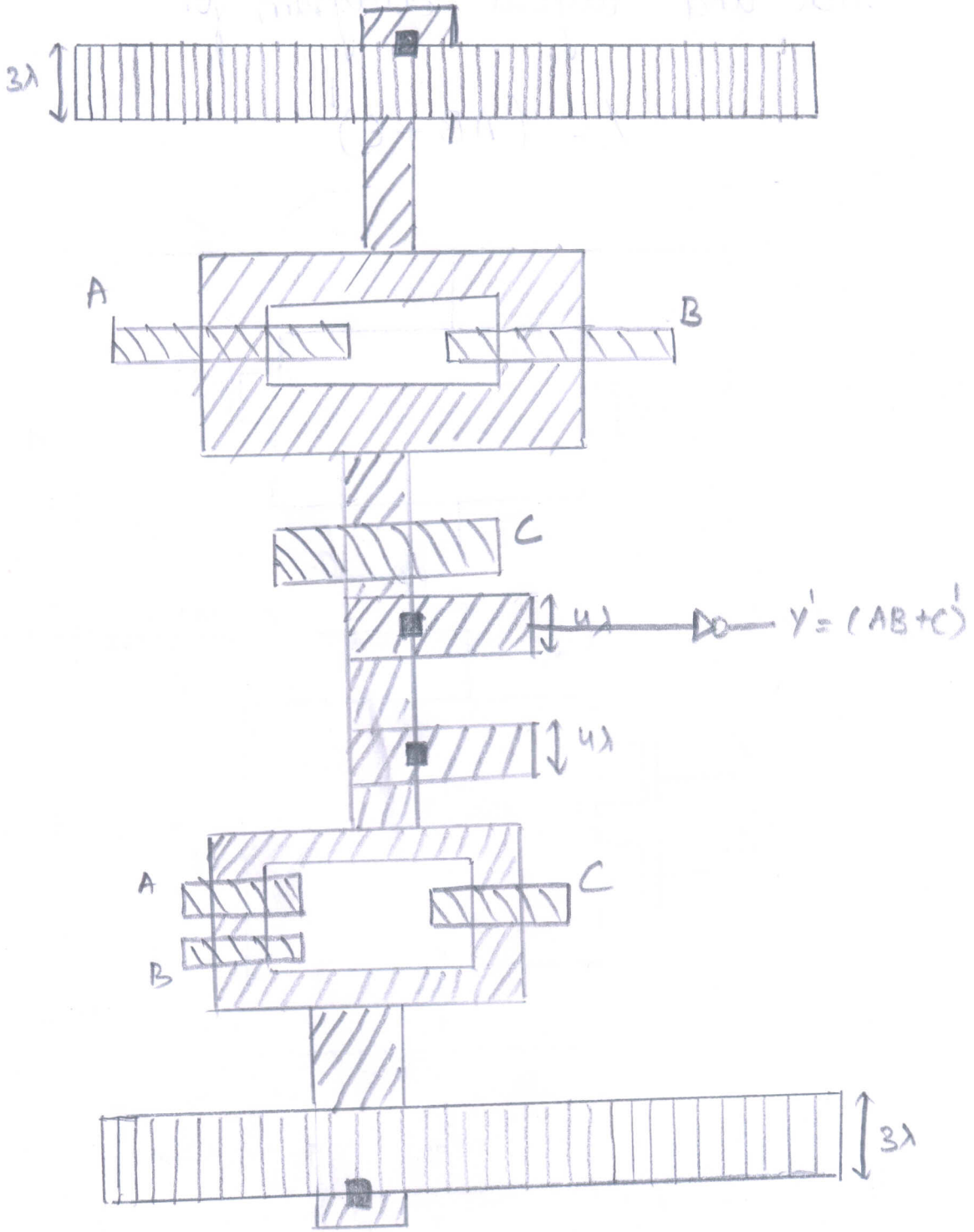
(i) n/p diff - poly
(ii) poly - metal 1
(iii) n/p diff - metal 1
(iv) metal 1 - metal 2

3) Stick and layout diagram for

$$Y = (AB + C)'$$






3λ
A
B
C
4λ
Y' = (AB+C)'
4λ
A
B
C
3λ

Scaling Factors for device Parameters:

(1) Gate Area (A_g):

$$A_g = W \cdot L$$

W & L scaled by $1/\alpha$.

$\therefore$ A_g scaled by $1/\alpha^2$

(2) Gate Cap Per Unit Area C_o or C_{ox}.

$$C_{ox} = \frac{\epsilon_{ox}}{D} = \frac{\epsilon_0 \epsilon_{ins}}{D}$$

$$\because \hat{D} = \frac{1}{\beta} D$$

$$\Rightarrow \boxed{\hat{C_{ox}} = \beta C_{ox}} \quad \text{(scaled by } \beta\text{)}$$

(3) Gate Cap C_g

$$C_g = C_{ox} \cdot W \cdot L$$

$$\hat{C_g} = \beta C_{ox} \cdot \frac{W}{\alpha} \cdot \frac{L}{\alpha}$$

$$\hat{C_g} = \frac{\beta}{\alpha^2} \cdot C_{ox} \cdot W \cdot L$$

$$\boxed{\hat{C_g} = \frac{\beta}{\alpha^2} C_g.} \quad \text{(Scaled by } \frac{\beta}{\alpha^2}\text{)}.$$

(4) Parasitic Capacitance C_x:

$$C_x \propto \frac{A_x}{d}$$

where d: depletion width around S & D

$$\hat{d} = d/\alpha.$$

A_x: Area of depletion region around src or drain

$$\hat{A_x} = A_x/\alpha^2.$$

$$\therefore \hat{C_x} = \frac{A_x}{\alpha^2} \cdot \frac{1}{d/\alpha} = \frac{A_x}{\alpha \cdot d}$$

$\therefore$ C_x scaled by $1/\alpha$

(5) Carrier Density in Channel Q_{on}:

$$Q_{on} = C_{ox} \cdot V_{gs}.$$

where Q_{on} = avg charge per unit area in ch. in 'on' state.

C_{ox} = Scaled by β

V_{gs} = " " $1/\beta$

$$\therefore \hat{Q_{on}} = \beta \cdot C_{ox} \cdot \frac{1}{\beta} \cdot V_{gs} = Q_{on}$$

$\Rightarrow$ Q_{on} is scaled by 1.

(6) Channel Resistance R_{on}:

$$R_{on} = \frac{L}{W} \cdot \frac{1}{\sigma_{on} \cdot \mu}$$

where μ = carrier mobility (const).

R_{on} scaled by $\frac{1}{\alpha} \cdot \frac{1}{1/\alpha} \cdot 1 = 1$

(7) Gate Delay T_d.

$$T_d \propto R_{on} \cdot C_g$$

$$\hat{T_d} \propto 1 \cdot R_{on} \cdot \frac{\beta}{\alpha^2} \cdot C_g$$

$$\hat{T_d} \propto \frac{\beta}{\alpha^2} \cdot T_d$$

$\therefore$ Scaled by $\frac{\beta}{\alpha^2}$

(8) Max Operating Frequency f_o:

$$f_o = \frac{W}{L} \cdot \frac{\mu C_o V_{DD}}{C_g}$$

$$f_o \propto \frac{1}{T_d}$$

$\therefore$ f_o scaled by $\frac{\alpha^2}{\beta}$

(9) Saturation Current I_{dss}.

$$I_{dss} = \frac{C_o \mu}{2} \cdot \frac{W}{L} (v_{gs} - v_t)^2$$

$\because$ v_{gs} & v_t scaled by $1/\beta$ & C_o by β

$\Rightarrow$ I_{dss} scaled by $\beta \left(\frac{1}{\beta}\right)^2 = 1/\beta$

(10) Current Density J:

$$J = \frac{I_{dss}}{A}$$

'A' (area) of ch scaled by $1/\alpha^2$

$$\hat{J} = \frac{I_{dss}/\beta}{A/\alpha^2} = \frac{\alpha^2}{\beta} \cdot \frac{I_{dss}}{A} = \frac{\alpha^2}{\beta} J$$

$\therefore$ J is scaled by $\frac{\alpha^2}{\beta}$.

(11) Switching Egy Per Gate E_g.

$$E_g = \frac{C_g}{2} (V_{DD})^2$$

E_g scaled by $\frac{\beta}{\alpha^2} \cdot \frac{1}{\beta^2} = \frac{1}{\alpha^2 \beta}$