

Digital CMOS Integrated Circuits

By Dr. Bharat Choudhary

Curriculum structure of B.Tech ECE Programme

(Department of Electronics & Communication Engineering)

SYLLABUS OF B. Tech. (ECE)

Course Code : ECT-304
LTP: 3-0-1 per week

Syllabus:

Introduction to MOSFETs technology: Construction and working of MOSFET, Current-Voltage Characteristics, Performance metrics for digital design, Scaling of MOSFETs, Fabrication flow of CMOS n-well process. [05h].

CMOS Inverter: Design , analysis of NMOS inverter (resistive, enhancement and depletion load) , CMOS inverters; transfer characteristics, Noise margins, , rationing of transistor size, logic voltage levels, rise and fall of delays, Propagation Delay, Power Consumption. [8h].

Combinational Circuits: Design of basic gates in NMOS technology; CMOS logic design styles: static CMOS logic (NAND, NOR gates), complex gates, Pass Transistor logic, Transmission gate, Dynamic MOS design: pseudo NMOS logic, clocked CMOS (C2 MOS) logic, domino logic, NORA, Half and Full adder, Multiplexer, XOR, XNOR [10h].

Logical Effort: Logical effort of different digital circuit design: Input capacitance, Logical and Electrical effort, parasitic delay, Single stage and Multistage with and without branch network. Design of minimum delay and optimization of best stages. [6h]

Layout and stick diagram: Layout design rules: Lambda and micron based design rules- stick diagram, Layer properties of various conducting layers in MOS and CMOS technology (diffusion, poly-silicon and metal), Layout design of different CMOS circuit, area estimation. [6h]

Sequential and Memory Design: Sequential MOS Logic and Memory Design: Static latches; Flip flops & Register. [5h]

Project: Introduction of open source tools: EDA. The class project is to design reasonably complex CMOS circuit. The project will be performed as a team of three or four students.

Course Outcomes:

Co1- Understanding of different types of modulation and demodulation techniques for digital communication

Co2- To learn the ISI and equalization techniques.

Co3-To analyse different types of channel coding schemes.

Co4-Understanding the performance of different digital communication systems

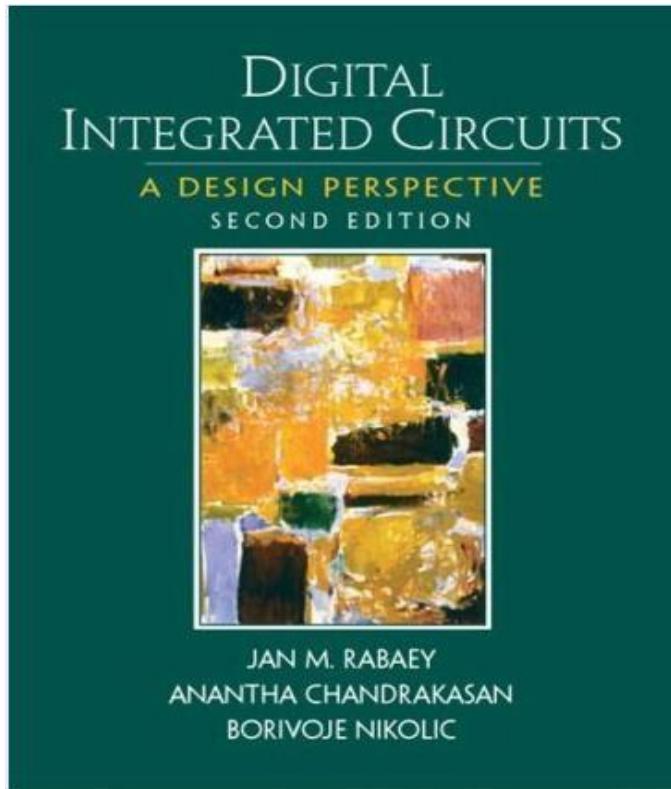
References:

1. Sung-Mo Kang & Yusuf Leblebici, CMOS Digital Integrated Circuits Analysis and Design, Second Edition, McGraw-Hill, 1999.
2. Rabaey, Chandrakasan and Milokic. Digital system design- A design perspective. Pearson education, India.
3. Neil H.E.Weste and Kamran Eshraghian, Principles of CMOS VLSI Design, A System Perspective, Pearson Education, India.
4. Ken Martin, Digital Integrated Circuits, Oxford Press.
4. CMOS Circuit Design, Layout and simulation: J. Baker, D.E. Boyce., IEEE press.

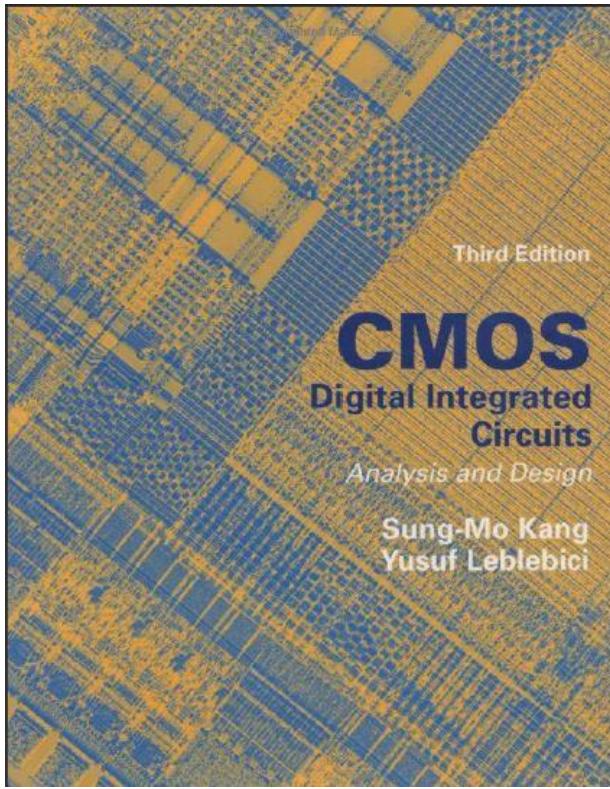
Course Outcome

- Analyze NMOS & CMOS inverter circuits
- Analyze and Design combinational and sequential circuits using MOS devices.
- Describe the process of fabrication of NMOS and CMOS devices
- Draw the stick diagrams and layouts for different MOS circuits.
- Describe various issues involved in subsystem design and estimate performance parameters.
- Analyze the impact of interconnects on the circuit performance

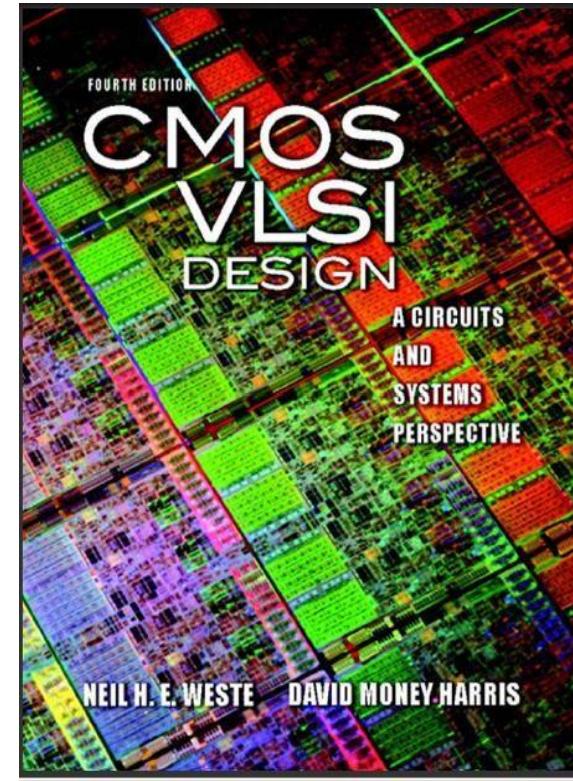
Reference Books



**Digital Integrated Circuits – A Design Perspective”,
2nd ed. by J. Rabaey, A. Chandrakasan, B. Nikolic**

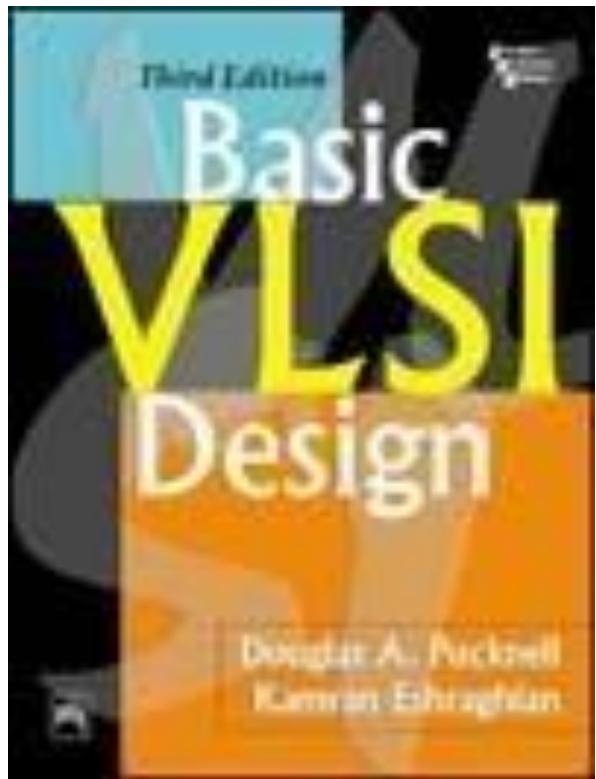


**“CMOS DIGITAL INTEGRATED CIRCUITS:Analysis and Design”
3rd Ed. McGraw-Hill.
By:SUNG-MO (STEVE) KANG**

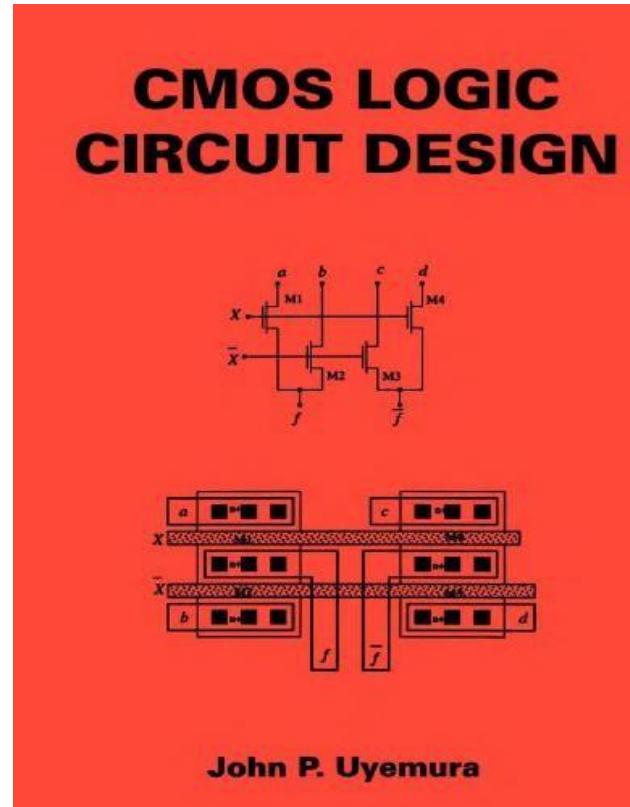


“CMOS VLSI Design:A Circuits and Systems Perspective” 4th ed. Addison-Wesley By: Neil H. E. Weste, David Money Harris

Reference Books



“Basic VLSI Design”, 3rd Ed., PHI
By: PUCKNELL DOUGLAS
A.ESHRAGHIAN,
KAMRAN



“CMOS LOGIC CIRCUIT DESIGN”
KLUWER ACADEMIC PUBLISHERS
By: John P. Uyemura

Introduction

What is expected ?

Complexity → ∞

Intel 4004 Processor: 2300 Transistor → 10µm Technology

Intel i7 Processor : 3,200,000,000 Transistors → 14nm Technology

Power → Min

Intel 4004 Processor: 1 W

Intel i7 Processor : 47W

Delay → Min

Intel 4004 Processor: 740KHz

Intel i7 Processor : 3.6 GHz

Cost → Min

Intel 4004 Processor: \$60

Intel i7 Processor : \$366

Size → Min

Intel 4004 Processor: 12 mm²

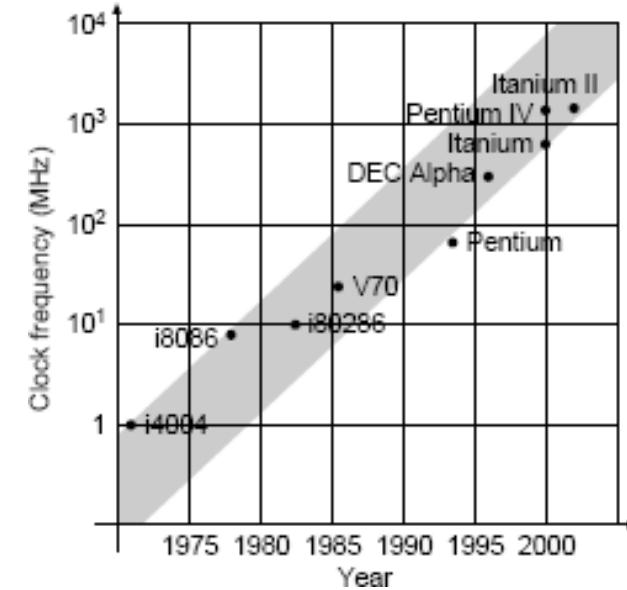
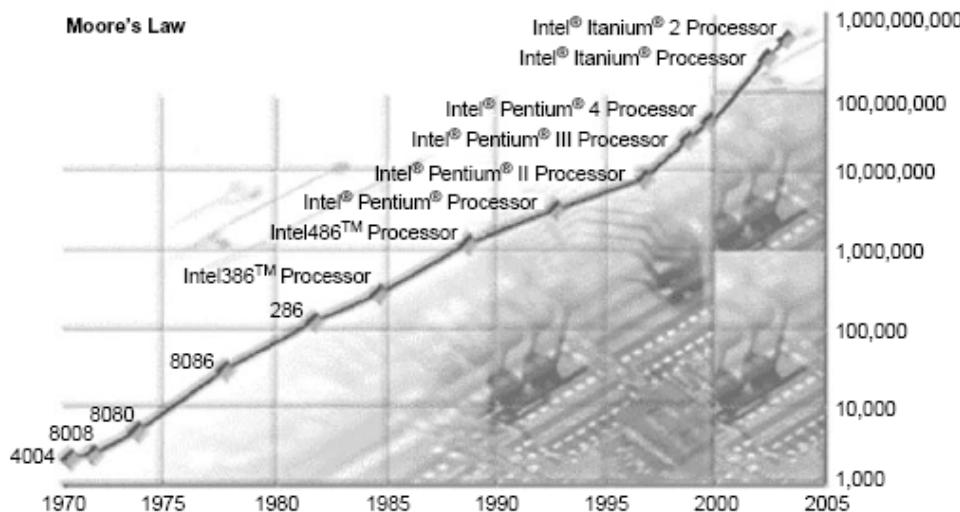
Intel i7 Processor : 246 mm²

TECHNOLOGY GENERATIONS

Integration level	Year	No. of transistors
SSI	1950s	Less than 10^2
MSI	1960s	$10^2 \approx 10^3$
LSI	1970s	$10^3 \approx 10^5$
VLSI	1980s	$10^5 \approx 10^7$
ULSI	1990s	$10^7 \approx 10^9$
SLSI	2000s	Over 10^9

Moore's law

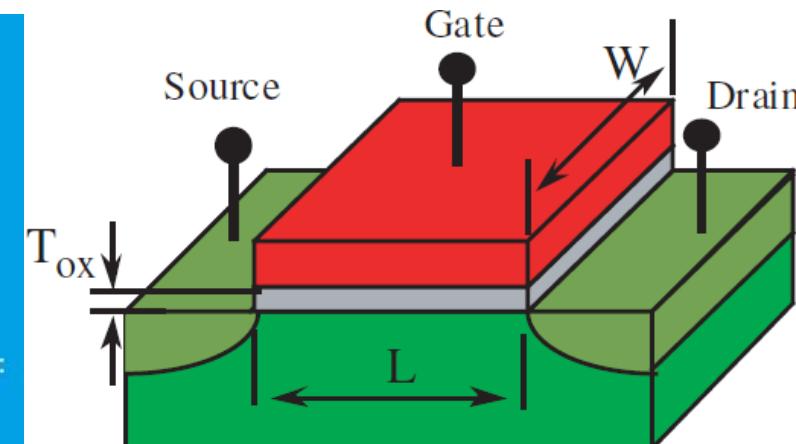
The evolution of MOS technology has followed the famous Moore's law that predicts a steady decrease in gate length. As predicted by Gordon Moore in the 1960s, integrated circuit (IC) **densities have been doubling** approximately **every 18 months**, and this doubling in size has been accompanied by a similar exponential increase in circuit speed (or more precisely, clock frequency).



On-chip transistor count increase for the Intel processors (Source: Intel).

Moore's Law and our expectation

1	2	3	4	5
1971	1972	1974	1978	1982
Intel® 4004 processor	Intel® 8008 processor	Intel® 8080 processor	Intel® 8086 processor	Intel® 286™ processor
Initial clock speed: 108KHz	Initial clock speed: 800KHz	Initial clock speed: 2MHz	Initial clock speed: 5MHz	Initial clock speed: 6MHz
Transistors: 2,300	Transistors: 3,500	Transistors: 4,500	Transistors: 29,000	Transistors: 134,000
Manufacturing technology: 10 micron	Manufacturing technology: 10 micron	Manufacturing technology: 6 micron	Manufacturing technology: 3 micron	Manufacturing technology: 1.5 micron



Assume initial $L=1, W=1$ and $T_{ox}=1$

$$W=0.7, L=0.7, T_{ox}=0.7$$

=> Lateral and vertical dimensions reduce 30 %

$$\text{Area Cap} = C = \frac{0.7 \times 0.7}{0.7} = 0.7$$

- **Area reduced by 50%**
- **Capacitance reduced by 30%**

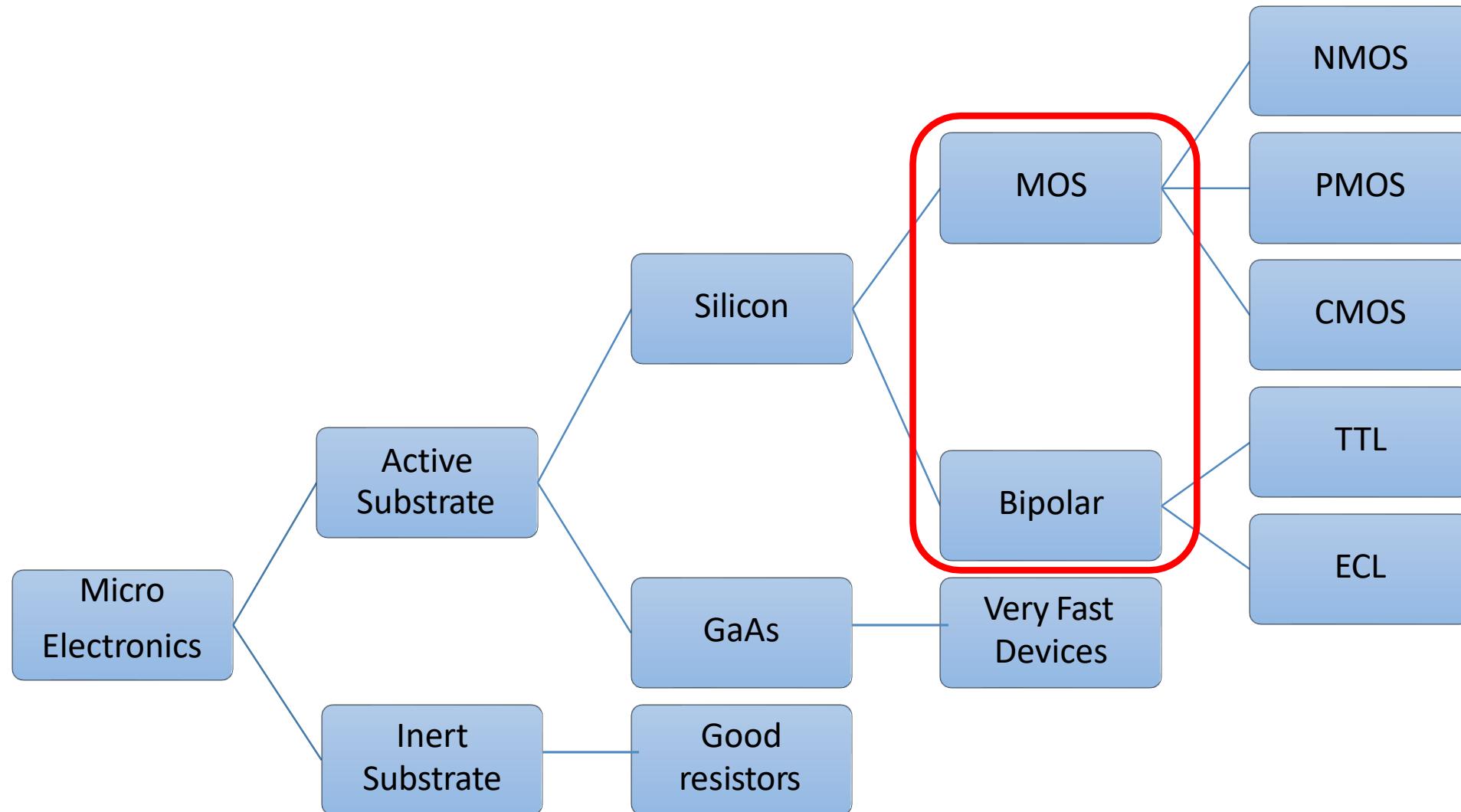
Assume initial $V_{dd}=1, V_t=1$ and $T_{ox}=1$

$$V_{dd}=0.7, V_t=0.7, T_{ox}=0.7,$$

$$T = \frac{C \times V_{dd}}{I} = 0.7, \text{Power} = CV^2f = \frac{0.7 \times 0.7^2}{0.7} = 0.7^2$$

=> **Delay reduces by 30 % and Power reduces by 50 %**

Microelectronics Technology

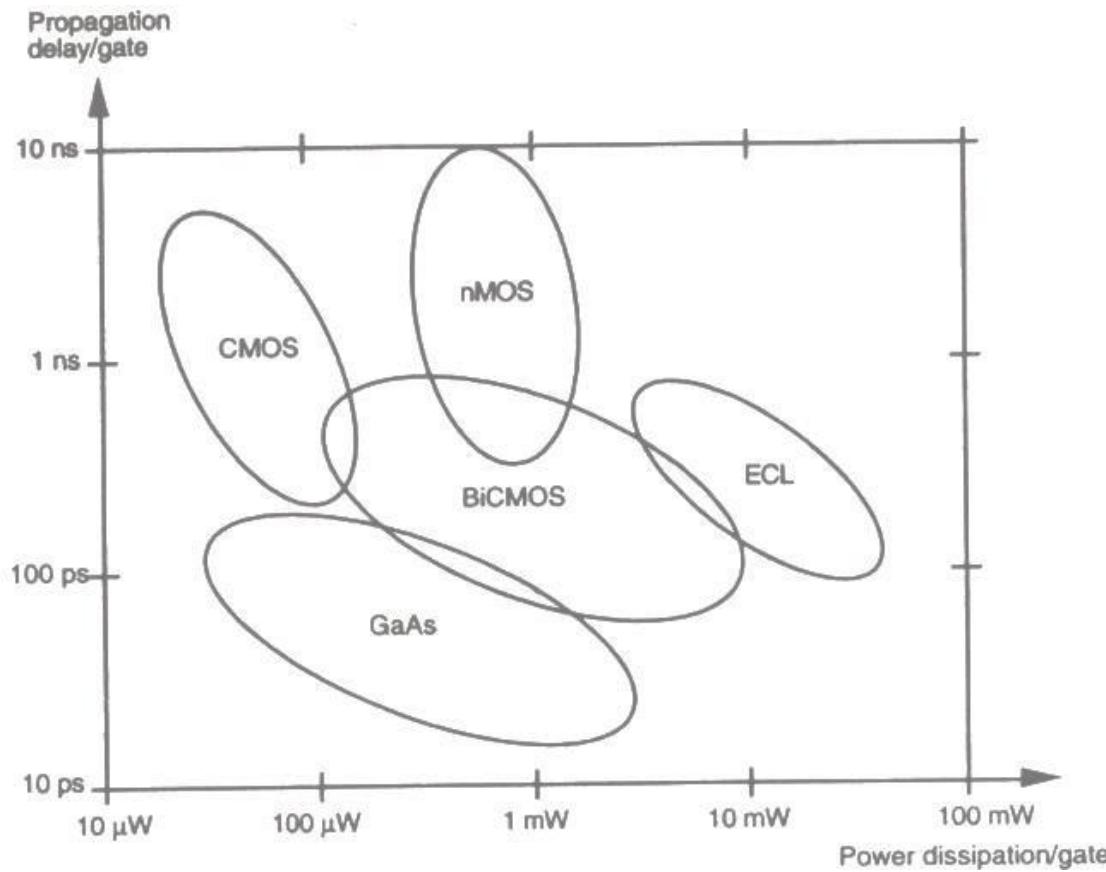


MOS Vs. BJT

Factors	CMOS	Bipolar
Static Power Dissipation	Low	High
Input Impedance	High	Low
Noise Margin	High	Low
Packing Density	High	Low
Fan-out	Low	High
Direction	Bidirectional	Unidirectional

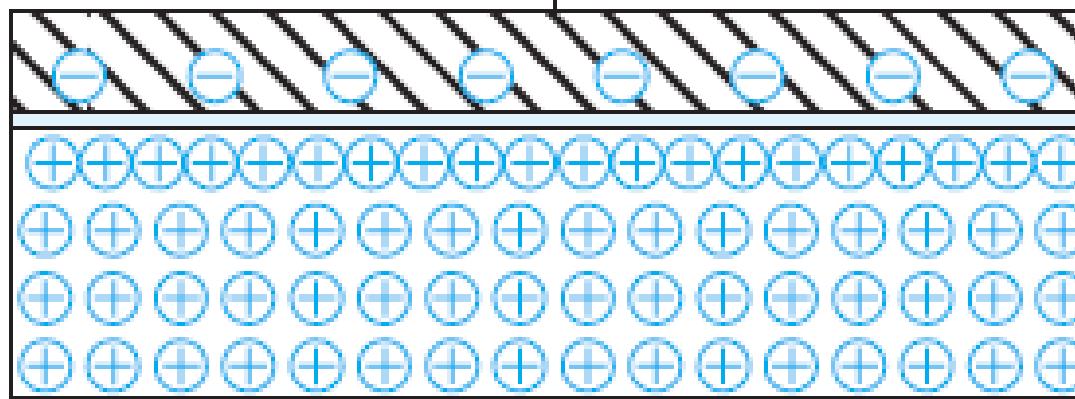
CMOS is superior!

Power Dissipation Vs. Delay



CMOS offers low powers dissipation with large delay

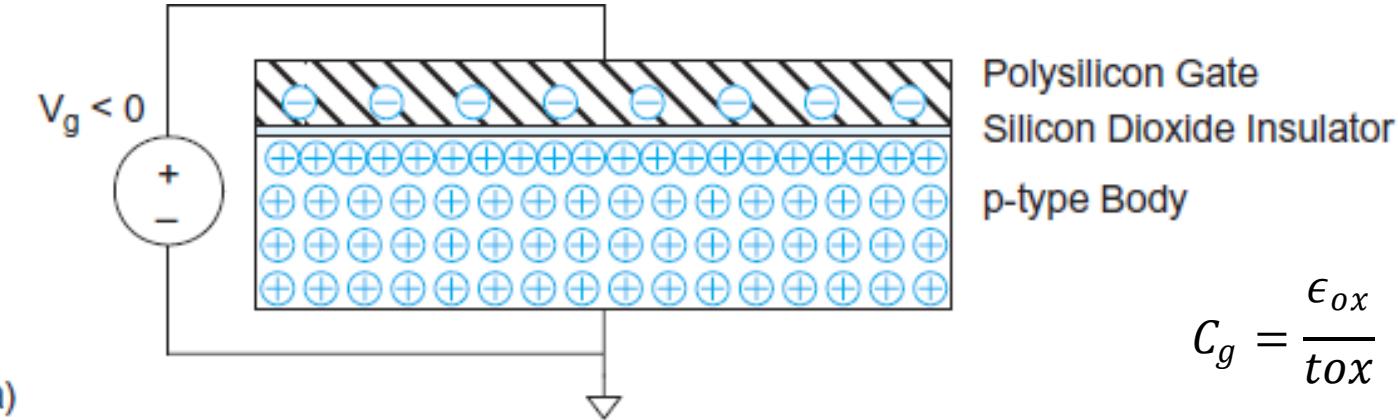
MOS Capacitor



Oxide thickness, Threshold voltage, and Doping levels, depend on the fabrication process, and cannot be changed by design; they are technology parameters.

MOS Capacitor

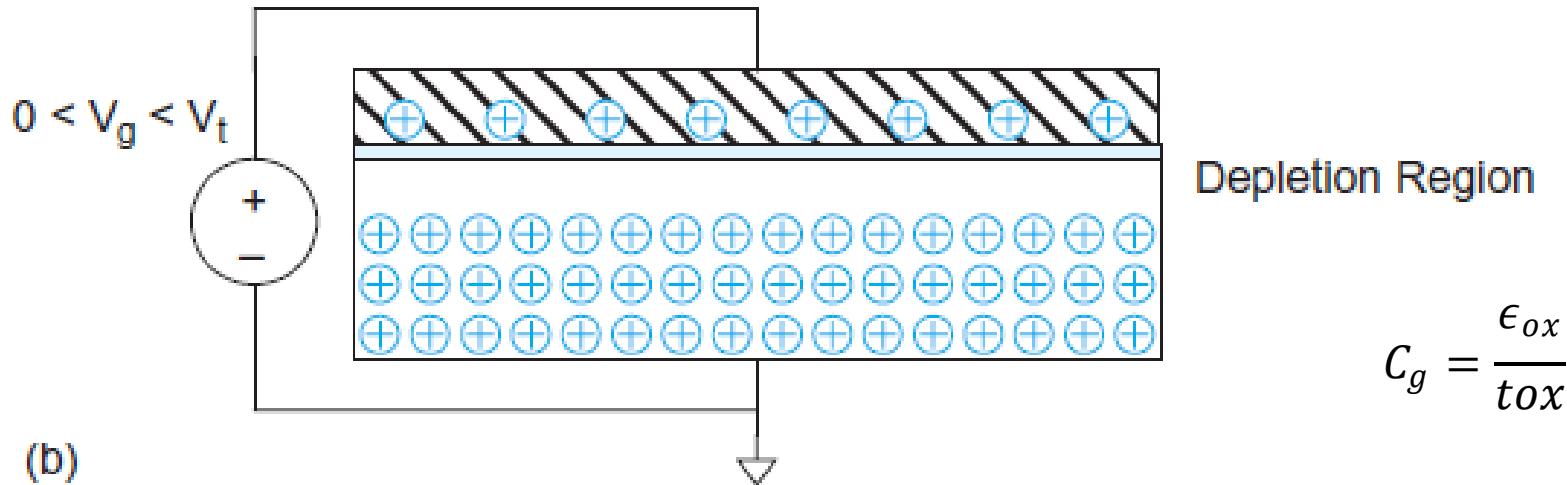
i. Accumulation



A negative voltage is applied to the gate, so there is negative charge on the gate. The mobile positively charged holes are attracted to the region beneath the gate. This is called the *accumulation* mode

MOS Capacitor

ii. Depletion

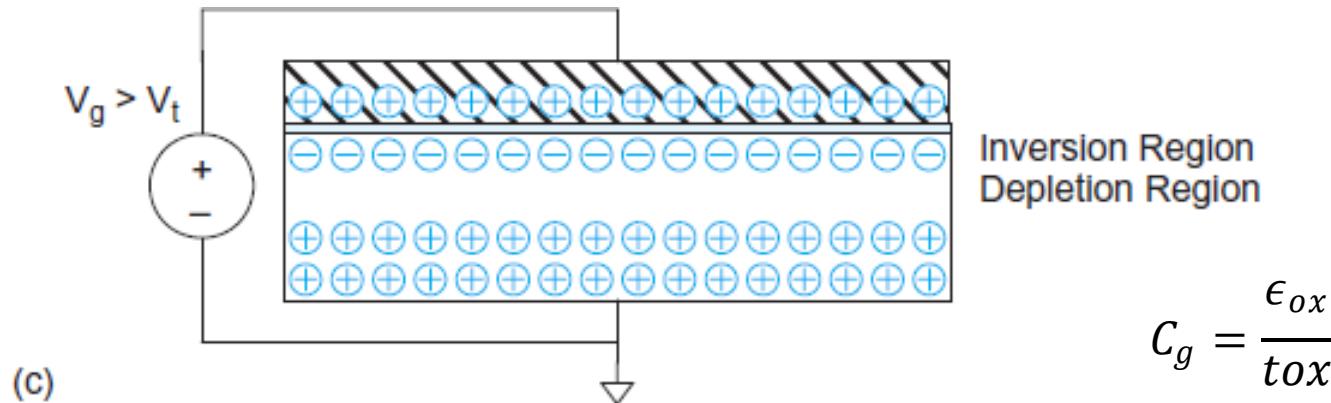


$$C_g = \frac{\epsilon_{ox}}{t_{ox}}$$

A small positive voltage is applied to the gate, resulting in some positive charge on the gate. The holes in the body are repelled from the region directly beneath the gate, resulting in a *depletion* region forming below the gate.

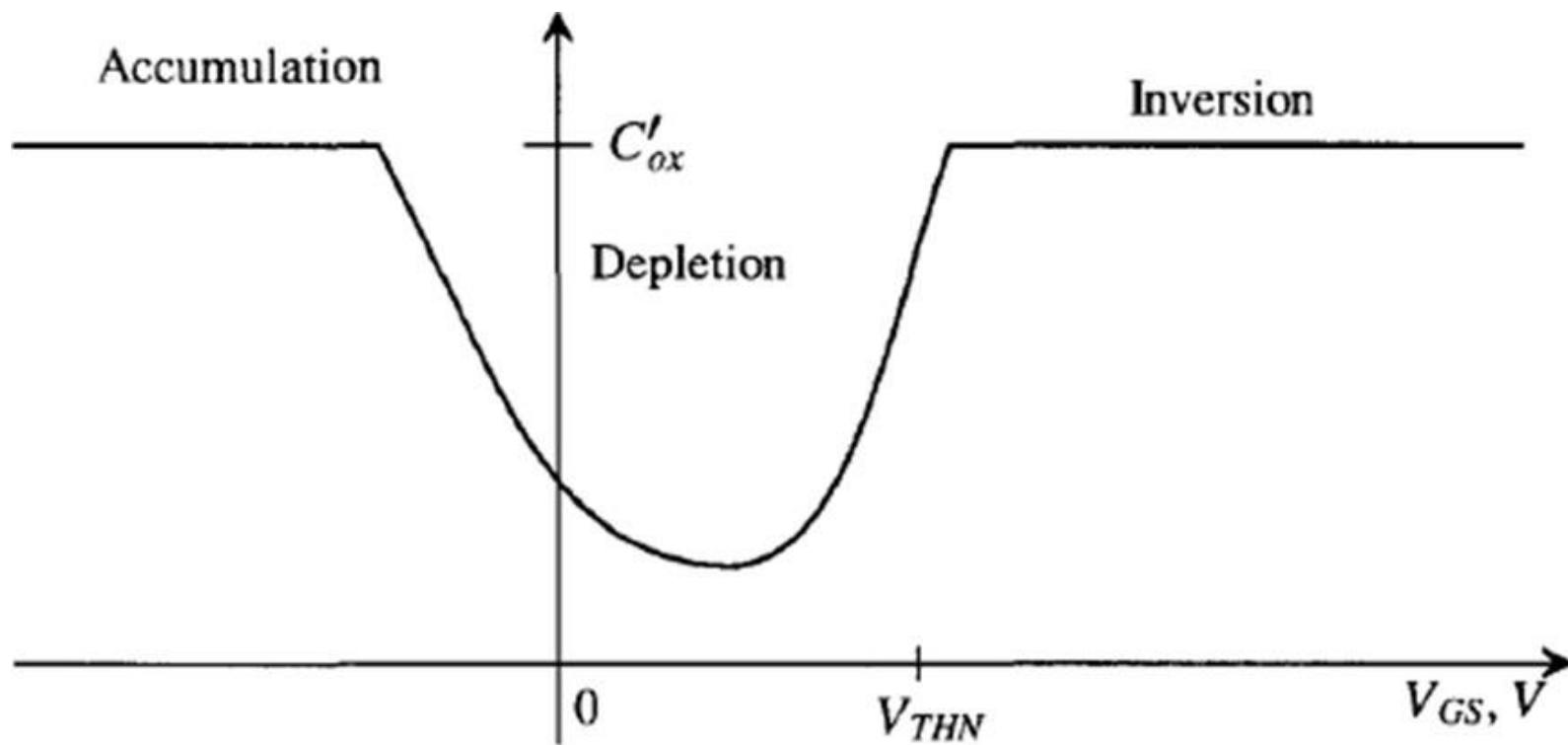
MOS Capacitor

iii. inversion



A higher positive potential exceeding a critical threshold voltage V_t is applied, attracting more positive charge to the gate. The holes are repelled further and some free electrons in the body are attracted to the region beneath the gate. This conductive layer of electrons in the p-type body is called the *inversion* layer

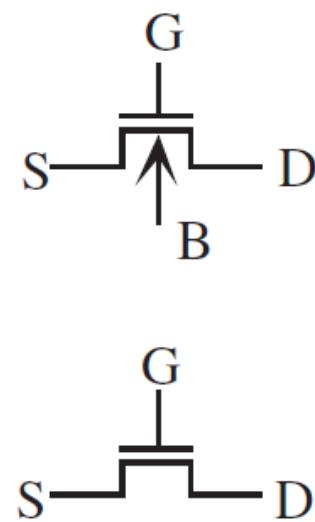
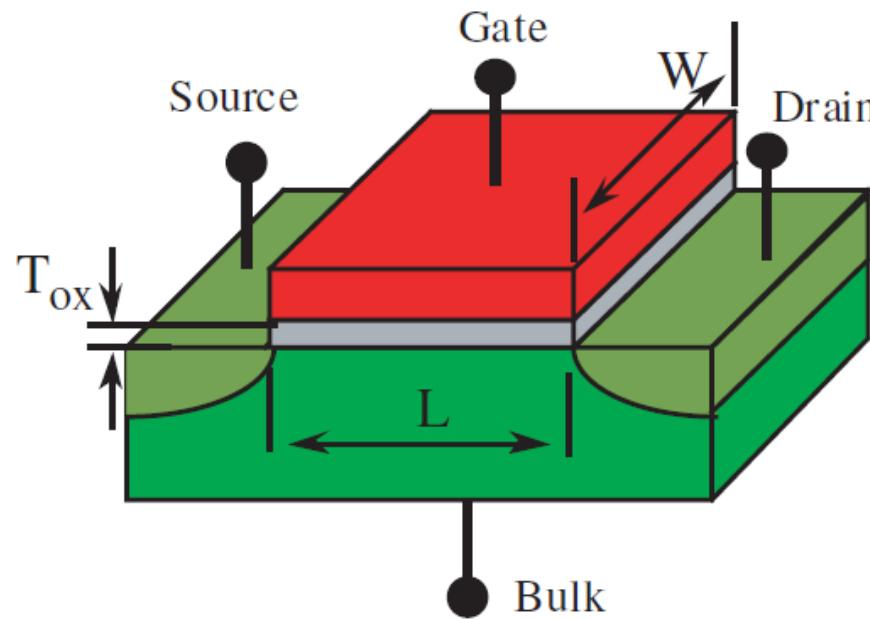
MOS Capacitor



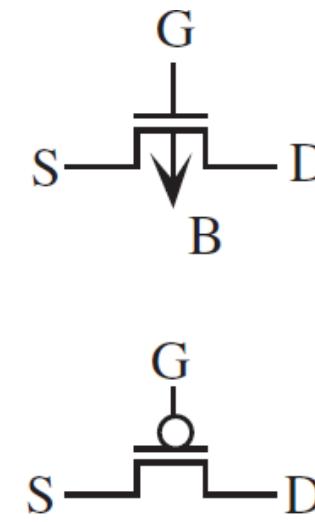
MOSFET

- Enhancement type MOSFET
 1. N-Channel MOSFET
 2. P-Channel MOSFET
- Depletion type MOSFET
 1. N-Channel MOSFET
 2. P-Channel MOSFET

Enhancement MOSFET



nMOS

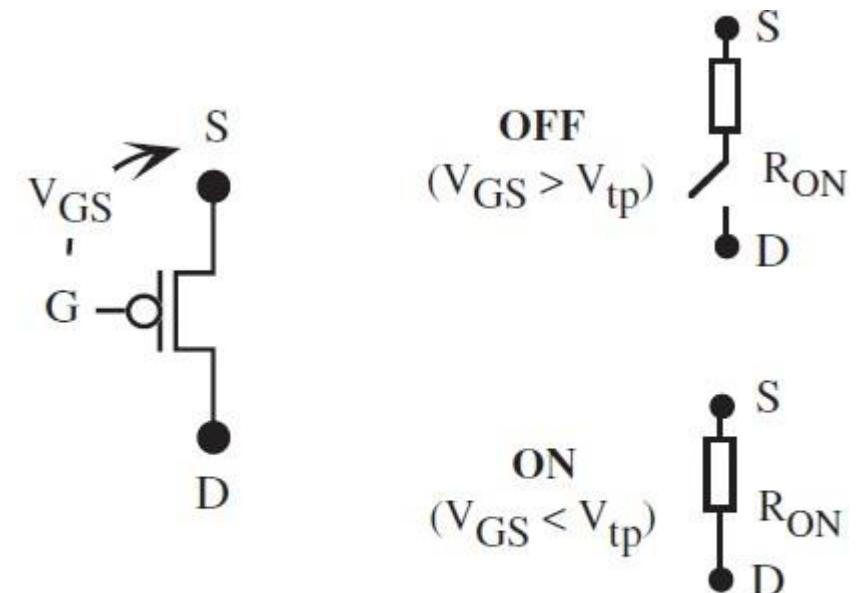
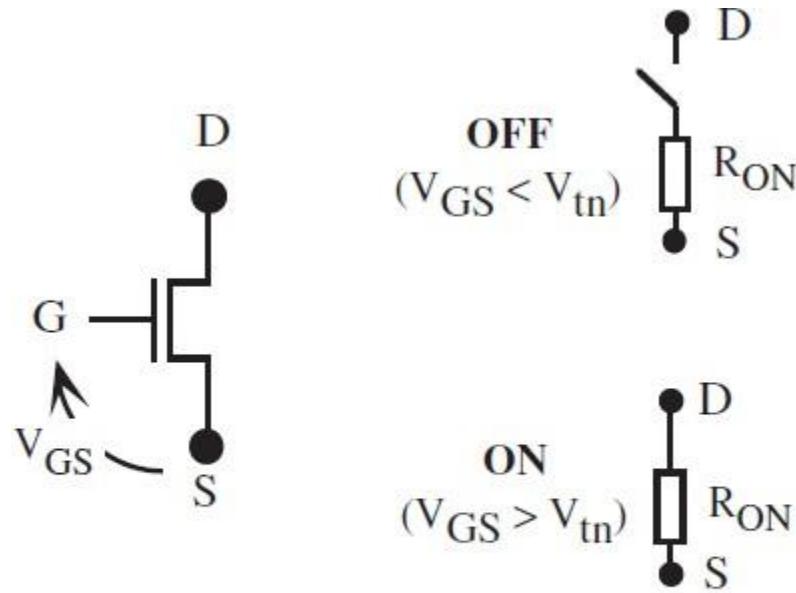


pMOS

$T_{ox} = 15\text{\AA}$ to 100\AA (Diameter of SiO_2 molecule is about 3.2\AA)

Arrows always point from P to N, so an NMOS (N-channel in P-well or P-substrate) has the arrow pointing in (from the bulk to the channel)

Enhancement MOSFET



V_{Th} is fixed for NMOS and PMOS devices for given fabrication process

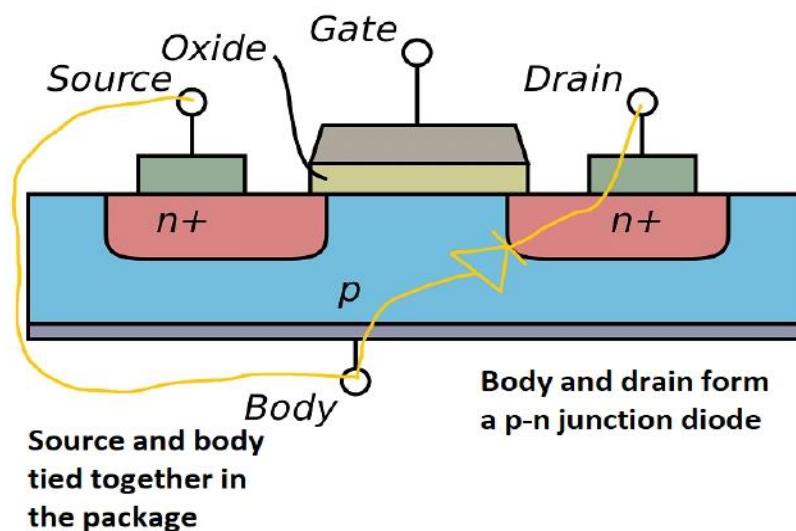
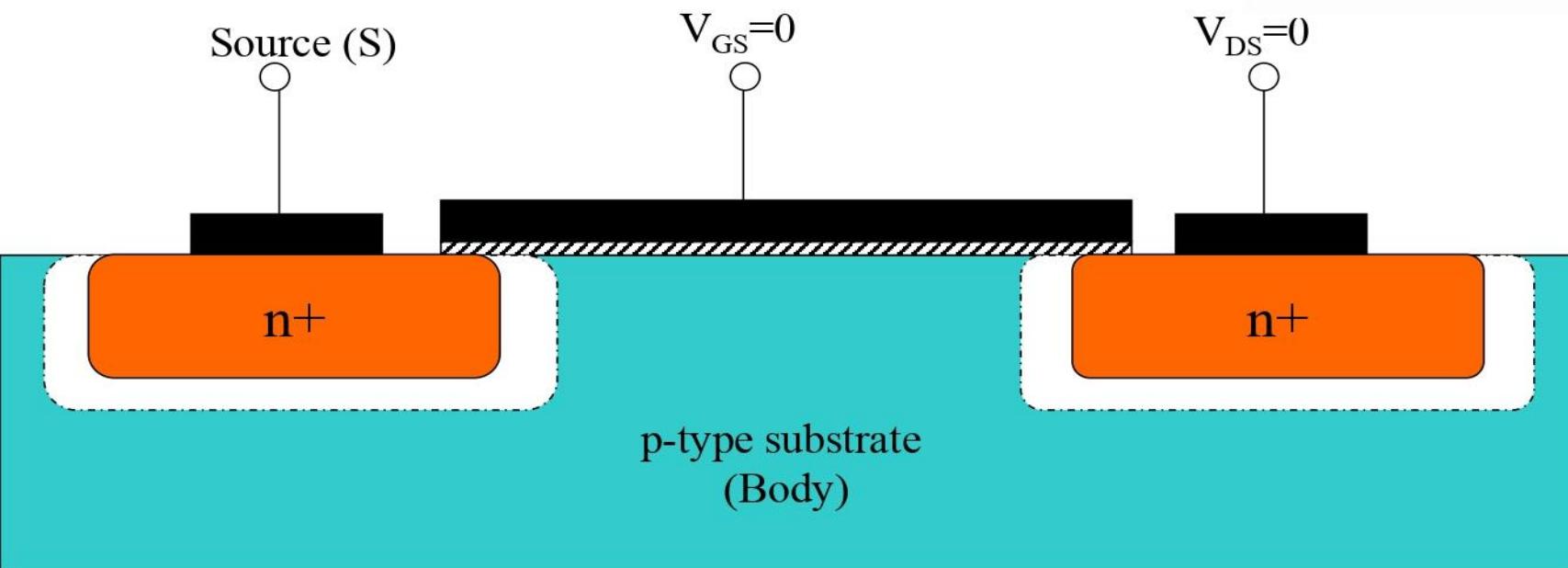
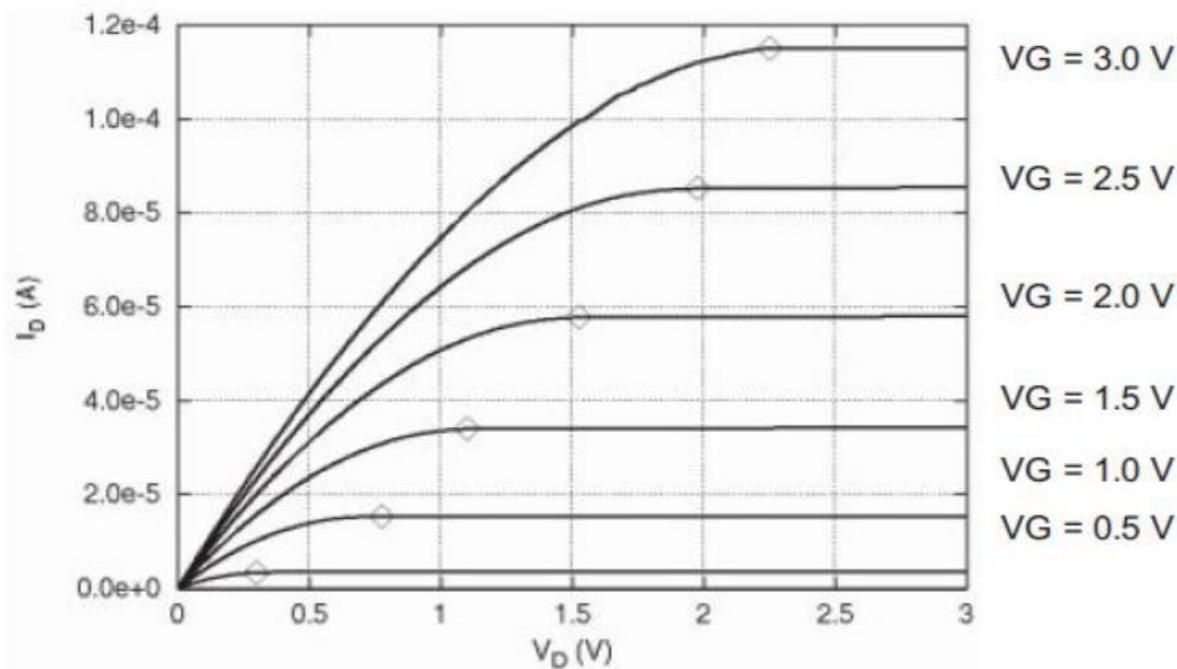
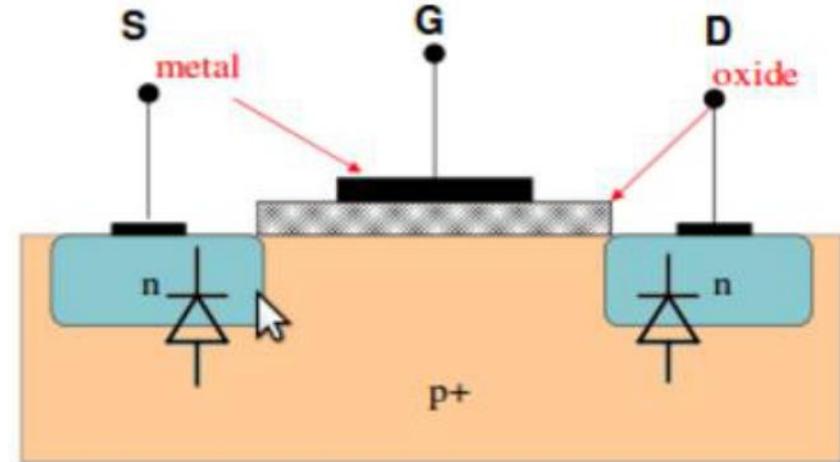
Enhancement NMOS Transistor

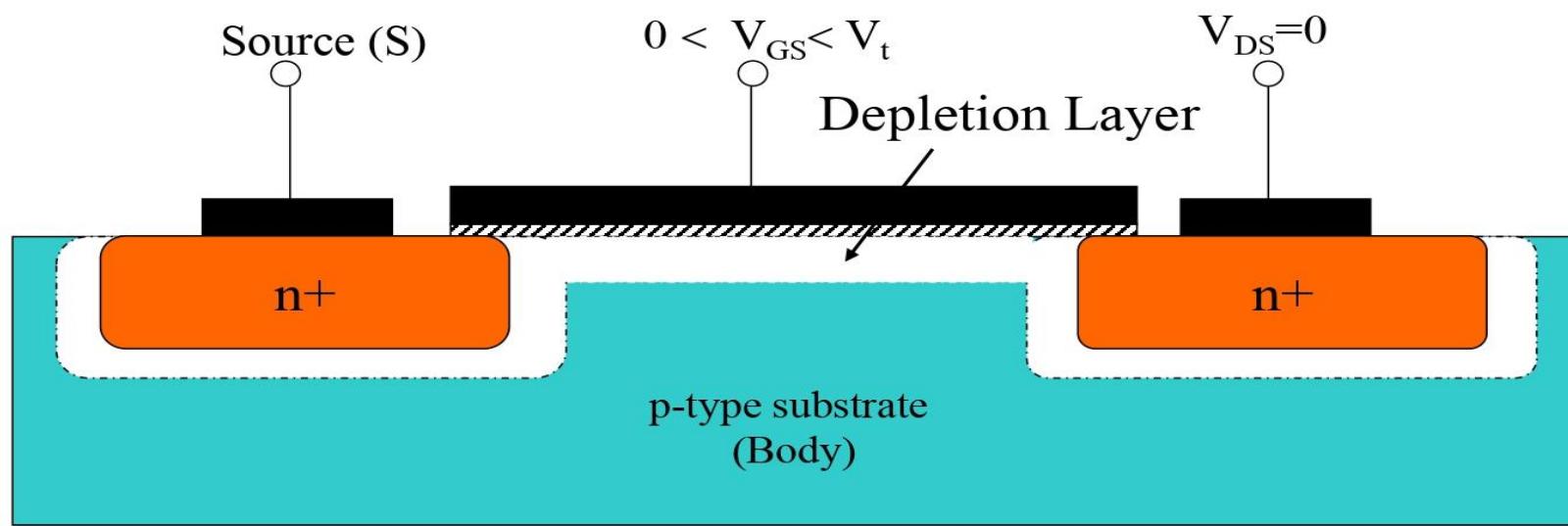
Operation Of N-MOS Transistor

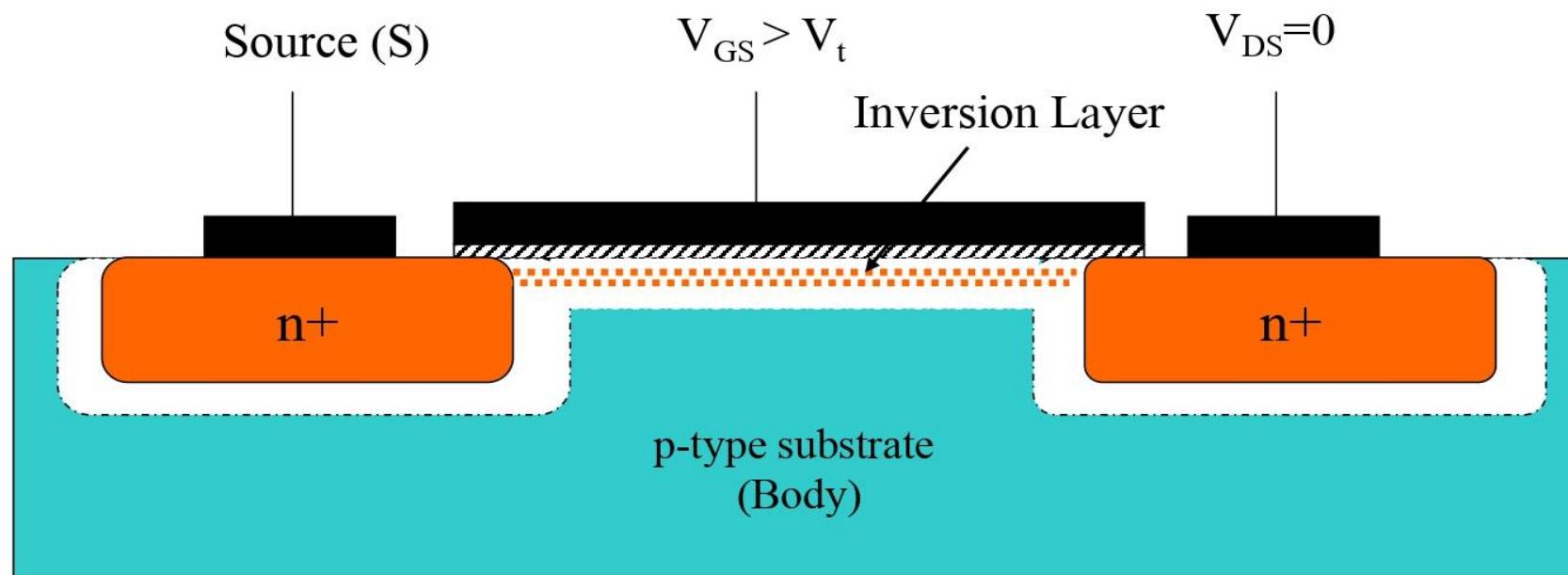
Depending on the relative voltages of the source, drain and gate, the NMOS transistor may operate in any of three regions viz :

- Cut off : Current flow is essentially zero (also called accumulation region)
- Linear : (Non saturated region)-It is weak inversion region drain current depends on gate and drain voltage.
- Saturation : It is strong inversion region where drain current is independent of drain-source voltage.

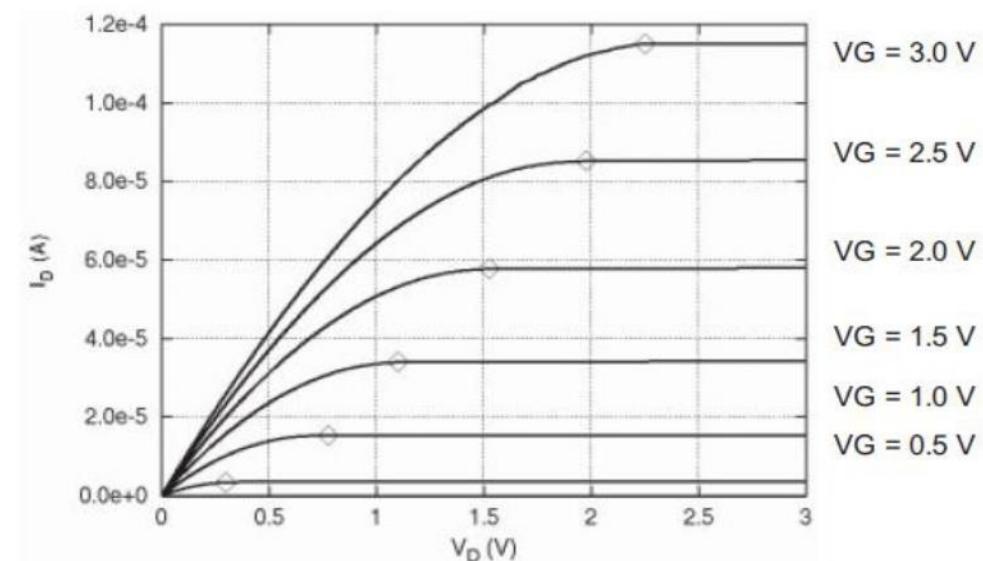
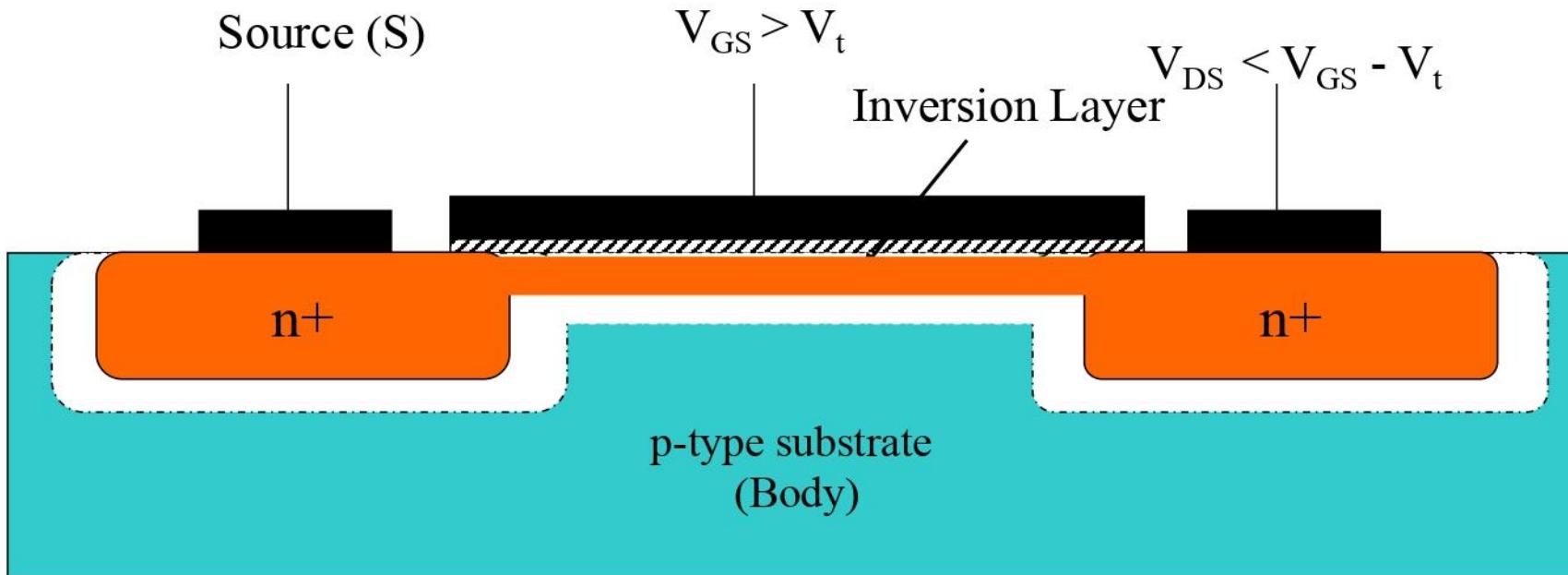
Cut-off Region

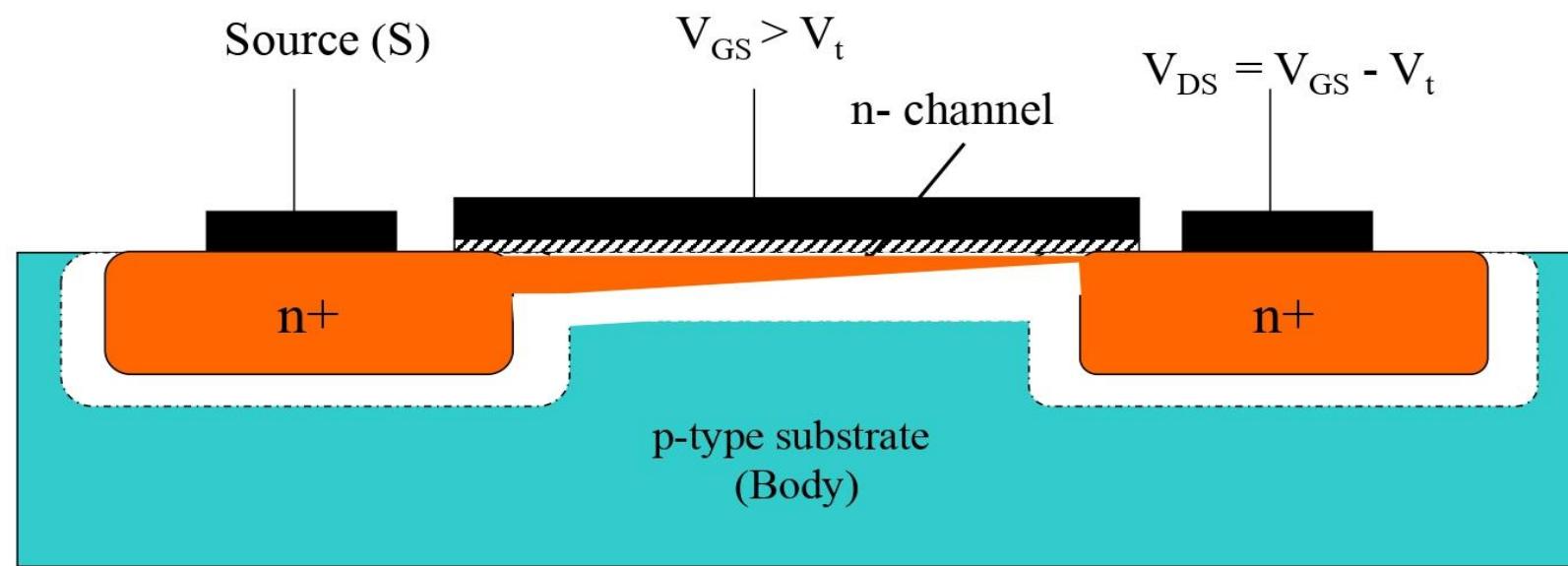




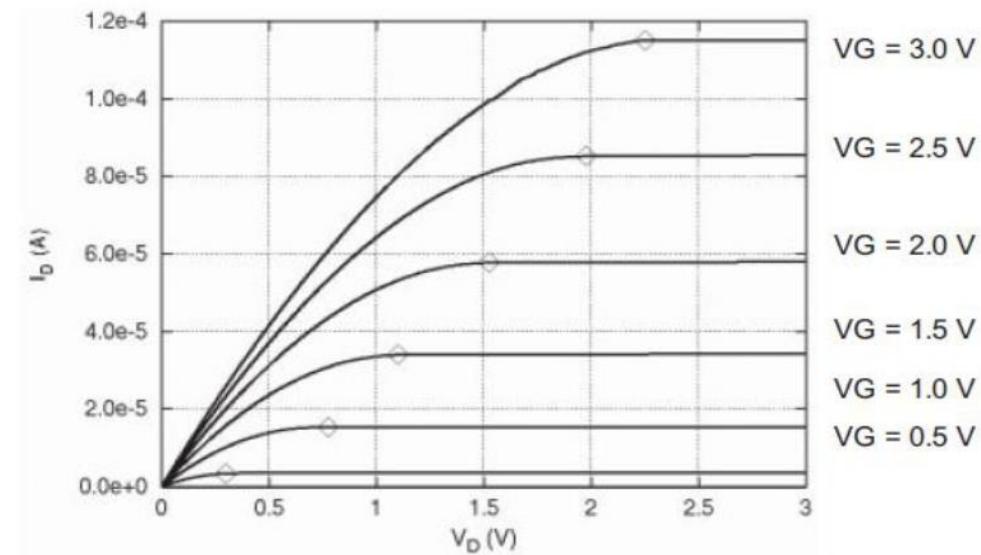
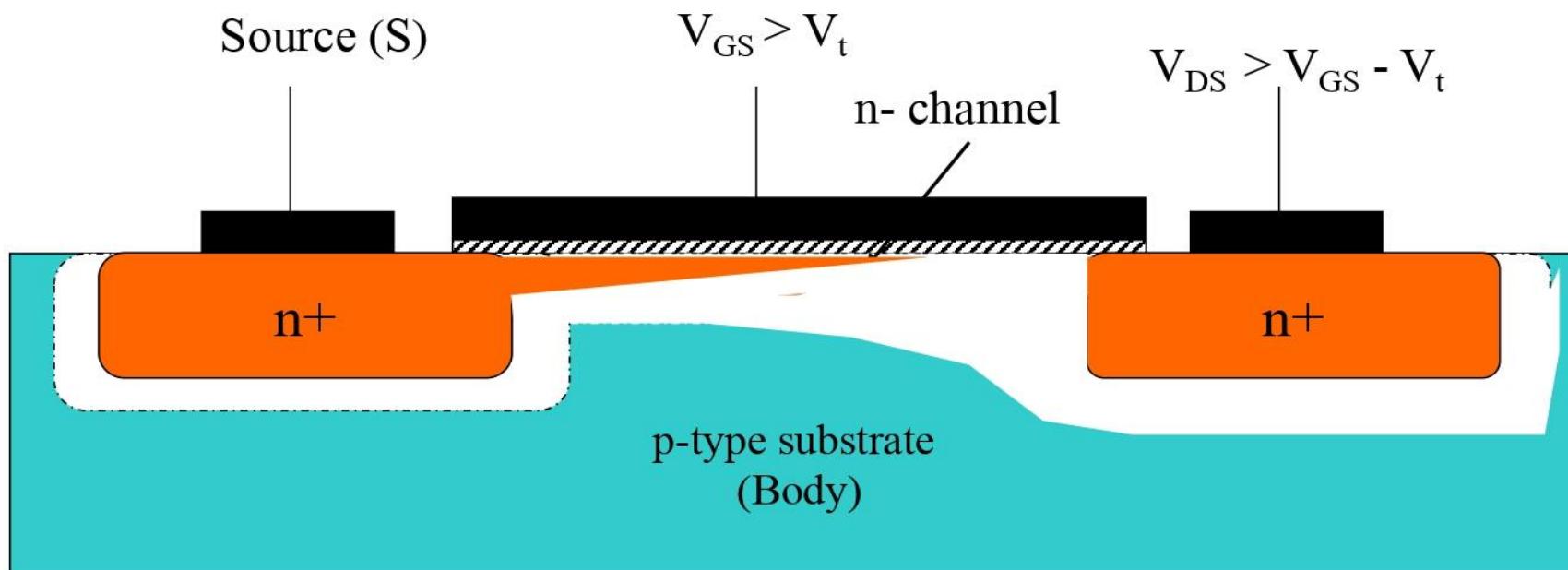


Linear region





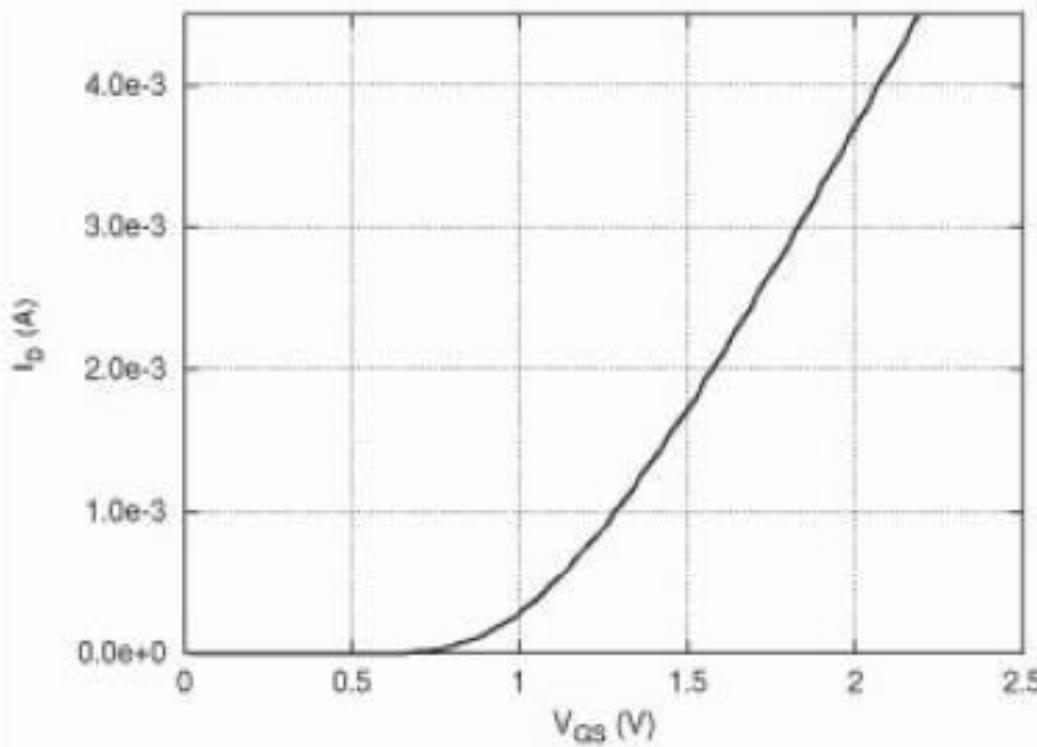
Saturation Region



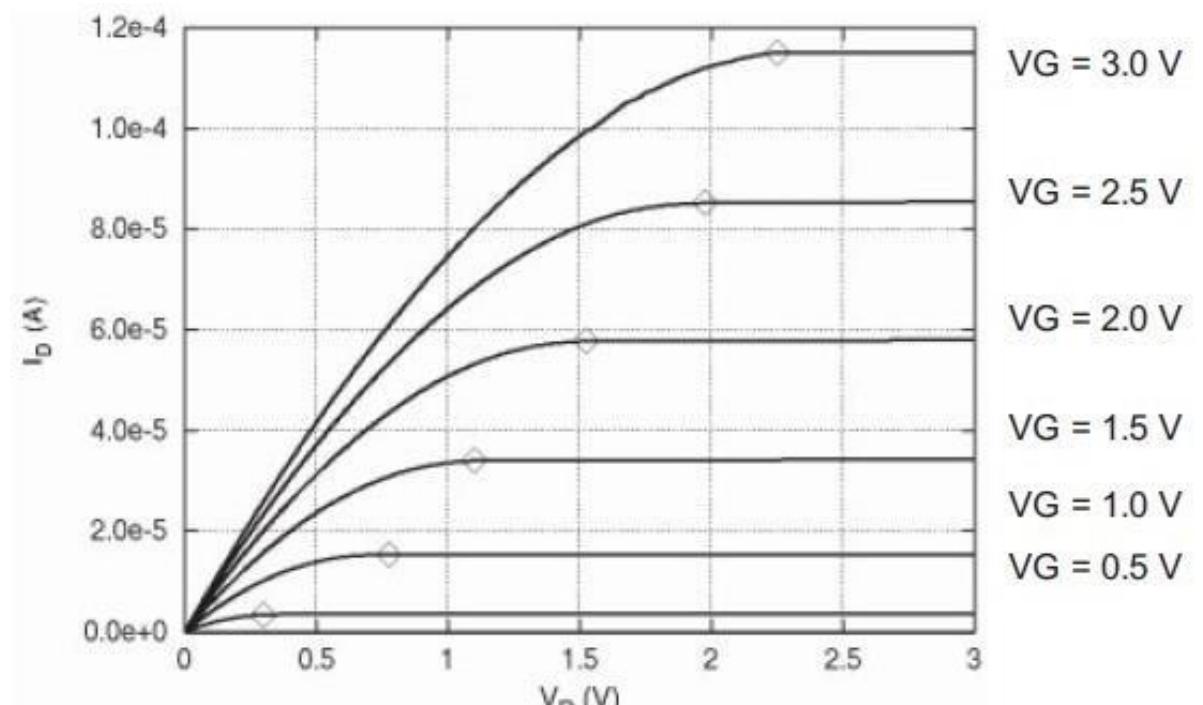
Important Observations

- The bulk or substrate of *n*MOS transistors must always be connected to the lower voltage that is the reference terminal.
- The positive convention current in an *n*MOS device is from the drain to the source, and is referred to as I_{DS} or just I_D , since drain and source current are equal.
- When a positive voltage is applied to the drain terminal, the drain current depends on the voltage applied to the gate control terminal.

Input and Output characteristics of nMOS



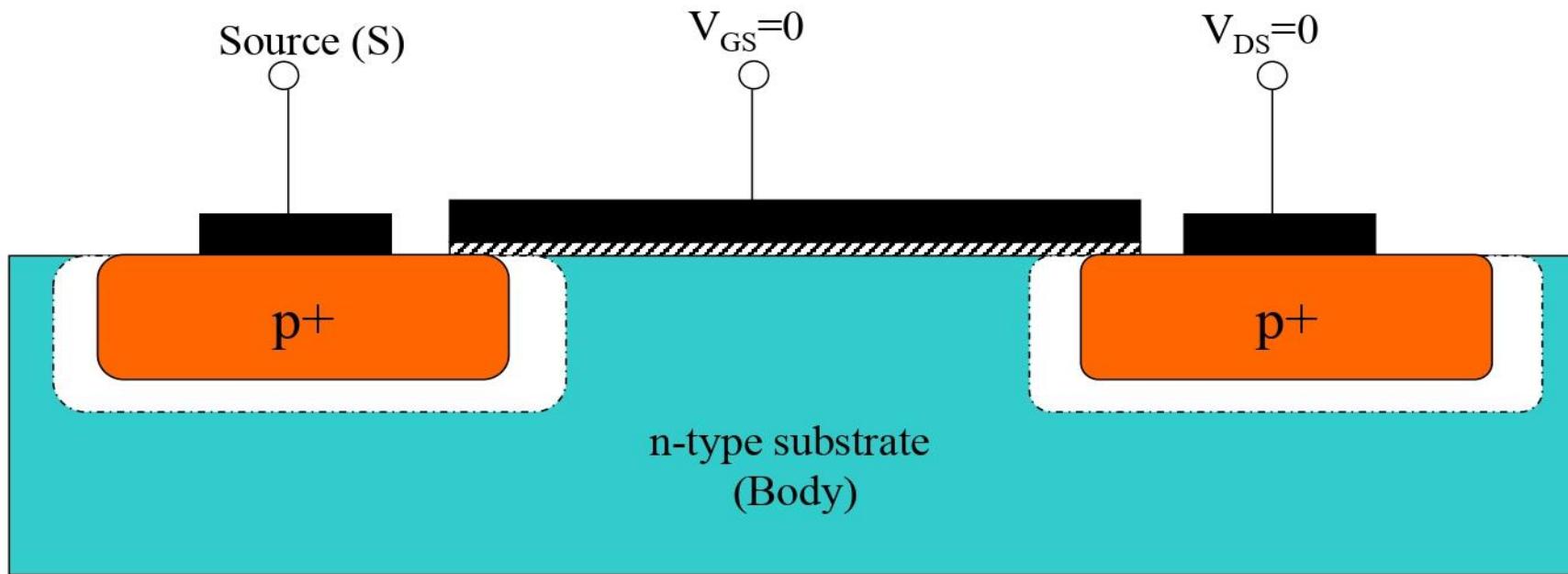
input characteristics (I_D vs. V_{GS}) for an nMOS,

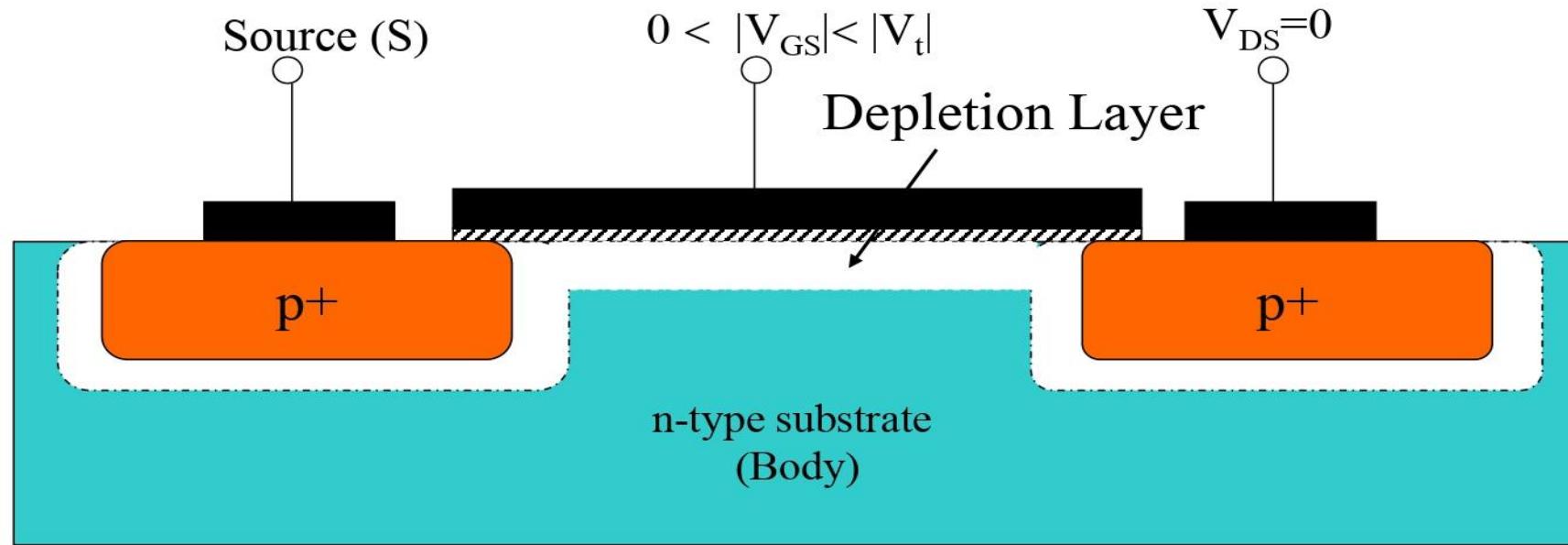


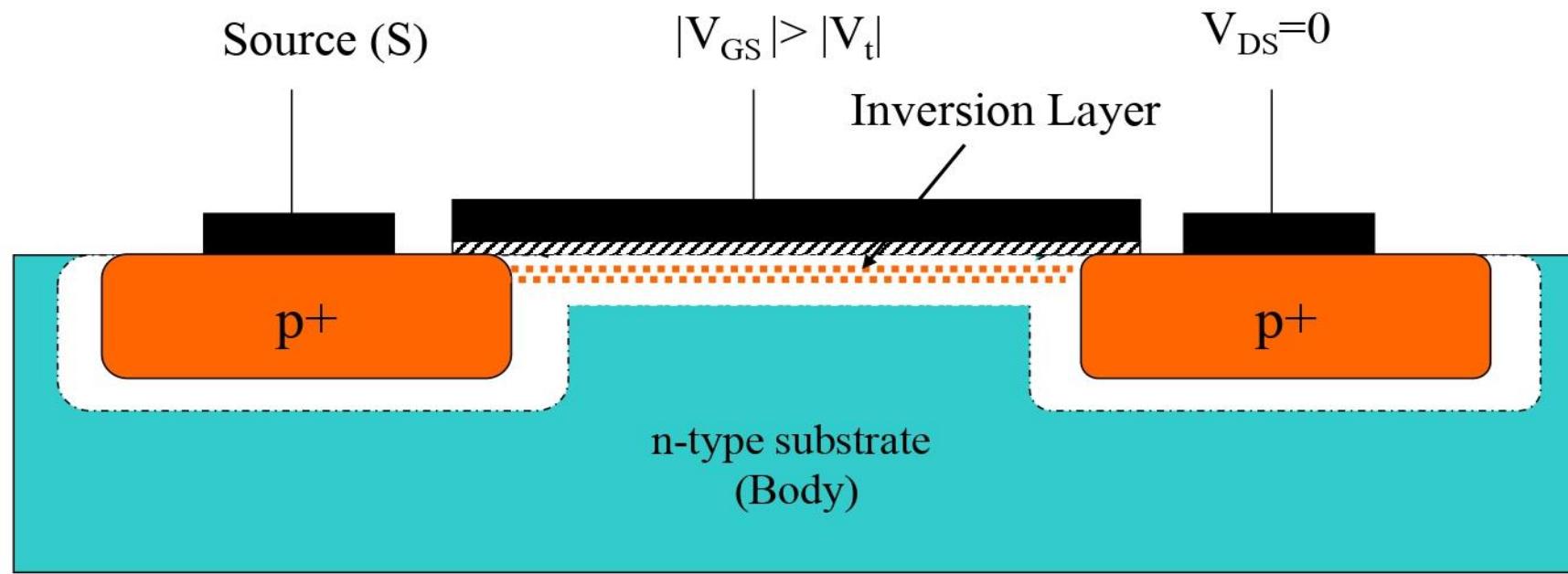
nMOS transistor output characteristics

Enhancement PMOS Transistor

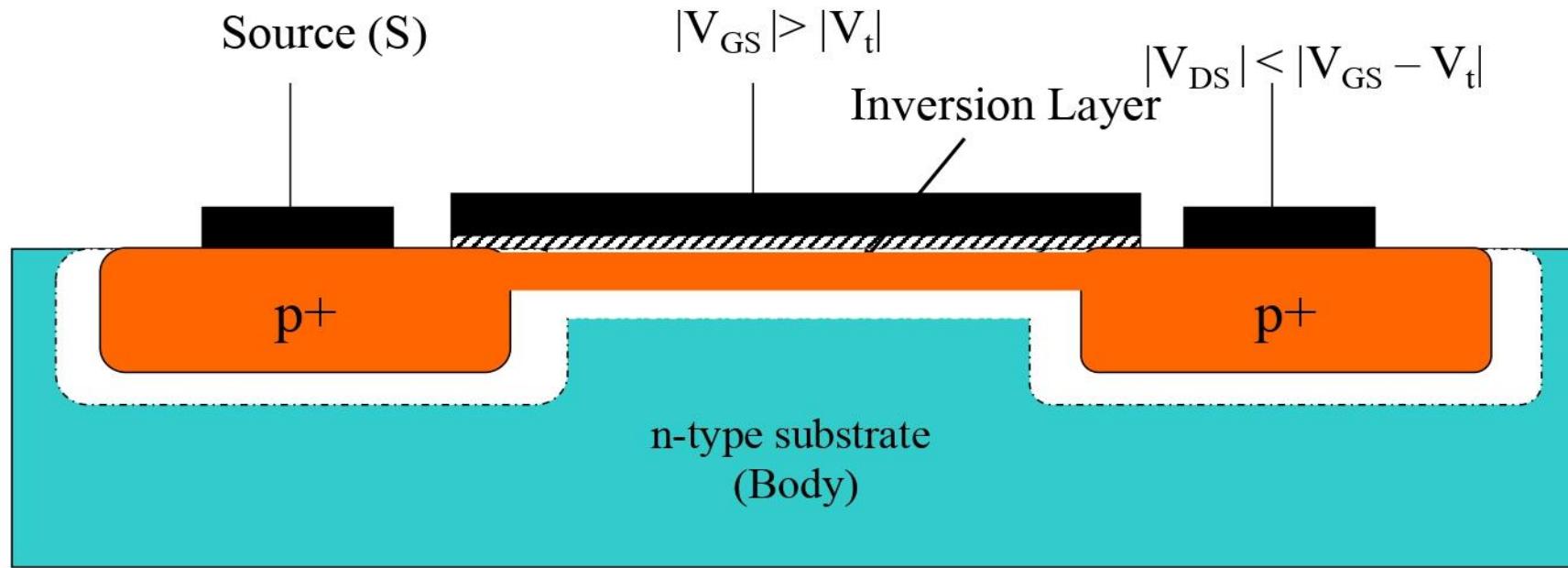
Cut-off Region

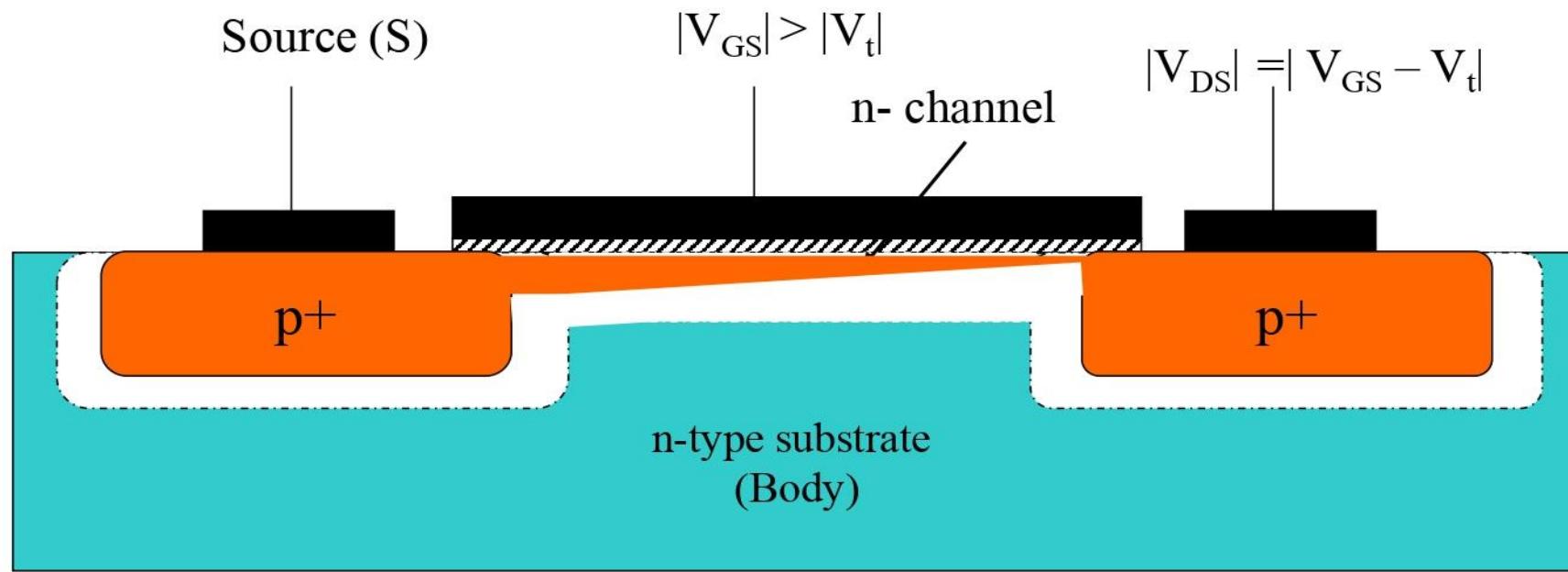




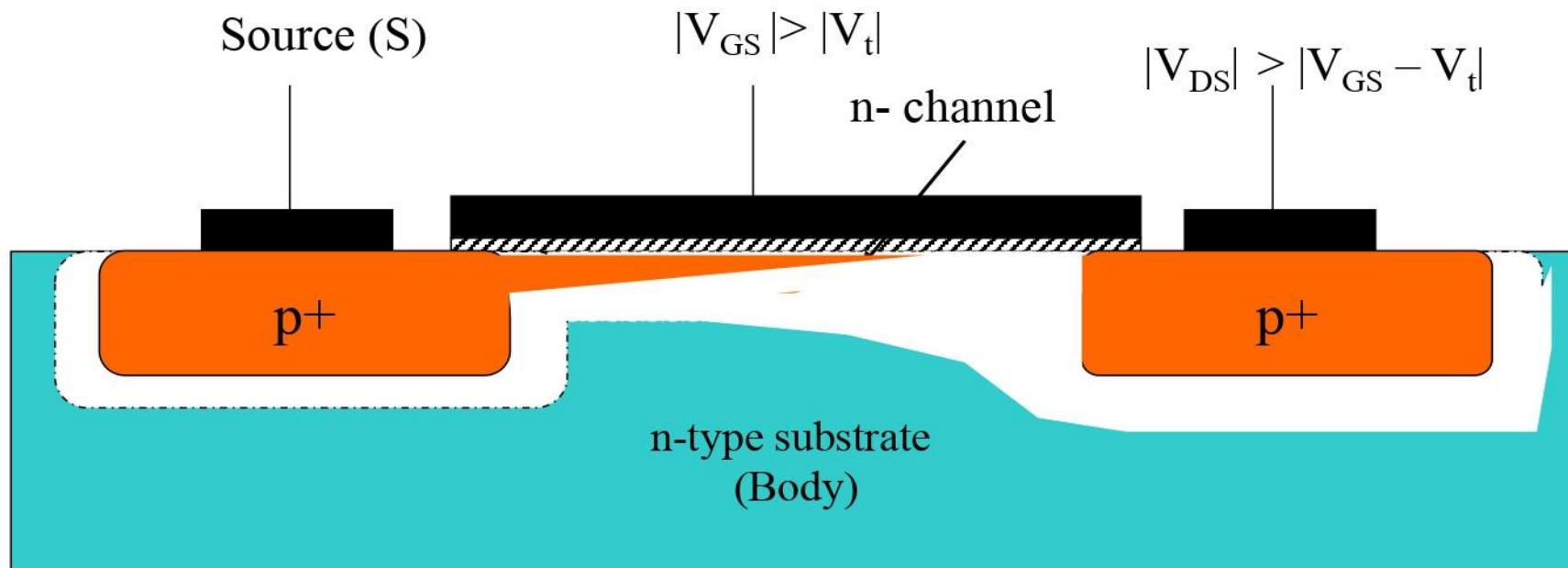


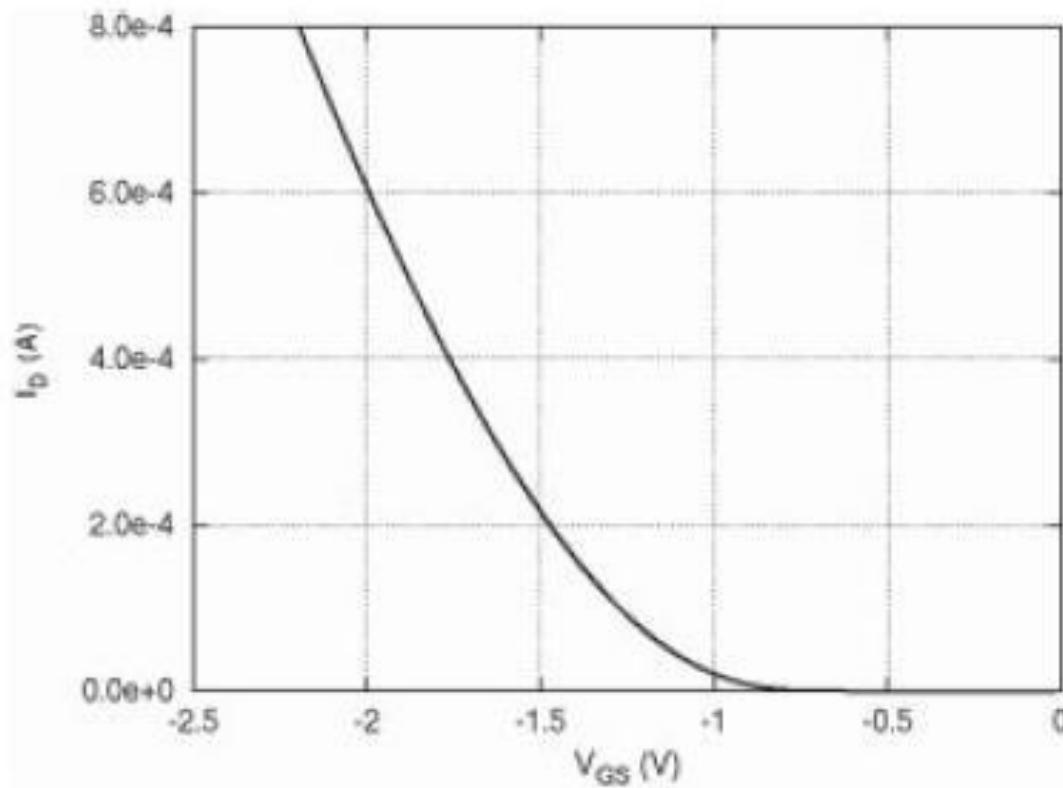
Linear region



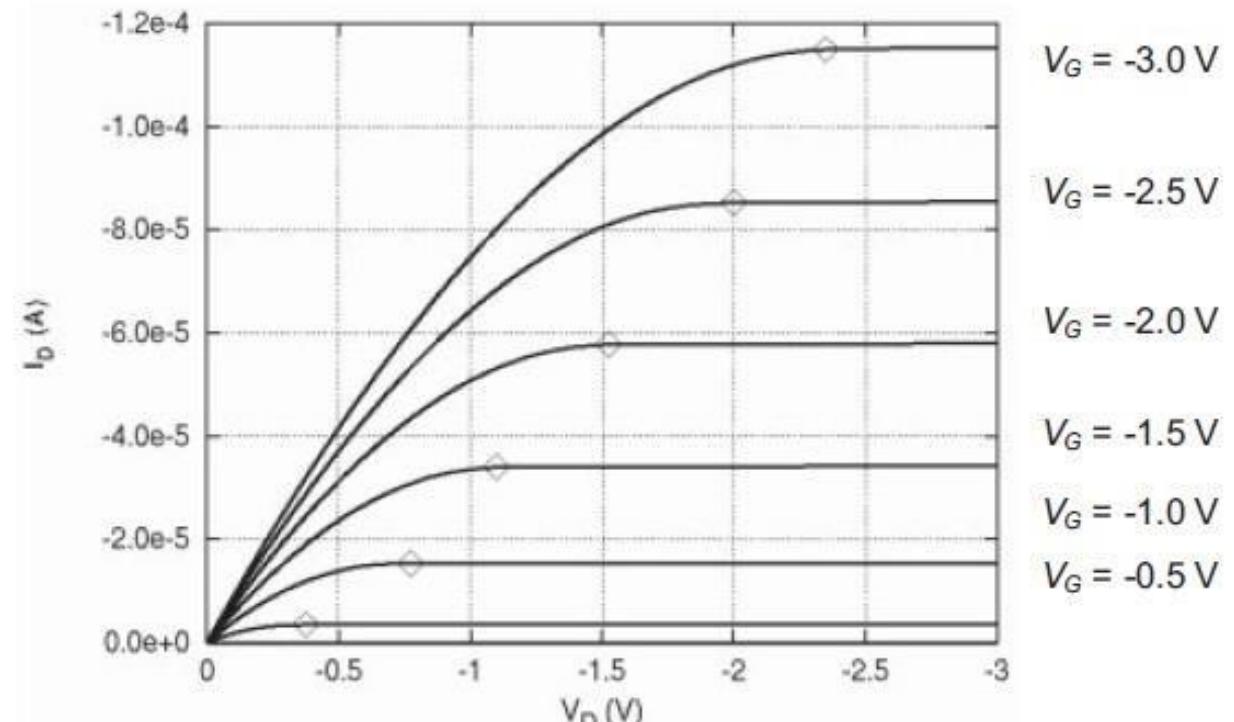


Saturation Region





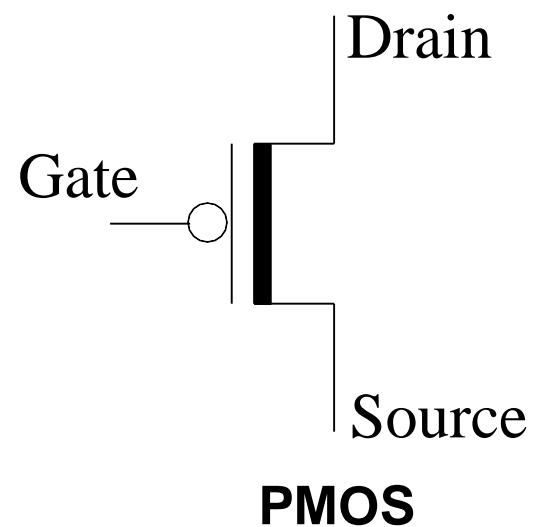
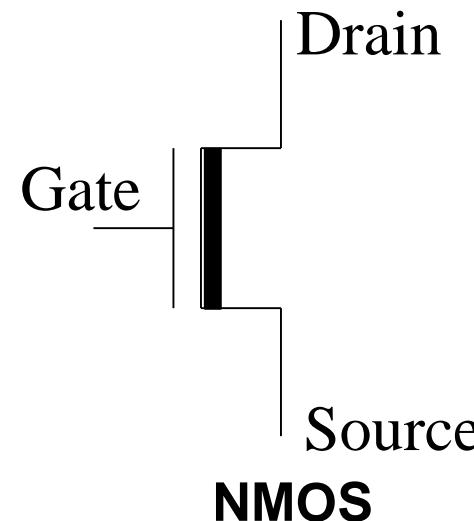
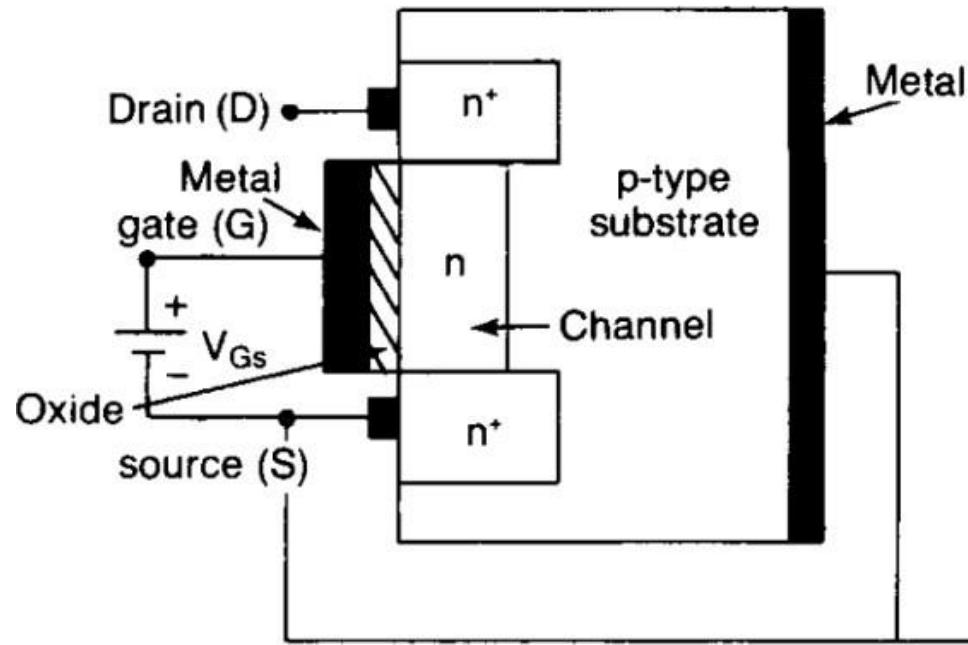
input characteristics (I_D vs. V_{GS}) for an *p*MOS,



*p*MOS transistor output characteristics

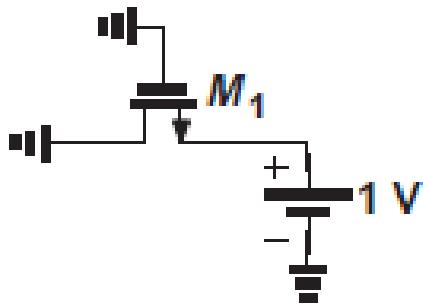
Depletion NMOS Transistor

Depletion Type MOS

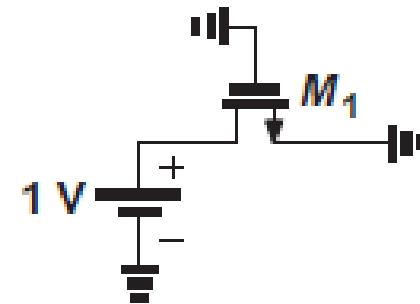


- In Depletion MOS structure, the source & drain are diffused on P- substrate as shown above.
- Positive voltages enhances number of electrons from source to drain.
- Negative voltage applied to gate reduces the drain current
- This is called as ' normally ON ' MOS.

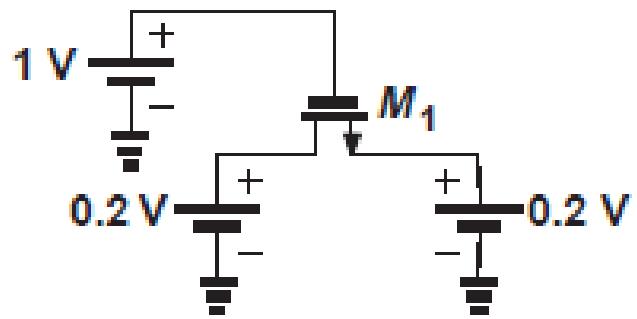
Determine the region of operation of M_1 in each of the circuits shown in Fig.



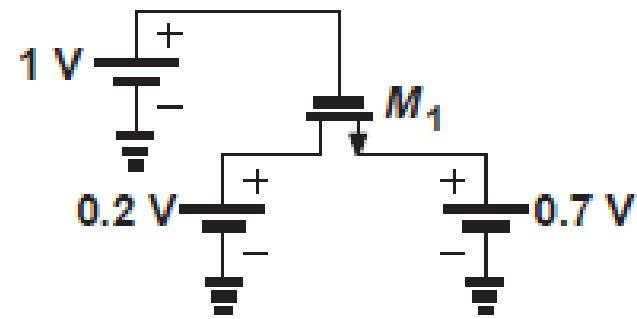
(a)



(b)

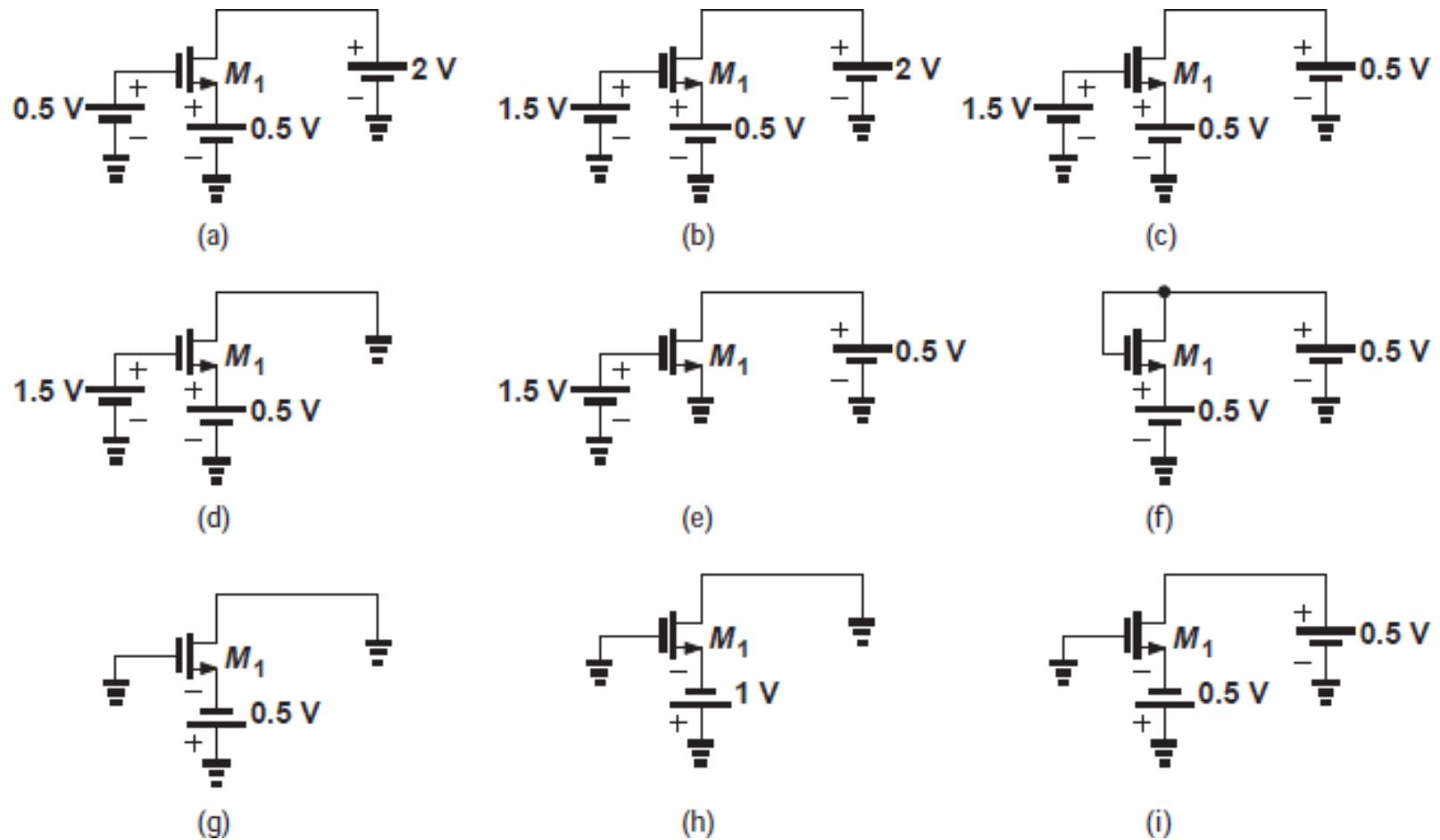


(c)



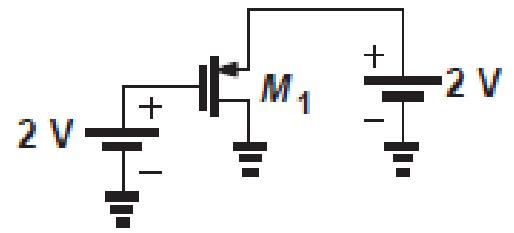
(d)

$V_{TH} = 0.4 \text{ V}$ for NMOS devices and -0.4 V for PMOS devices.

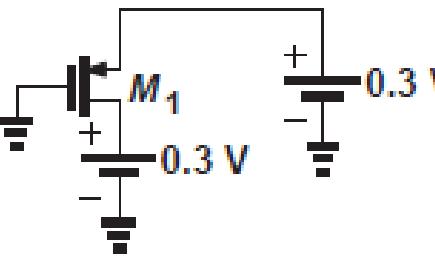


$V_{TH} = 0.4 \text{ V}$ for NMOS devices and -0.4 V for PMOS devices.

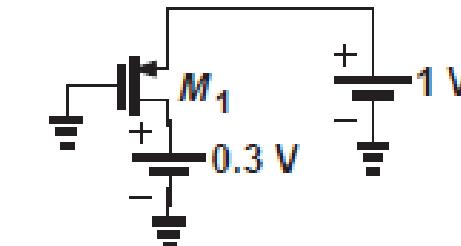
Determine the region of operation of M_1 in each of the circuits shown in Fig.



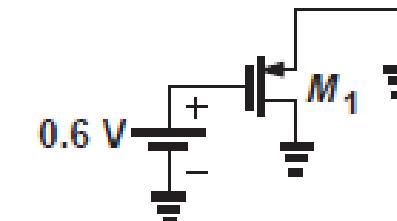
(a)



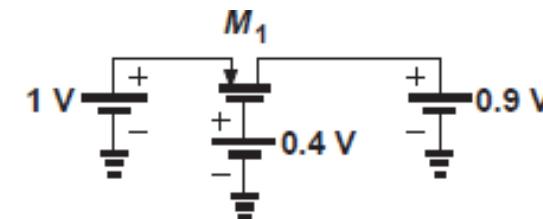
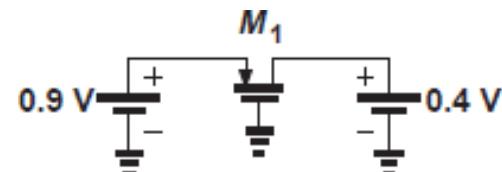
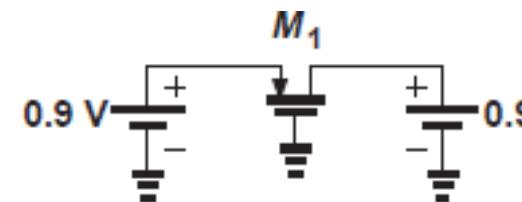
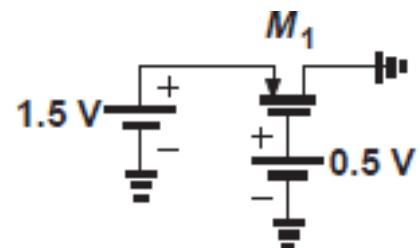
(b)



(c)

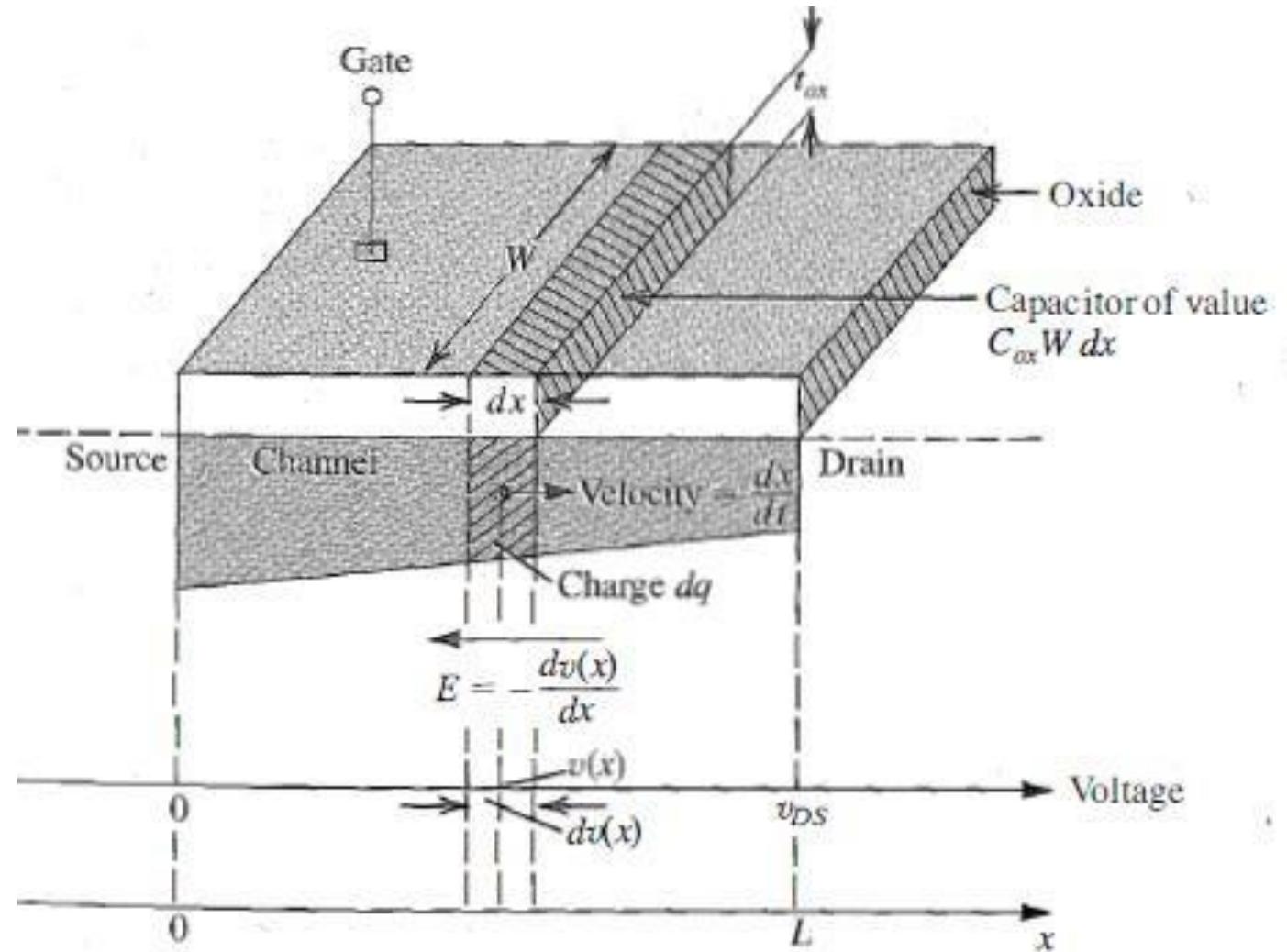


(d)



Current Derivation For Enhancement NMOS

Drain to Source Current I_{DS}



The gate and the channel region form a parallel plate capacitor for which the oxide layer serves as a dielectric.

$$C_o = \frac{\epsilon_{ox}}{t_{ox}}$$

$$\epsilon_{ox} = 3.9\epsilon_0 = 3.9 \times 8.854 \times 10^{-12} = 3.45 \times 10^{-11} \text{ F/m}$$

consider the infinitesimal strip of the gate at distance x from the source. The capacitance of this strip is

$$C_{ox} W dx$$

To find the charge stored on this infinitesimal strip of the gate capacitance, we multiply the capacitance by the effective voltage between the gate and the channel at point x ,

$$V_{GS} - V(x) - V_t$$

Charge dq in the infinitesimal portion of the channel at point x is

$$dq = -C_{ox}Wdx[V_{GS} - V(x) - V_t]$$

Negative sign accounts for the fact that dq is a negative charge

The voltage V_{DS} produces an electric field along the channel in the negative x direction

$$E(x) = -\frac{dV(x)}{dx}$$

The electric field $E(x)$ causes the electron charge dq to drift toward the drain with a velocity $\frac{dx}{dt}$

$$\frac{dx}{dt} = -\mu_n E(x) = \mu_n \frac{dV(x)}{dx}$$

Where μ_n is the mobility of electrons in the channel (called surface mobility).

The resulting drift current i

$$i = \frac{dq}{dt} = \frac{dq}{dx} \frac{dx}{dt}$$

$$i = -\mu_n C_{ox} W [V_{GS} - V(x) - V_t] \frac{dV(x)}{dx}$$

Thus i must be equal to the source-to-drain current. Since we are interested in the drain-to-source current i_D , we can find it as

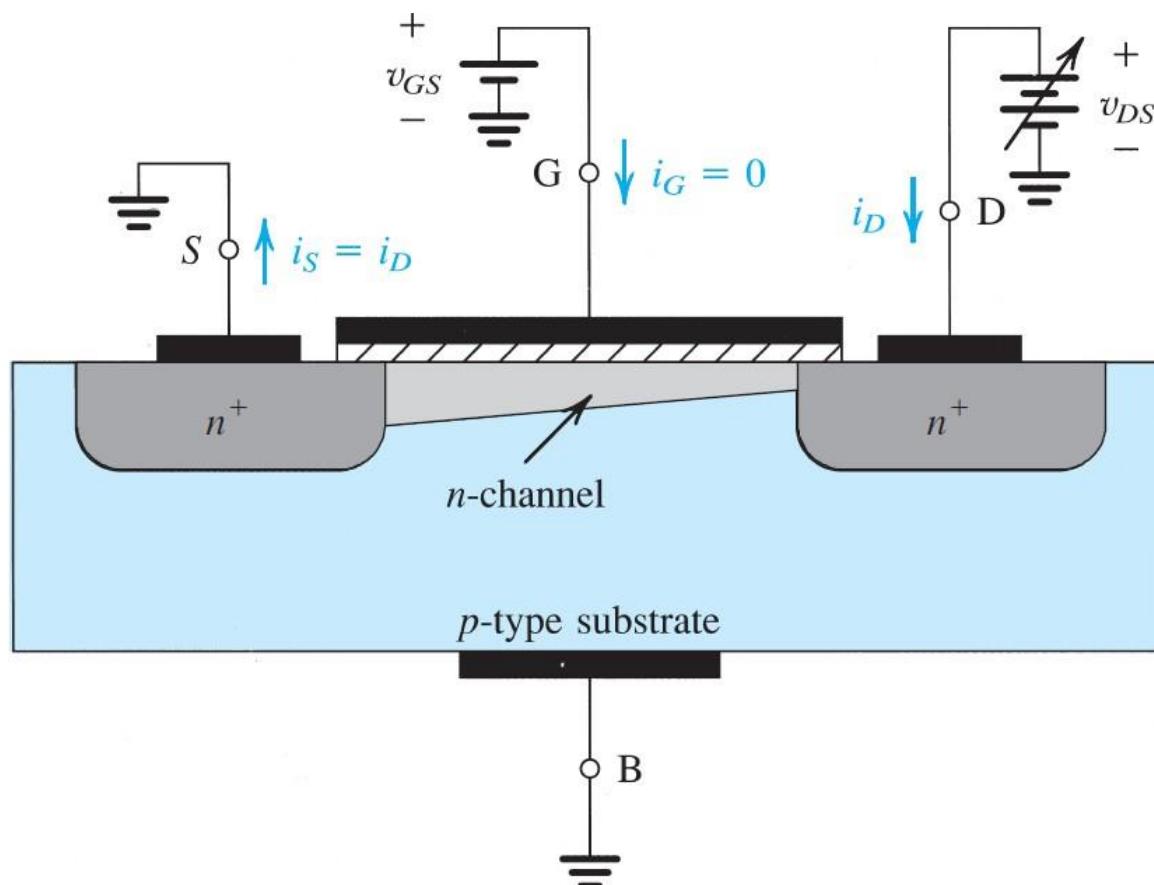
$$i_D = \mu_n C_{ox} W [V_{GS} - V(x) - V_t] \frac{dV(x)}{dx}$$

$$i_D dx = \mu_n C_{ox} W [V_{GS} - V(x) - V_t] dV(x)$$

Integrating both sides of this equation from $x = 0$ to $x = L$ and, correspondingly, for $V(0) = 0$ to $V(L) = V_{DS}$,

$$\int_0^L i_D dx = \int_0^{V_{DS}} \mu_n C_{ox} W [V_{GS} - V(\textcolor{brown}{x}) - V_t] dV(x)$$

Linear region: $I = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_t)V_{DS} - \frac{V_{DS}^2}{2}]$



Drain to Source Current I_{DS}

Saturation Region $V_{DS} \geq (V_{GS} - V_T)$

$$I = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_t)(V_{GS} - V_t) - \frac{(V_{GS} - V_t)^2}{2}]$$

$$I = \mu_n C_{ox} \frac{W}{L} [(V_{GS} - V_t)^2 - \frac{(V_{GS} - V_t)^2}{2}]$$

$$I = \mu_n C_{ox} \frac{W}{2L} (V_{GS} - V_t)^2$$

$$\mathcal{C}_G = \frac{\varepsilon_{ins}\varepsilon_0WL}{t_{ox}}$$

$$K'=\frac{\mu C_G}{WL}$$

$$\mathcal{C}_G = \mathcal{C}_{ox}WL$$

1. A 0.18- μm fabrication process is specified to have $t_{\text{ox}}= 4\text{nm}$, $\mu_n=450\text{cm}^2/\text{Vs}$ and $V_T=0.5\text{V}$. Find the value of $\mu_n C_{\text{ox}}$ (Also known as k_n' process transconductance) For a MOSFET with minimum length fabricated in this process, find the required value of W so that the device exhibits a channel resistance r_{DS} of 1K at $V_{\text{GS}}=1\text{V}$. Given $e_{\text{ox}}=3.45\text{e}-11 \text{ F/m}$

$$K_n' = 388 \mu\text{A/V}^2$$

$$W = 0.93 \mu\text{m}$$

2. Consider a process technology for which $L_{min} = 0.4 \mu\text{m}$, $t_{ox} = 8 \text{ nm}$, $\mu_n = 450 \text{ cm}^2/\text{V}\cdot\text{s}$, and $V_t = 0.7 \text{ V}$.

(a) Find C_{ox} and K_n' . $C_{ox} = 4.31 \times 10^{-3} \text{ F/m}^2$ $K_n' = 194.1 \mu\text{A/V}^2$

(b) For a MOSFET with $W/L = 8 \mu\text{m}/0.8 \mu\text{m}$ calculate the values of V_{ov} , V_{GS} , and V_{DSmin} needed to operate the transistor in the saturation region with a dc current $I_D = 100 \mu\text{A}$ $V_{DSmin} = V_{ov} = 0.32 \text{ V}$ $V_{GS} = 1.015 \text{ V}$

(c) For the device in (b), find the values of V_{ov} and V_{GS} required to cause the device to operate as a 1000-resistor for very small v_D $V_{ov} = 0.51 \text{ V}$ $V_{GS} = 1.215 \text{ V}$

3. For a 0.8- μm process technology for which $t_{ox} = 15 \text{ nm}$ and $\mu_n = 550 \text{ cm}^2/\text{V}\cdot\text{s}$, find C_{ox} , K'_n , and the overdrive voltage V_{ov} required to operate a transistor having $W/L=20$ in saturation with $I_D = 0.2 \text{ mA}$. What is the minimum value of V_{DS} needed?

$$C_{ox} = 2.301 \times 10^{-3} \text{ F/m}^2 \quad K'_n = 126.5 \mu\text{A/V}^2 \quad V_{DSmin} = V_{ov} = .397 \text{ V} \approx 0.4 \text{ V}$$

4. A circuit designer intending to operate a MOSFET in saturation is considering the effect of changing the device dimensions and operating voltages on the drain current I_D . Specifically, by what factor does I_D change in each of the following cases?

- (a) The channel length is doubled.
- (b) The channel width is doubled.
- (c) The overdrive voltage is doubled.
- (d) The drain-to-source voltage is doubled.

Ans. 0.5; 2; 4; no change

5. An enhancement type NMOS transistor with $V_t = 0.7V$ has its source terminal grounded and a 1.5-V DC is applied to the gate. In what region does the device operate :for (a) $V_D = +0.5 V$ (b) $V_D = +0.9 V$ (c) $V_D = +3 V$

If the NMOS device in $\mu_n C_{ox} = 100 \mu\text{A/V}^2$, $W = 10 \mu\text{m}$, and $L = 1 \mu\text{m}$, find the value of drain current that results in each of the three cases (a), (b), and (c).

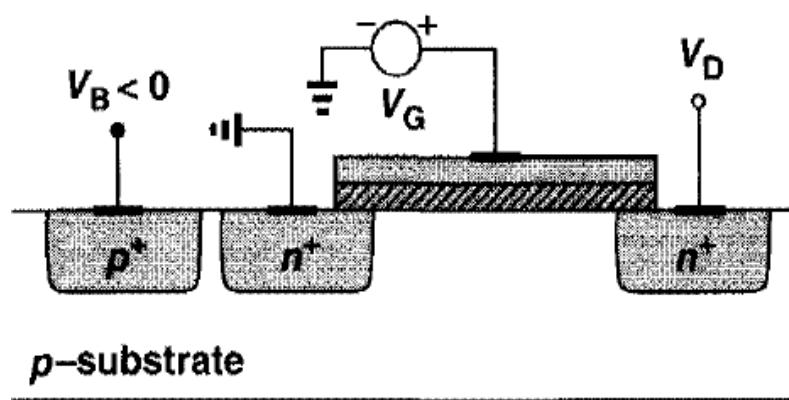
- (a) Non Sat $\rightarrow 275 \mu\text{A}$
- (b) Sat $\rightarrow 320 \mu\text{A}$
- (c) Sat $\rightarrow 320 \mu\text{A}$

Second-order Effects

- Body Effect.
- Sub threshold conduction
- Channel length modulation
- Mobility variation
- Fowler-Nordheim tunneling
- Drain punchthrough
- Impact Ionization-Hot electrons.

Body Effect

What happens if the bulk voltage of an N-MOSFET drops below the source voltage ?



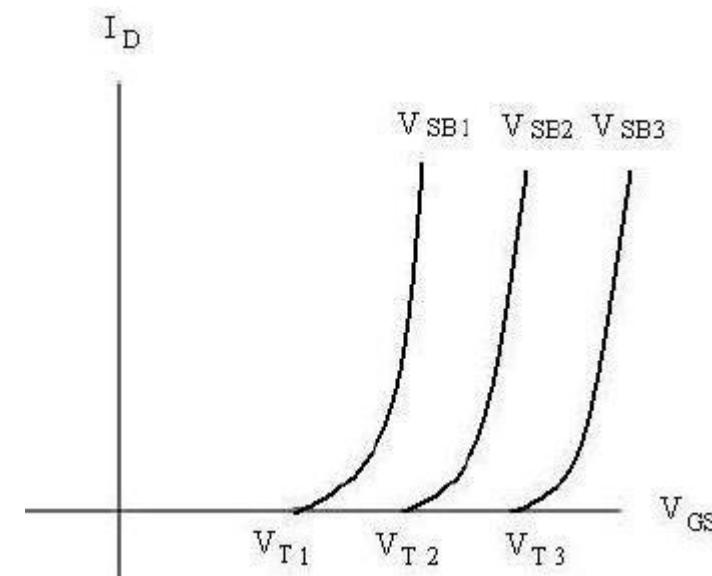
$$V_{TH} = V_{TH0} + \gamma \left(\sqrt{|2\Phi_F + V_{SB}|} - \sqrt{|2\Phi_F|} \right)$$

$$\gamma = \sqrt{2q\epsilon_{si}N_{sub}}/C_{ox}$$

$$V_{TH0} = \varphi_{MS} + 2\varphi_F + \frac{Q_{Dep}}{C_{OX}}$$

where φ_{MS} is the difference between the work functions of the polysilicon gate and the silicon substrate

φ_F = Fermi potential



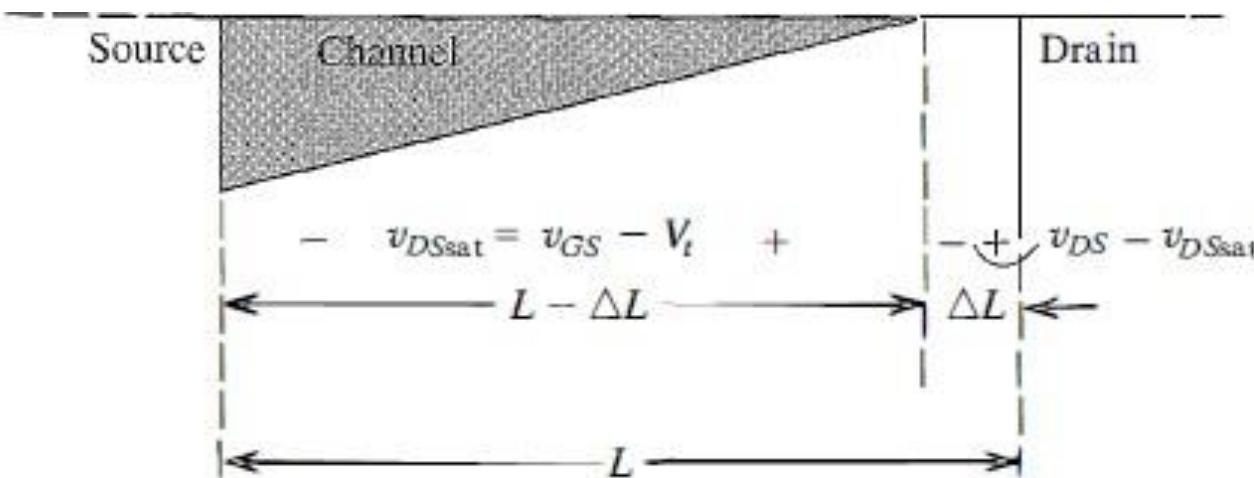
Sub threshold conduction

For $V_{GS} \approx V_{TH}$, a "weak" inversion layer still exists and some current flows from D to S. Even for $V_{GS} < V_{TH}$, I_D is finite, but it exhibits an *exponential* dependence on V_{GS} . Called "subthreshold conduction". This effect can be formulated for V_{DS} greater than roughly 200 mV as

$$I_D = I_0 \exp \frac{V_{GS}}{\xi V_T},$$

$\xi > 1$ is a nonideality factor

Channel length modulation



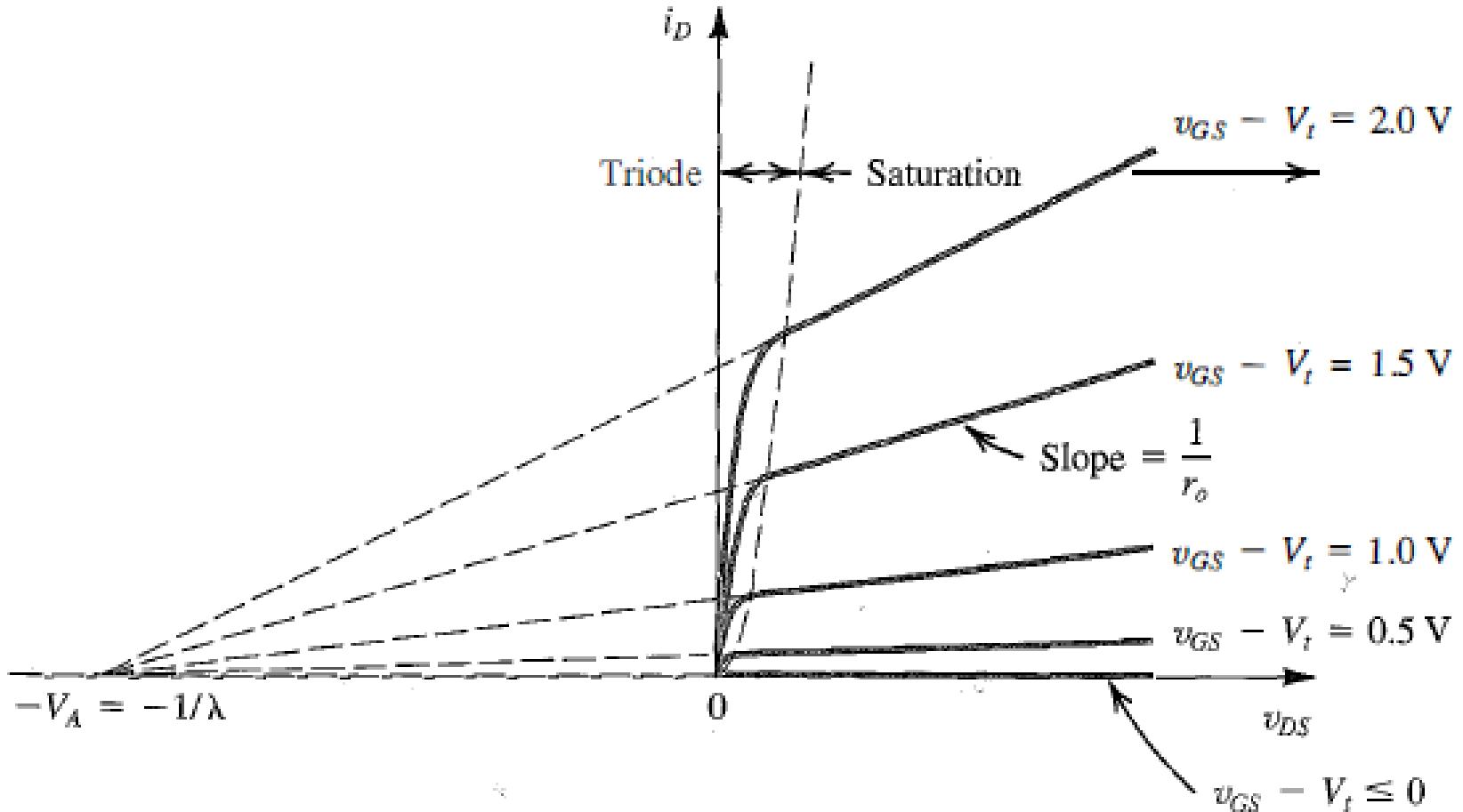
$$\begin{aligned} i_D &= \frac{1}{2} k'_n \frac{W}{L - \Delta L} (v_{GS} - V_t)^2 \\ &= \frac{1}{2} k'_n \frac{W}{L} \frac{1}{1 - (\Delta L/L)} (v_{GS} - V_t)^2 \\ &\approx \frac{1}{2} k'_n \frac{W}{L} \left(1 + \frac{\Delta L}{L}\right) (v_{GS} - V_t)^2 \\ \Delta L &= \lambda' v_{DS} \end{aligned}$$

$$i_D = \frac{1}{2} k'_n \frac{W}{L} \left(1 + \frac{\lambda'}{L} v_{DS}\right) (v_{GS} - V_t)^2$$

$$\lambda = \frac{\lambda'}{L}$$

λ is a process-technology parameter
with the dimensions of V^{-1}

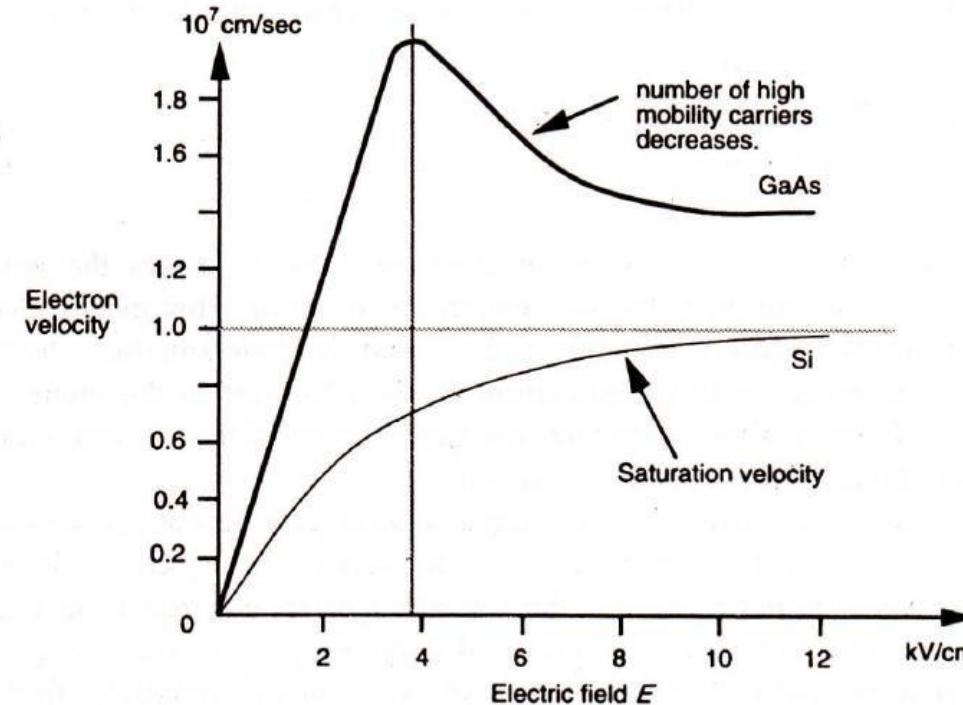
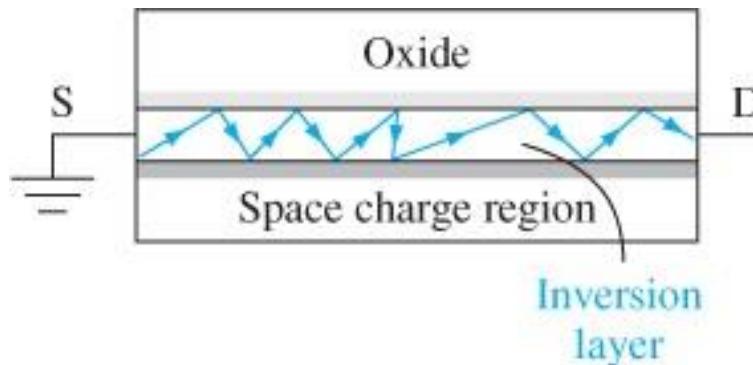
$$i_D = \frac{1}{2} k'_n \frac{W}{L} (v_{GS} - V_t)^2 (1 + \lambda v_{DS})$$



V_A is a process-technology parameter with the dimensions of V .

Mobility variation

Mobility is defined as the ease with which the charge carriers drift in the substrate material. Mobility decreases with increase in doping concentration and increase in temperature. Mobility is the ratio of average carrier drift velocity and electric field. Mobility is represented by the symbol μ .



Fowler Nordhiem tunneling:

When the gate oxide is very thin there can be a current between gate and source or drain by electron tunneling through the gate oxide. This current is proportional to the area of the gate of the transistor.

Drain punchthrough

When the drain is at a high voltage, the depletion region around the drain may extend to the source, causing the current to flow even if the gate voltage is zero. This is known as Punchthrough condition.

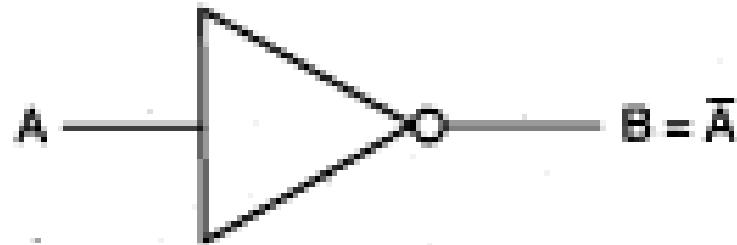
Impact ionization-Hot electrons

When the length of the transistor is reduced, the electric field at the drain increases. The field can become so high that electrons are imparted with enough energy we can term them as hot. These hot electrons impact the drain, dislodging holes that are then swept toward the negatively charged substrate and appear as a substrate current. This effect is known as Impact Ionization.

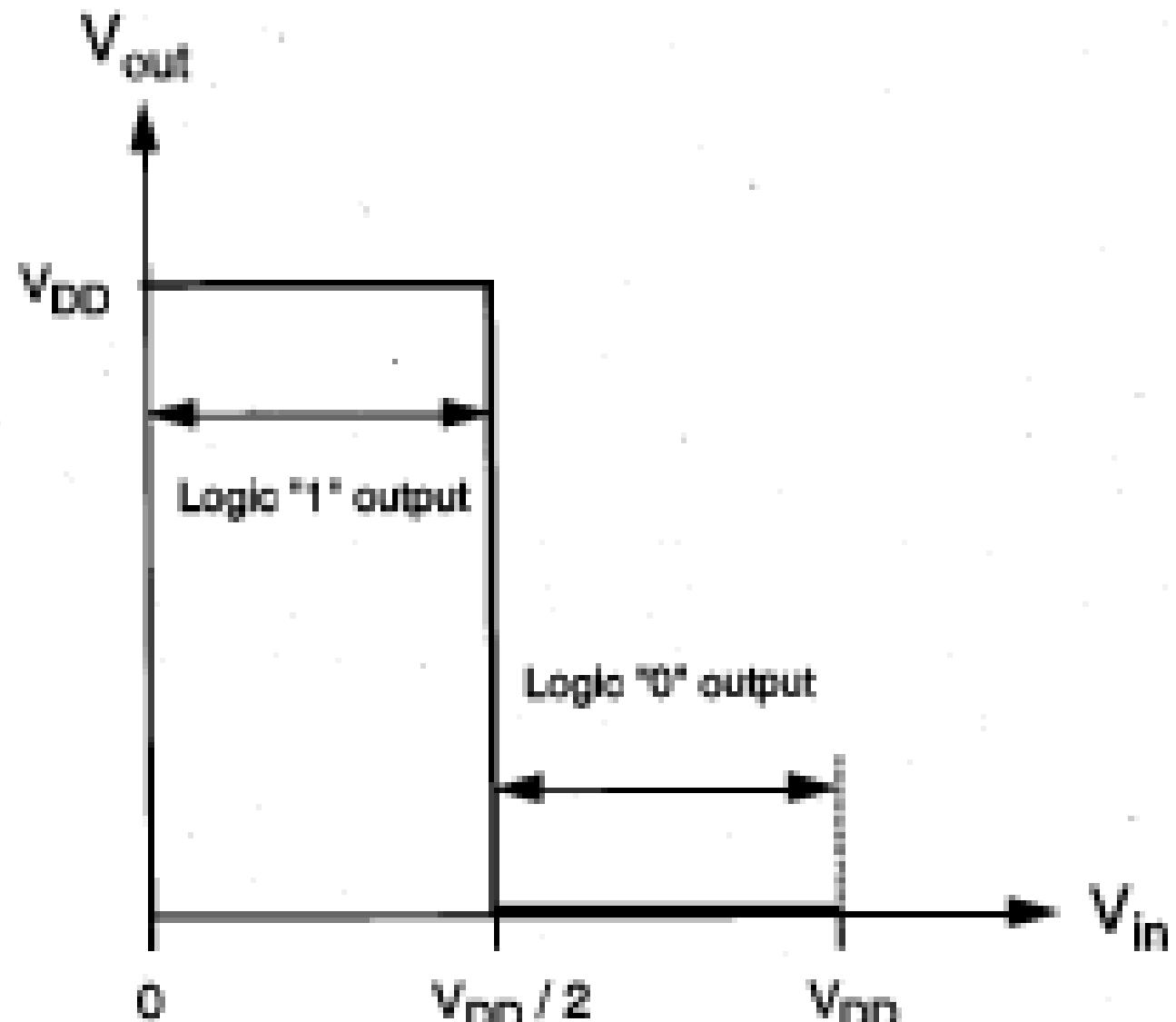
Inverters

1. Ideal inverter
2. Resistive load inverter
3. n-type MOSFET load
 - i. Enhancement load inverter
 - ii. Depletion load inverter
4. CMOS inverter

Ideal Inverters

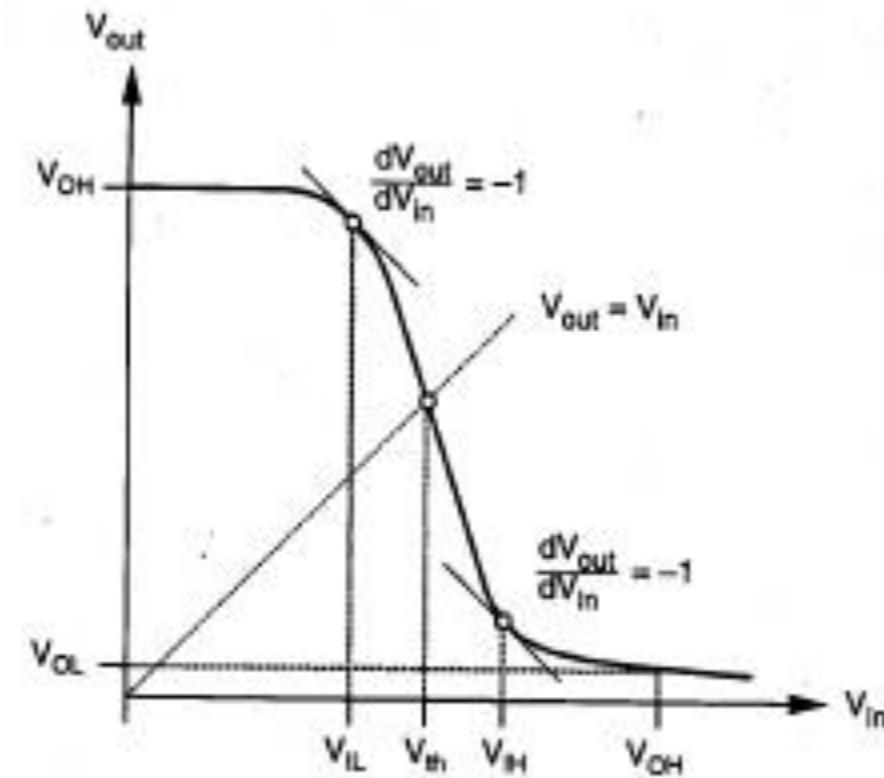
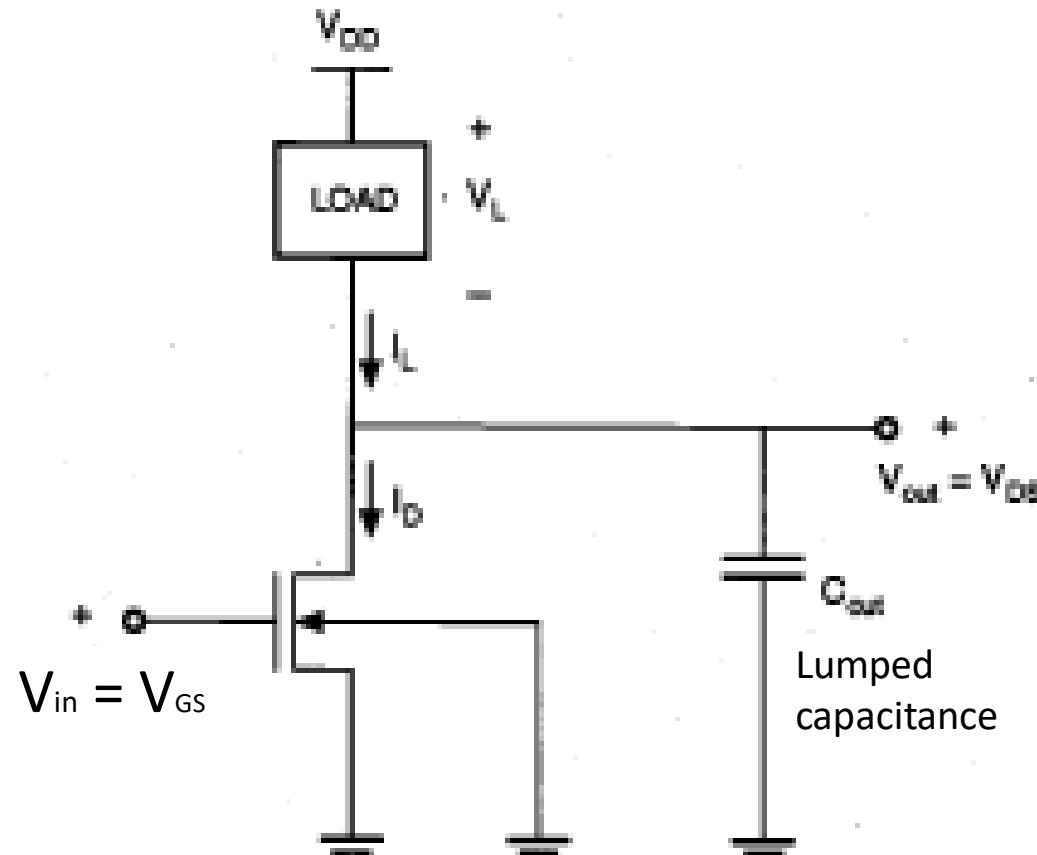


A	B
0	1
1	0



DC Voltage transfer characteristics (VTC)
of an ideal inverter

General circuit of an nMOS Inverter



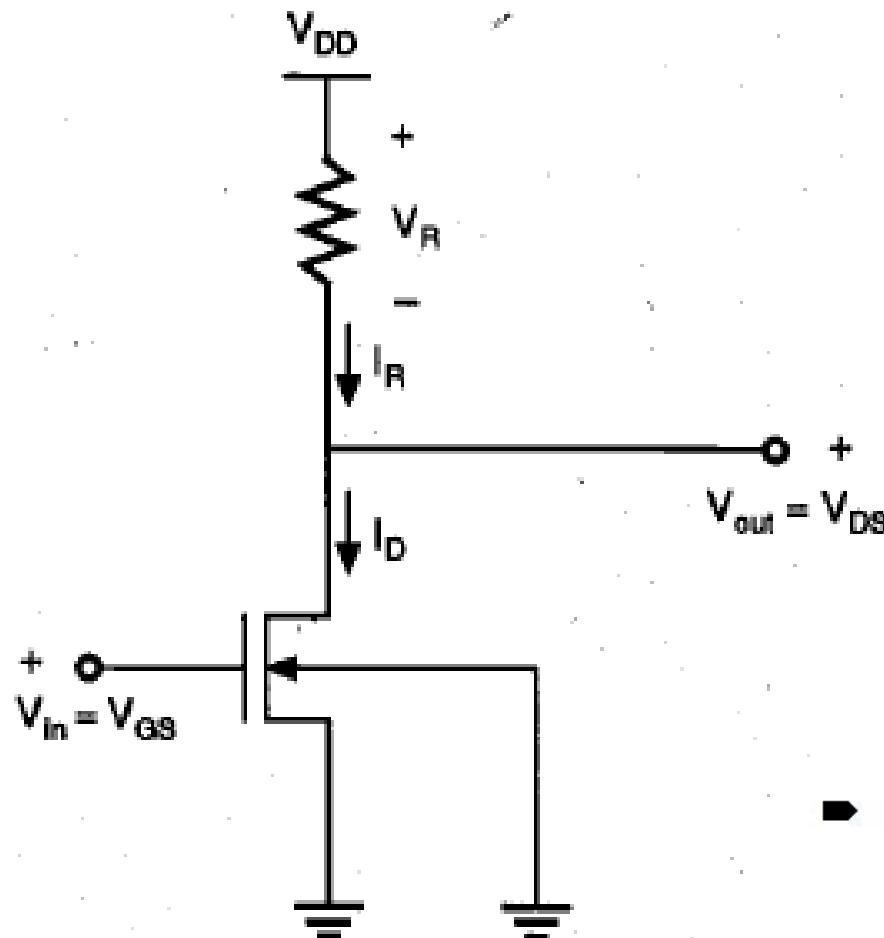
V_{OH} : Maximum output voltage when the output level is logic "1"

V_{OL} : Minimum output voltage when the output level is logic "0"

V_{IL} : Maximum input voltage which can be *interpreted* as logic "0"

V_{IH} : Minimum input voltage which can be *interpreted* as logic "1"

Resistive Load Inverter



Input Voltage Range	Operating Mode
$V_{in} < V_{T0}$	cut-off
$V_{T0} \leq V_{in} < V_{out} + V_{T0}$	saturation
$V_{in} \geq V_{out} + V_{T0}$	linear

output voltage V_{out} is

$$V_{out} = V_{DD} - I_R R$$

$$V_{out} = V_{DD} - I_D R$$

So, Drain current equation will be

$$I_D = \frac{V_{DD} - V_{out}}{R}$$

- if MOSFET is there in saturation region then

$$V_{in} - V_{T0} < V_{out} \text{ and } I_D = \frac{K}{2} (V_{GS} - V_{TO})^2$$

- If MOSFET is there in linear region then

$$V_{in} - V_{T0} > V_{out} \text{ and } I_D = \frac{K}{2} [2(V_{GS} - V_{TO})V_{DS} - V_{DS}^2]$$

Calculation of V_{OH} , V_{OL} , V_{IL} , V_{IH}

V_{OH} : Maximum output voltage when the output level is logic "1"

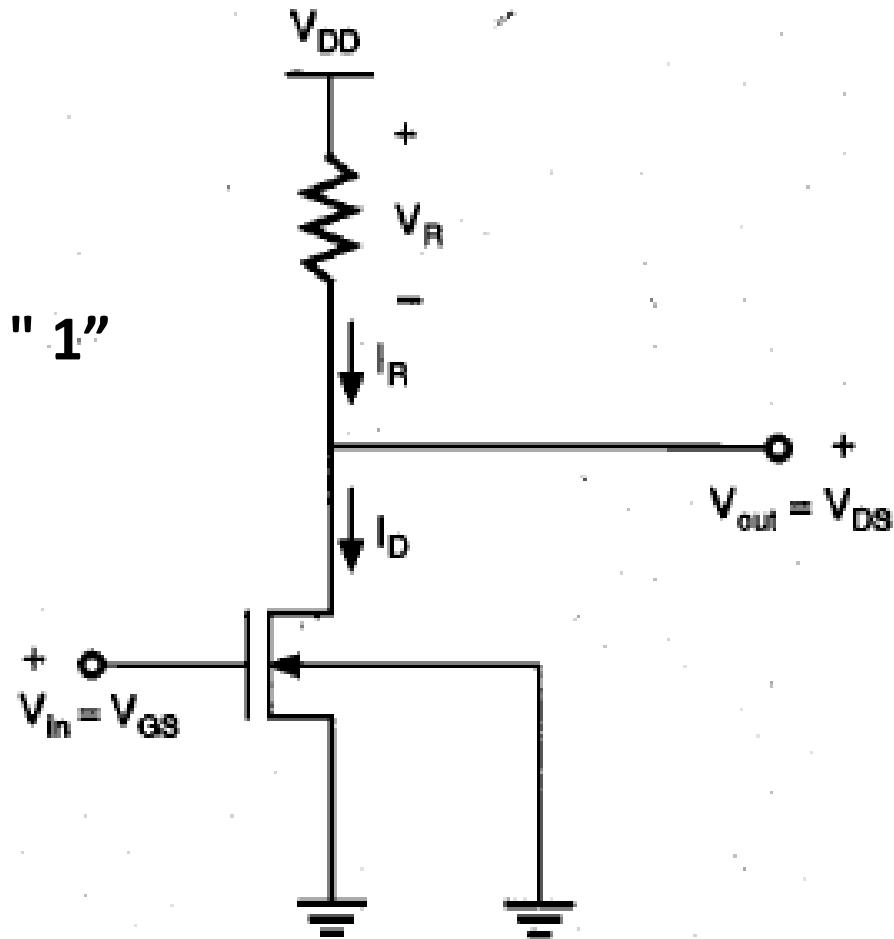
V_{OH} -

- Output voltage V_{out} is

$$V_{out} = V_{DD} - I_D R$$

$$\Rightarrow V_{out} = V_{DD}$$

$$\Rightarrow V_{OH} = V_{DD}$$



Calculation of V_{OH} , V_{OL} , V_{IL} , V_{IH}

V_{OL} : Minimum output voltage when the output level is logic “0”

V_{OL} :

- When $V_{in} - V_{TO} > V_{out}$, MOSFET is there in linear region, so, drain current will be

$$I_D = \frac{K}{2} [2(V_{in} - V_{TO})V_{out} - V_{out}^2]$$

- According to kirchoff's law in the drain current is

$$I_D = \frac{V_{DD} - V_{out}}{R}$$

- If we compare these two equations

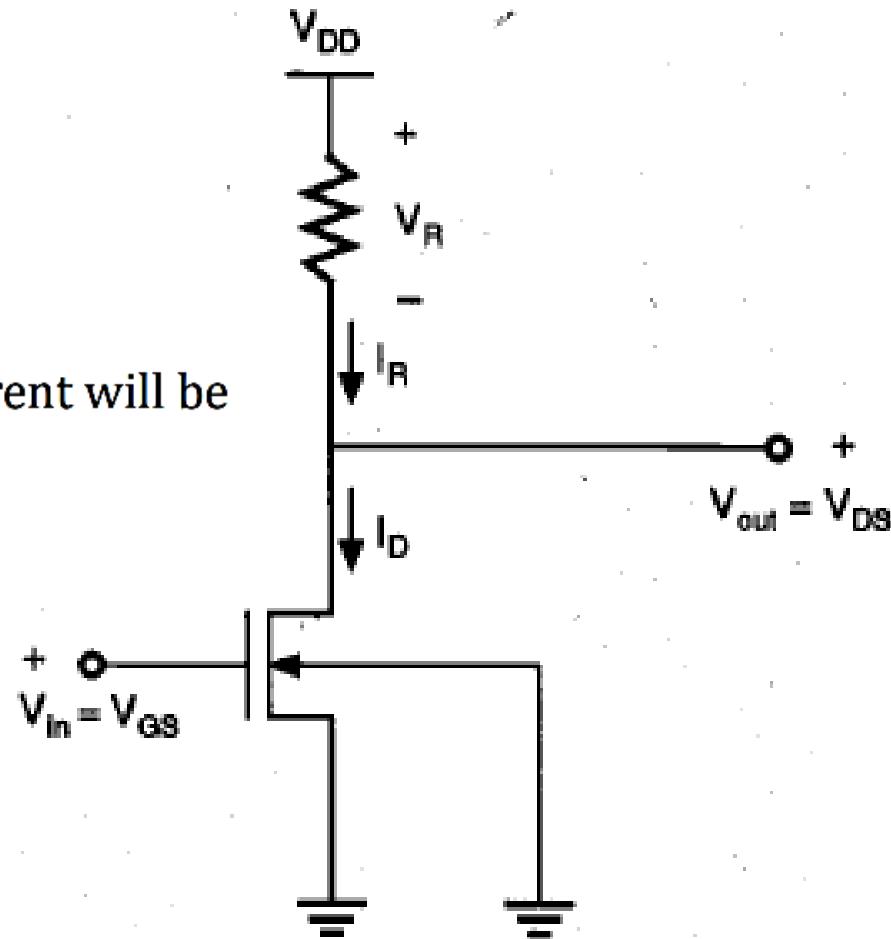
$$\frac{V_{DD} - V_{out}}{R} = \frac{K}{2} [2(V_{in} - V_{TO})V_{out} - V_{out}^2]$$

$$\Rightarrow \frac{V_{DD} - V_{OL}}{R} = \frac{K}{2} [2(V_{DD} - V_{TO})V_{OL} - V_{OL}^2]$$

$$\Rightarrow V_{OL}^2 - 2(V_{DD} - V_{TO} + \frac{1}{KR})V_{OL} + \frac{2}{KR}V_{DD} = 0$$

- If we solve the above equation, we get

$$V_{OL} = V_{DD} - V_{TO} + \frac{1}{KR} - \sqrt{(V_{DD} - V_{TO} + \frac{1}{KR})^2 - \frac{2V_{DD}}{KR}}$$



Calculation of V_{OH} , V_{OL} , V_{IL} , V_{IH}

V_{IL} : Maximum input voltage which can be *interpreted* as logic "0"

V_{IL} -

- When $V_{in} - V_{TO} < V_{out}$ MOSFET is there in saturation region, so drain current will be

$$I_D = \frac{K}{2} (V_{in} - V_{TO})^2$$

- Again compare the equation with circuit drain current equation

$$\frac{V_{DD} - V_{out}}{R} = \frac{K}{2} (V_{in} - V_{TO})^2$$

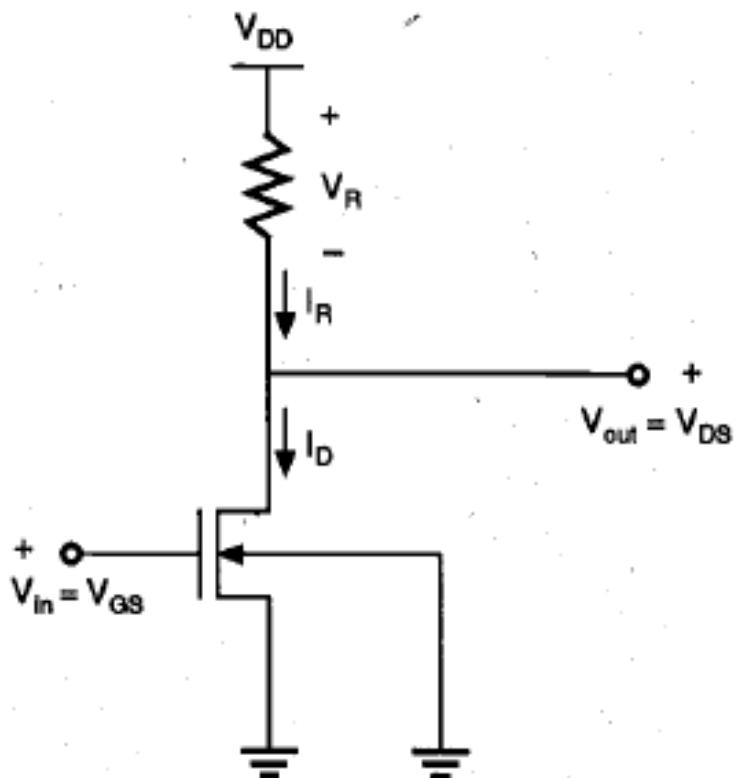
- We have to differentiate it with respect to V_{in}

$$\left(\frac{V_{DD} - V_{out}}{R} \right) \frac{dV_{out}}{dV_{in}} = \left[\frac{K}{2} (V_{in} - V_{TO})^2 \right] \frac{dV_{out}}{dV_{in}}$$

$$\Rightarrow -\frac{1}{R} \frac{dV_{out}}{dV_{in}} = K (V_{in} - V_{TO})$$

$$\Rightarrow \frac{1}{R} = K (V_{IL} - V_{TO})$$

$$\Rightarrow V_{IL} = V_{TO} + \frac{1}{KR}$$



Calculation of V_{OH} , V_{OL} , V_{IL} , V_{IH}

V_{IH} : Minimum input voltage which can be *interpreted* as logic "1"

V_{IH} -

- When $V_{in} - V_{TO} > V_{out}$ MOSFET is there in linear region, so drain current will be

$$I_D = \frac{K}{2} [2(V_{in} - V_{TO})V_{out} - V_{out}^2]$$

Compare the equation with circuit drain current equation

$$\frac{V_{DD} - V_{out}}{R} = \frac{K}{2} [2(V_{in} - V_{TO})V_{out} - V_{out}^2]$$

- We have to differentiate it with respect to V_{in}

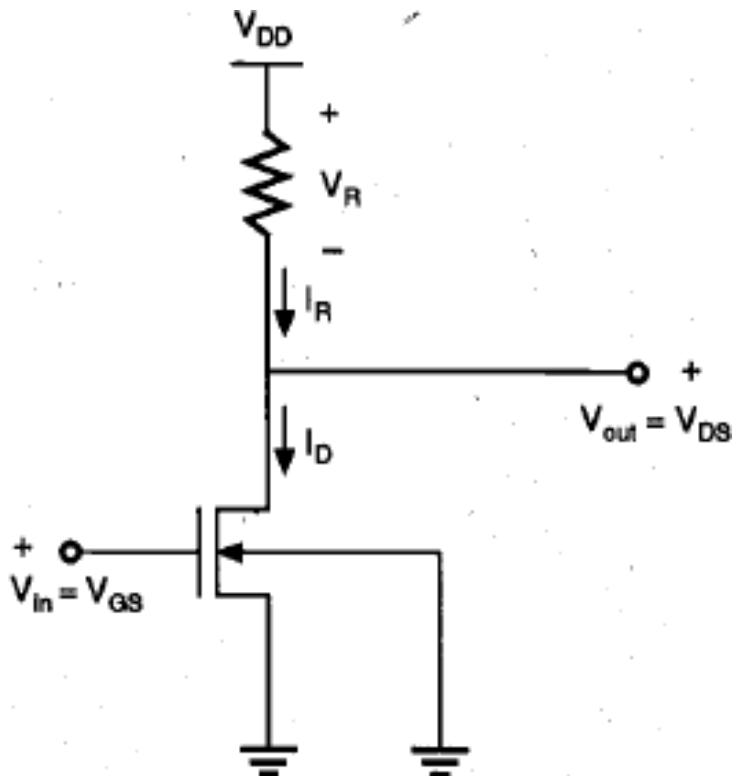
$$-\frac{1}{R} \frac{dV_{out}}{dV_{in}} = K/2 [2(V_{in} - V_{TO}) \frac{dV_{out}}{dV_{in}} + 2V_{out} - 2V_{out} \frac{dV_{out}}{dV_{in}}]$$

$$\Rightarrow \frac{1}{R} = K/2 [-2(V_{IH} - V_{TO}) + 4 V_{out}]$$

$$\Rightarrow \frac{1}{KR} = -(V_{IH} - V_{TO}) + 2 V_{out}$$

$$\Rightarrow \frac{1}{KR} = -V_{IH} + V_{TO} + 2V_{out}$$

$$\Rightarrow V_{IH} = V_{TO} + 2V_{out} - \frac{1}{KR}$$



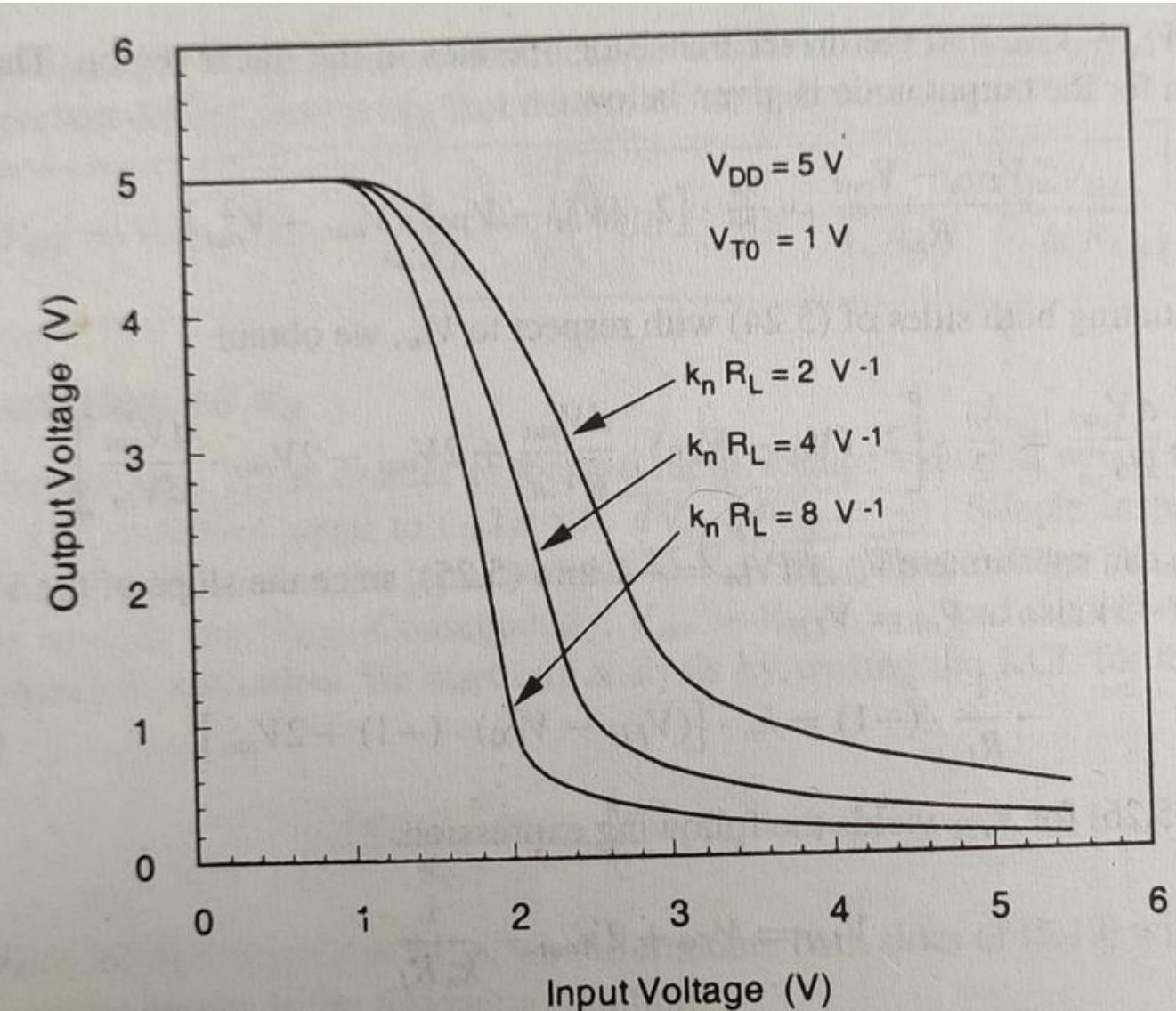
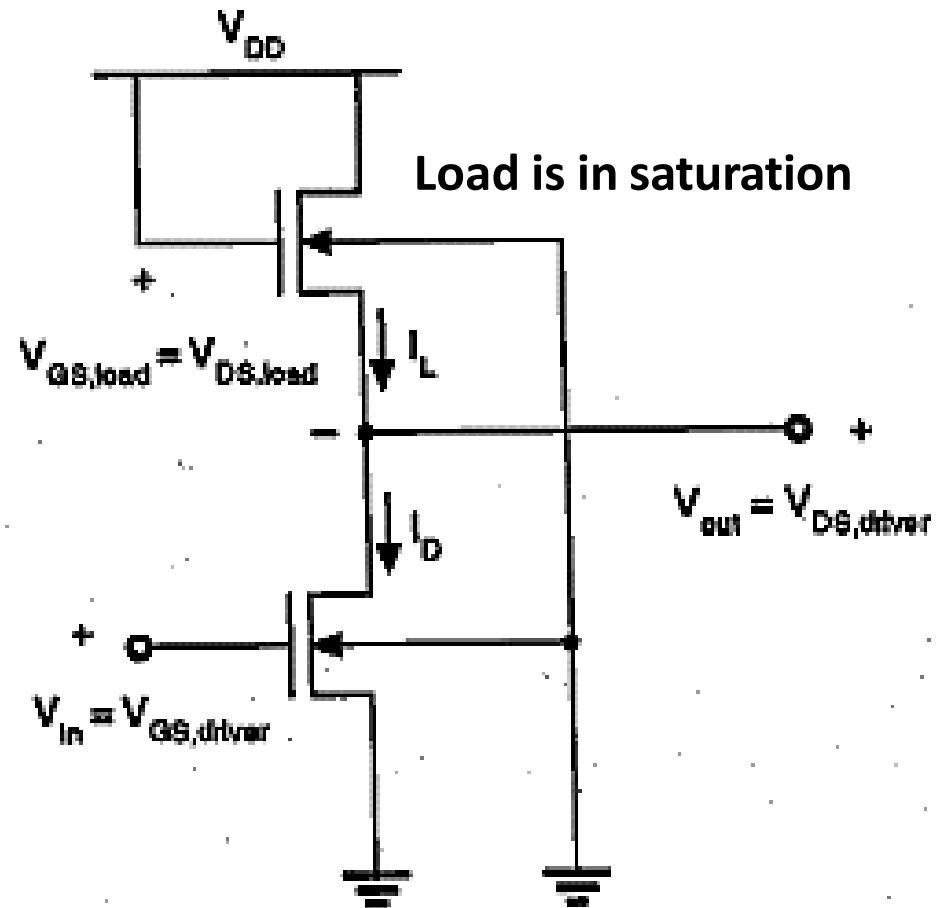
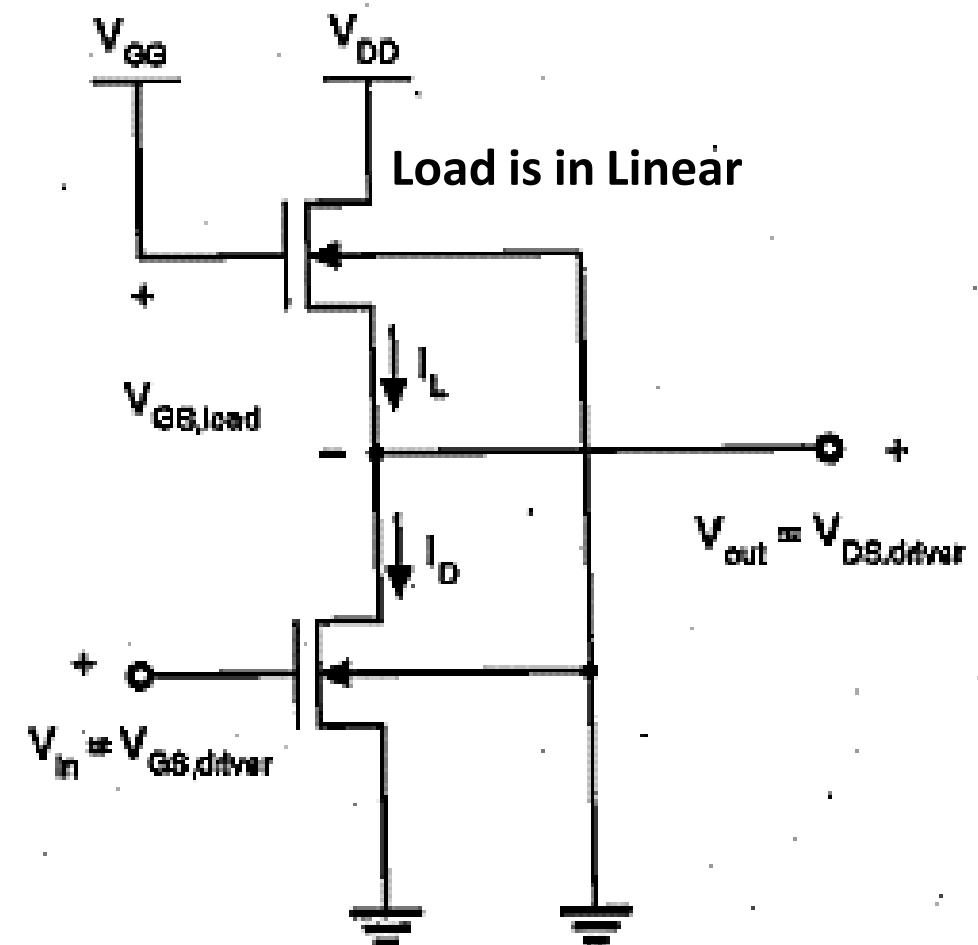


Figure 5.9 Voltage transfer characteristics of the resistive-load inverter, for different values of the parameter ($k_n R_L$).

Enhancement load NMOS inverter

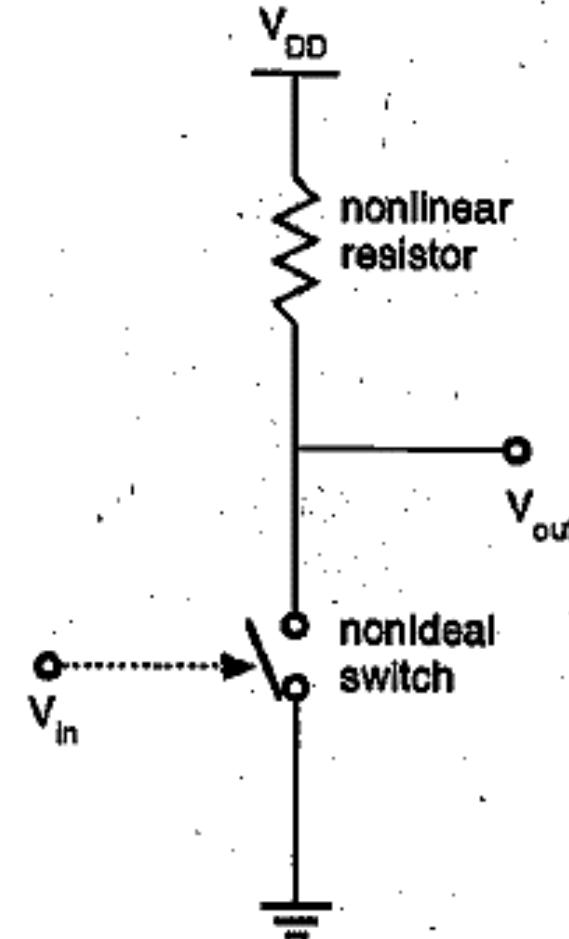
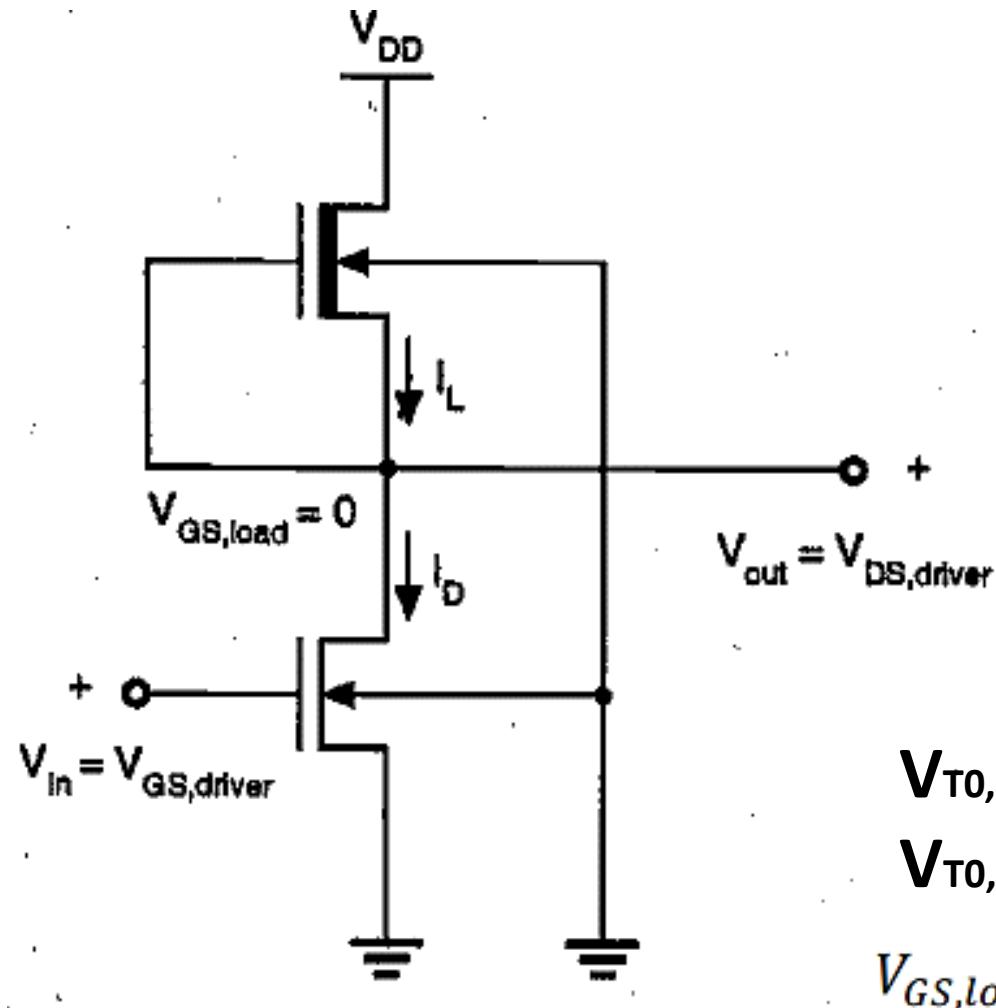


$$V_{OH} = V_{DD} - V_{T,load}$$



$$V_{OH} = V_{DD}$$

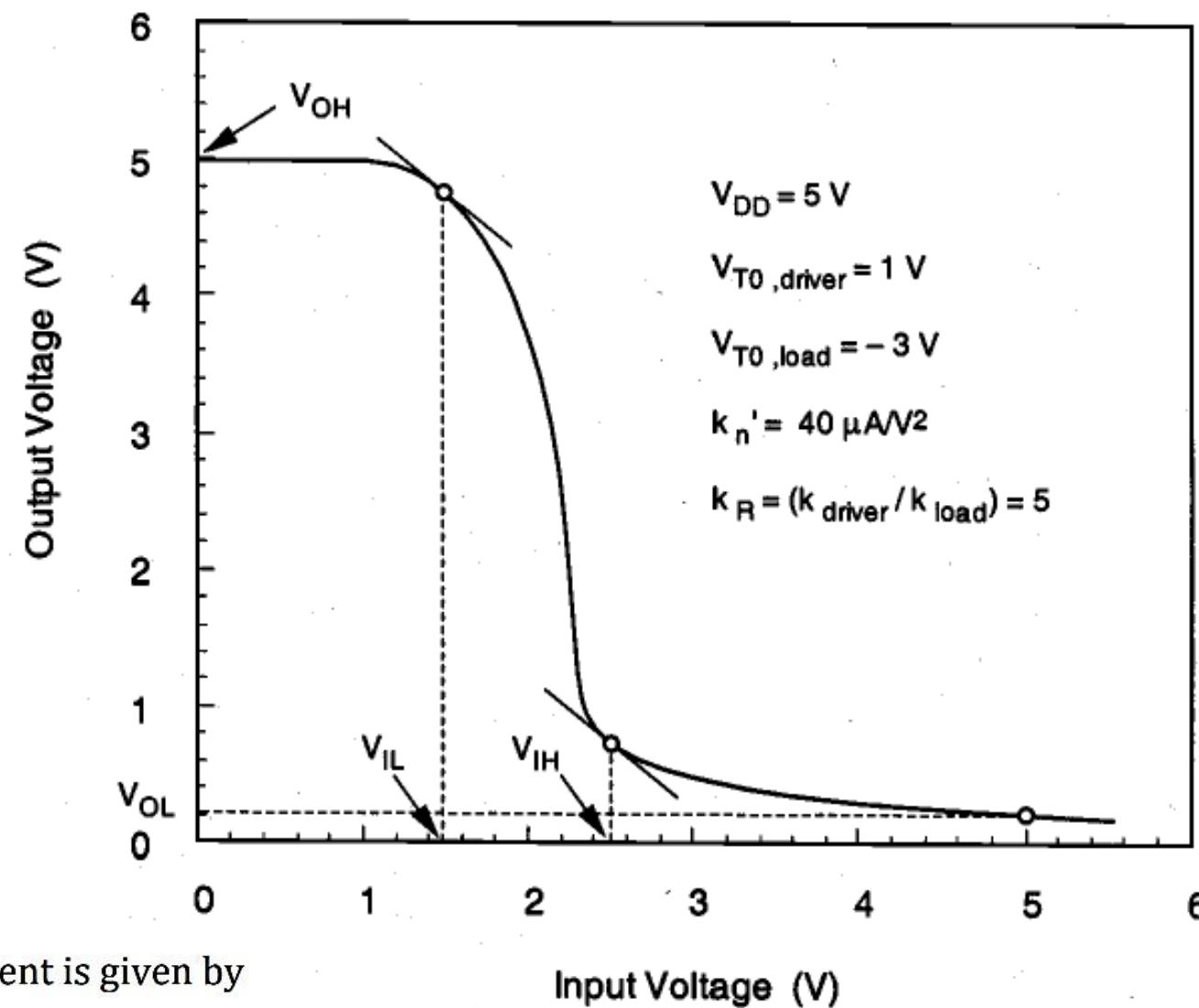
Depletion Load NMOS Inverter



At this condition the load will always be in ON state.

Depletion Load NMOS Inverter

V_{in}	V_{out}	Driver operating region	Load operating region
V_{OL}	V_{OH}	cut-off	linear
V_{IL}	$\approx V_{OH}$	saturation	linear
V_{IH}	small	linear	<i>saturation</i>
V_{OH}	V_{OL}	linear	saturation



- When the load transistor is in saturation region, the load current is given by

$$I_{D,load} = \frac{k_{n,load}}{2} [-V_{T,load}(V_{out})]$$

- When the load transistor is in linear region, the load current is given by

$$I_{D,load} = \frac{k_{n,load}}{2} [2|V_{T,load}(V_{out})| \cdot (V_{DD} - V_{out}) - (V_{DD} - V_{out})^2]$$

Depletion Load NMOS Inverter

Calculation of V_{OH}

When the input voltage V_{in} is smaller than the driver threshold voltage V_{T0} , the driver transistor is turned off and does not conduct any drain current. Consequently, the load device, which operates in the linear region, also has zero drain current.

Substituting V_{OH} for V_{out}

$$I_{D,load} = \frac{k_{n,load}}{2} \cdot \left[2|V_{T,load}(V_{OH})| \cdot (V_{DD} - V_{OH}) - (V_{DD} - V_{OH})^2 \right] = 0$$

The only valid solution in the linear region is $V_{OH} = V_{DD}$.

Depletion Load NMOS Inverter

Calculation of V_{OL}

We assume that the input voltage V_{in} of the inverter is equal to $V_{OH} = V_{DD}$

Driver transistor → linear region

Depletion-type load → saturation region

$$\frac{k_{driver}}{2} \cdot [2 \cdot (V_{OH} - V_{T0}) \cdot V_{OL} - V_{OL}^2] = \frac{k_{load}}{2} \cdot [-V_{T,load}]^2$$

This second-order equation in V_{OL} can be solved by temporarily neglecting the dependence of $V_{T,load}$ on V_{OL} , as follows.

$$V_{OL} = V_{OH} - V_{T0} - \sqrt{(V_{OH} - V_{T0})^2 - \left(\frac{k_{load}}{k_{driver}} \right) \cdot |V_{T,load}|^2}$$

Depletion Load NMOS Inverter

Calculation of V_{IL}

By definition, the slope of the VTC is equal to (-1) , i.e., $dV_{out}/dV_{in} = -1$ when the input voltage is $V_{in} = V_{IL}$. Note that in this case, the driver transistor operates in saturation while the load transistor operates in the linear region. Applying KCL for the output node, we obtain the following current equation:

$$\frac{k_{driver}}{2} \cdot (V_{in} - V_{T0})^2 = \frac{k_{load}}{2} \cdot [2|V_{T,load}| \cdot (V_{DD} - V_{out}) - (V_{DD} - V_{out})^2] \quad (5.39)$$

To satisfy the derivative condition at V_{IL} , we differentiate both sides of (5.39) with respect to V_{in} :

$$\begin{aligned} k_{driver} \cdot (V_{in} - V_{T0}) &= \frac{k_{load}}{2} \cdot \left[2|V_{T,load}| \left(-\frac{dV_{out}}{dV_{in}} \right) \right. \\ &\quad \left. + 2(V_{DD} - V_{out}) \left(-\frac{dV_{T,load}}{dV_{in}} \right) - 2(V_{DD} - V_{out}) \left(-\frac{dV_{out}}{dV_{in}} \right) \right] \end{aligned} \quad (5.40)$$

Depletion Load NMOS Inverter

Calculation of V_{IL}

In general, we can assume that the term $(dV_{T,load}/dV_{in})$ is negligible with respect to the others. Substituting V_{IL} for V_{in} , and letting $dV_{out}/dV_{in} = -1$, we obtain V_{IL} as a function of the output voltage V_{out} .

$$V_{IL} = V_{T0} + \left(\frac{k_{load}}{k_{driver}} \right) \cdot [V_{out} - V_{DD} + |V_{T,load}|] \quad (5.41)$$

Depletion Load NMOS Inverter

Calculation of V_{IH}

V_{IH} is the larger of the two voltage points on the VTC at which the slope is equal to (-1) . Since the output voltage corresponding to this operating point is relatively small, the driver transistor is in the linear region and the load transistor is in saturation.

$$\frac{k_{driver}}{2} \cdot [2 \cdot (V_{in} - V_{T0}) \cdot V_{out} - V_{out}^2] = \frac{k_{load}}{2} \cdot [-V_{T,load}]^2 \quad (5.42)$$

Differentiating both sides of (5.42) with respect to V_{in} , we obtain :

Depletion Load NMOS Inverter

Calculation of V_{IH}

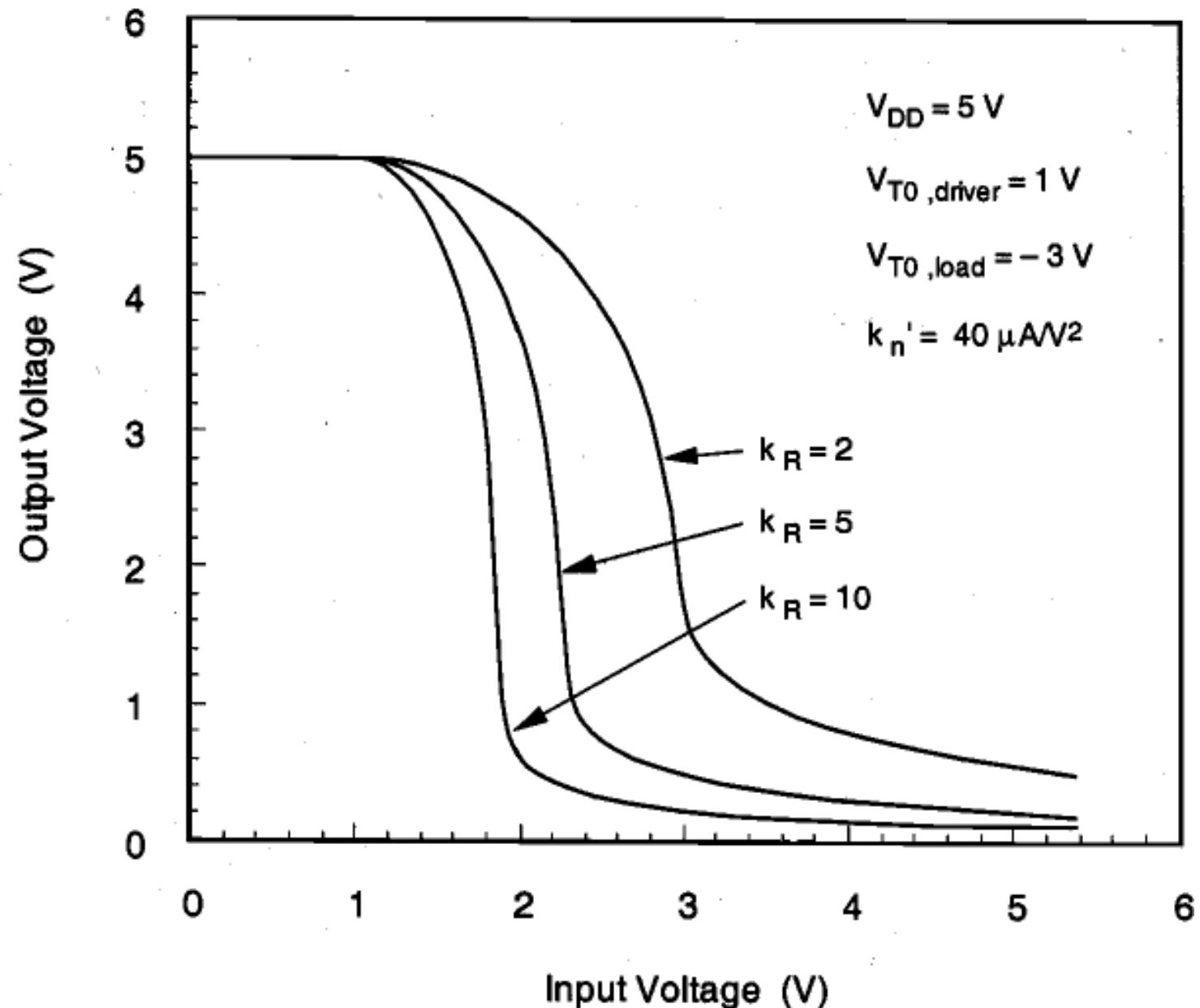
$$\begin{aligned} & k_{driver} \cdot \left[V_{out} + (V_{in} - V_{T0}) \left(\frac{dV_{out}}{dV_{in}} \right) - V_{out} \left(\frac{dV_{out}}{dV_{in}} \right) \right] \\ & = k_{load} \cdot [-V_{T,load}] \cdot \left(\frac{dV_{T,load}}{dV_{out}} \right) \cdot \left(\frac{dV_{out}}{dV_{in}} \right) \end{aligned} \quad (5.43)$$

Now, substitute $dV_{out} / dV_{in} = -1$ into (5.43), and solve for $V_{in} = V_{IH}$.

$$V_{IH} = V_{T0} + 2V_{out} + \left(\frac{k_{load}}{k_{driver}} \right) \cdot [-V_{T,load}] \cdot \left(\frac{dV_{T,load}}{dV_{out}} \right) \quad (5.44)$$

Note that the derivative of the load threshold voltage with respect to the output voltage cannot be neglected in this case.

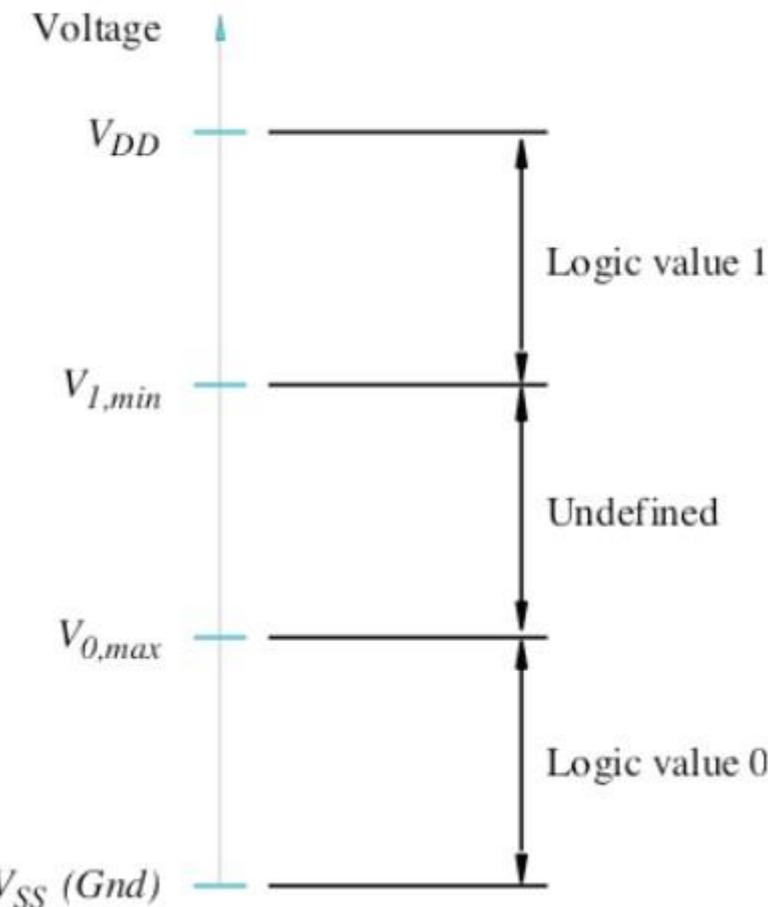
$$\frac{dV_{T,load}}{dV_{out}} = \frac{\gamma}{2\sqrt{|2\phi_F| + V_{out}}} \quad (5.45)$$



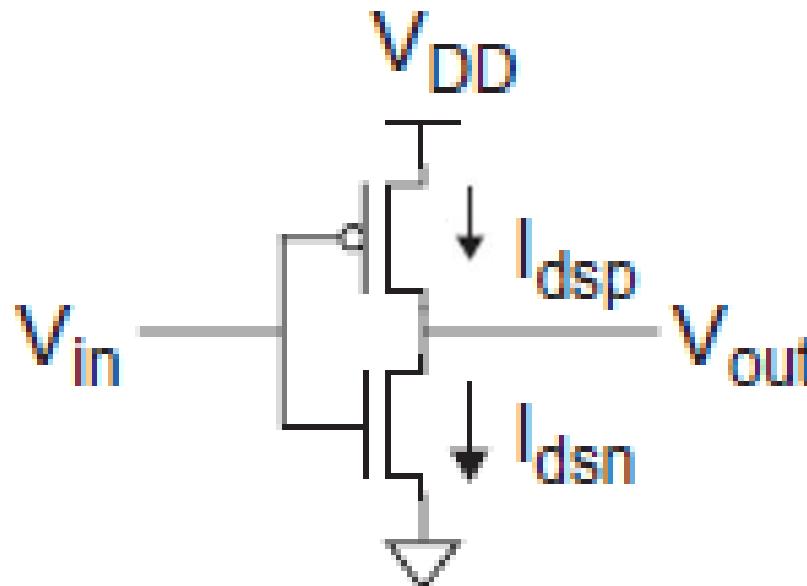
Pass Transistors

Ref:-CMOS LOGIC CIRCUIT DESIGN by John P. Uyemura

Transistor as a Switch



CMOS Inverter



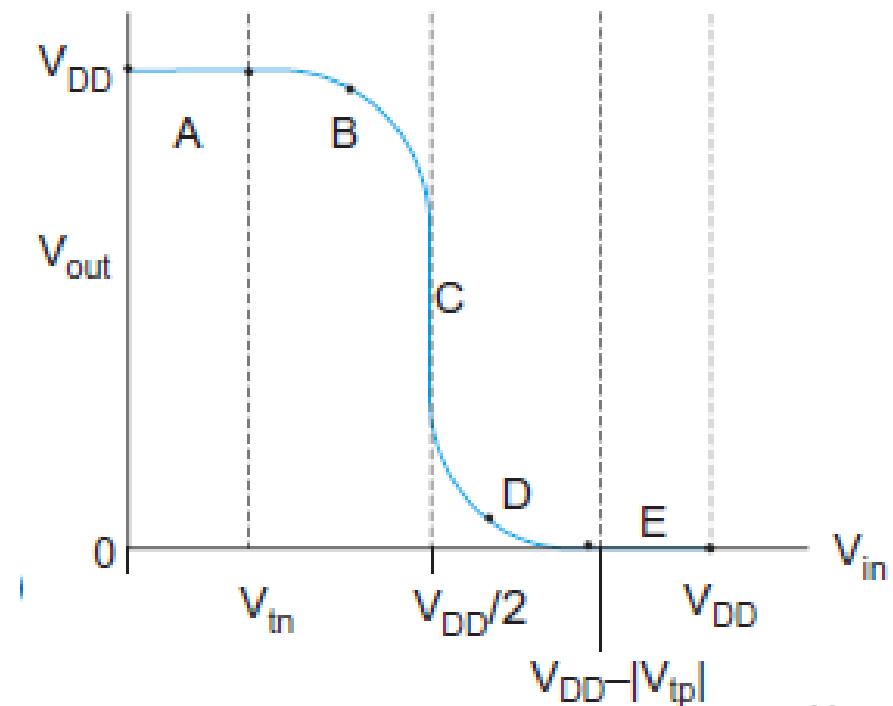
Region	V_{in}	V_{out}	nMOS	pMOS
A	$< V_{T0,n}$	V_{OH}	cut-off	linear
B	V_{IL}	high $\approx V_{OH}$	saturation	linear
C	V_{th}	V_{th}	saturation	saturation
D	V_{IH}	low $\approx V_{OL}$	linear	saturation
E	$> (V_{DD} + V_{T0,p})$	V_{OL}	linear	cut-off

V_{OH} : Maximum output voltage when the output level is logic "1"

V_{OL} : Minimum output voltage when the output level is logic "0"

V_{IL} : Maximum input voltage which can be *interpreted* as logic "0"

V_{IH} : Minimum input voltage which can be *interpreted* as logic "1"



Circuit operation

$$V_{GS,n} = V_{in}$$

$$V_{DS,n} = V_{OUT}$$

$$V_{GS,P} = V_{in} - V_{DD}$$

$$V_{DS,P} = V_{out} - V_{DD}$$

The nMOS operates in the saturation region if $V_{in} > V_{TO}$ and if following conditions are satisfied.

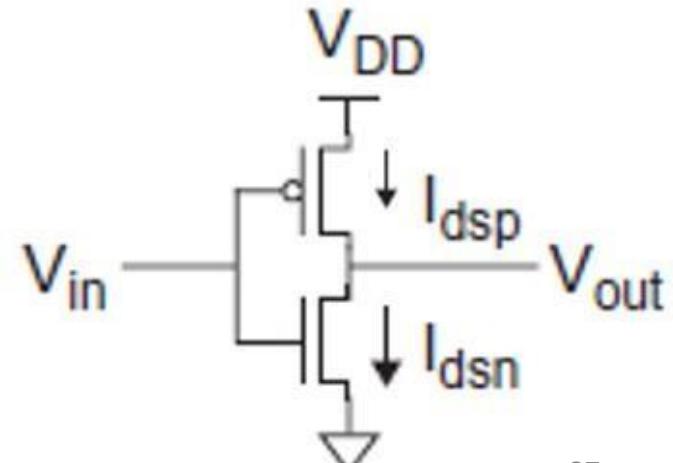
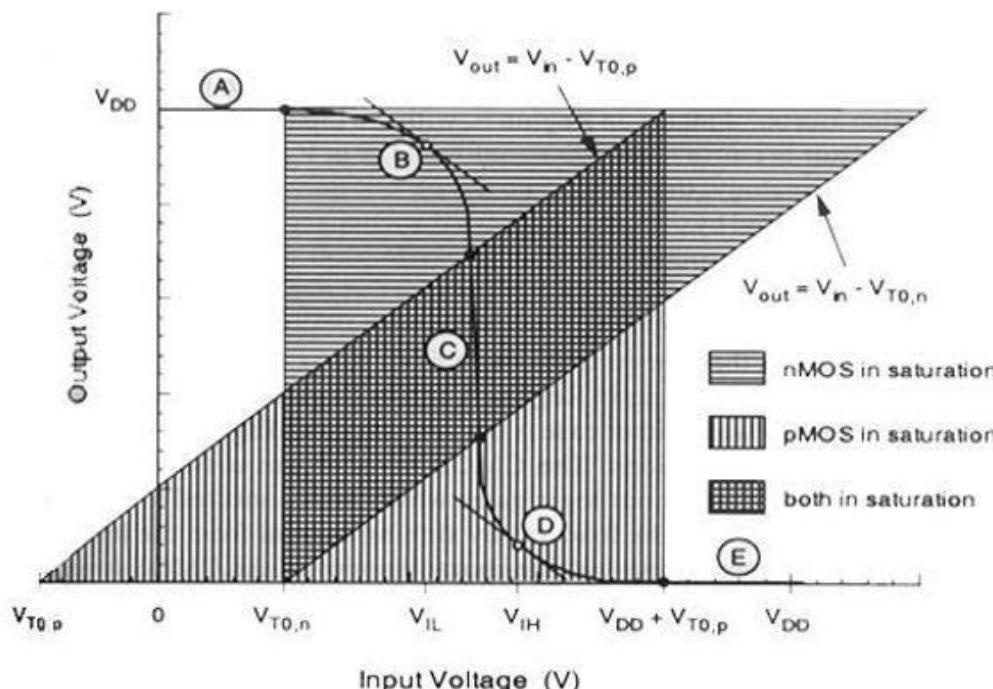
$$V_{DS,n} \geq V_{GS,n} - V_{TO,n}$$

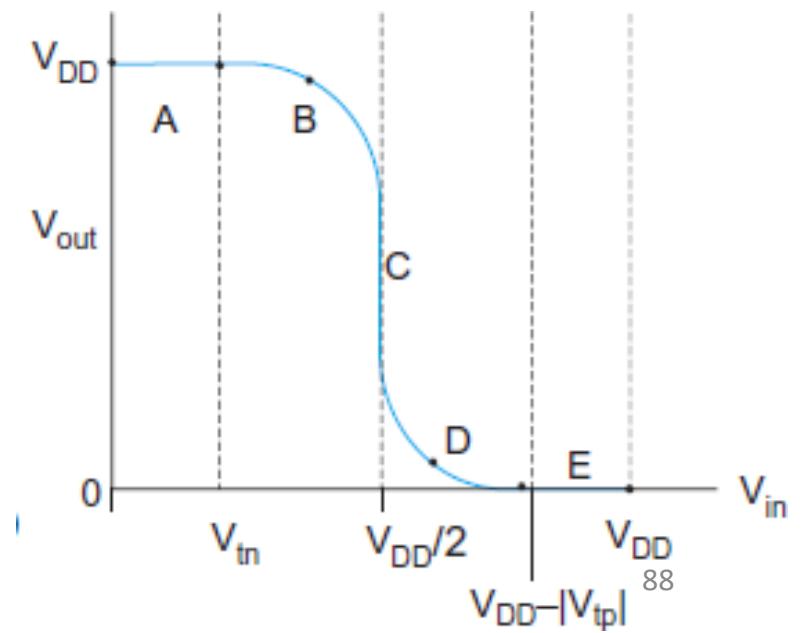
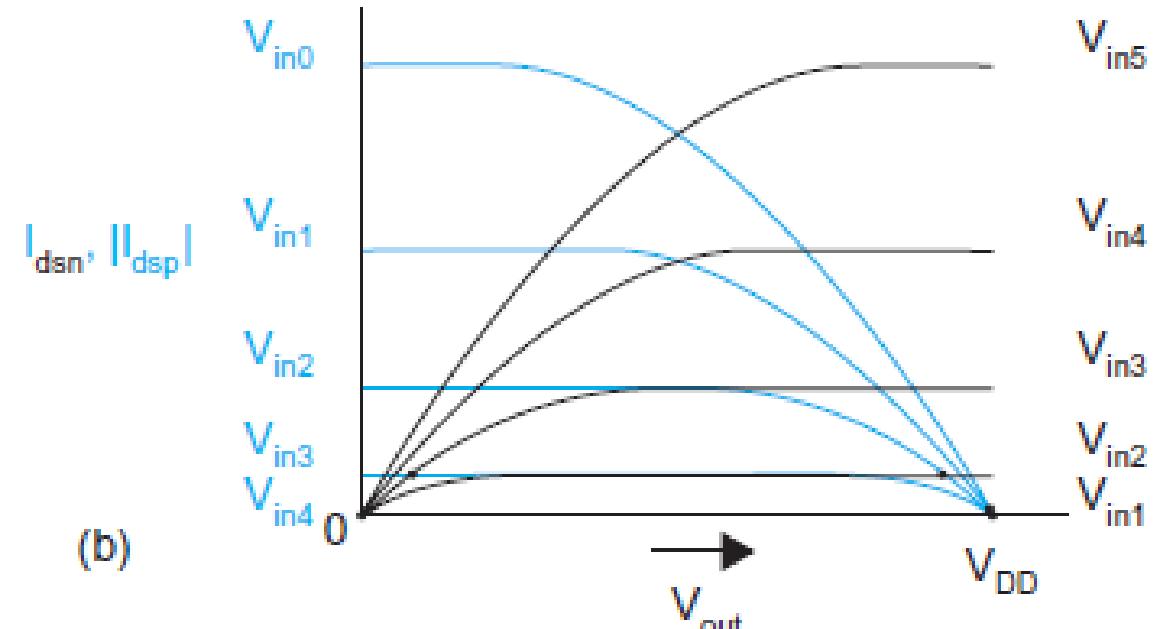
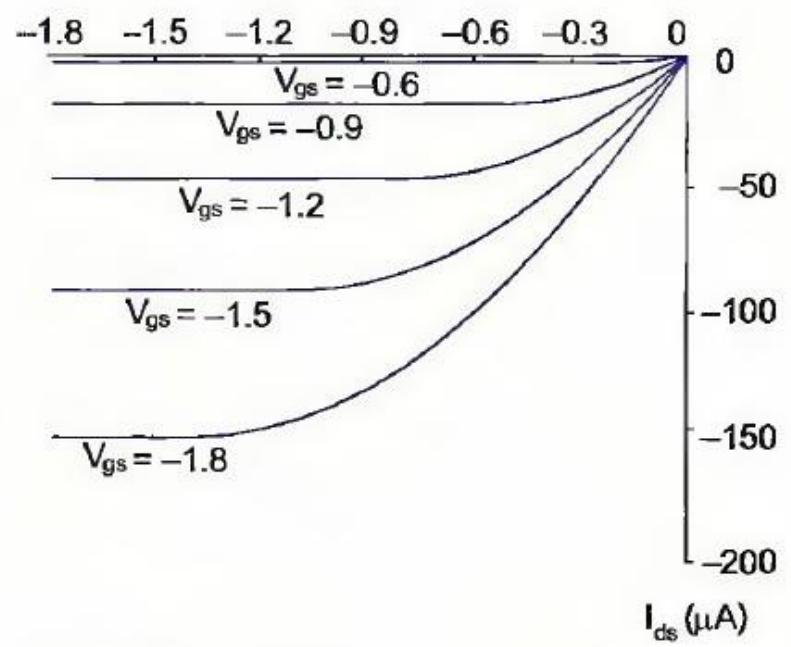
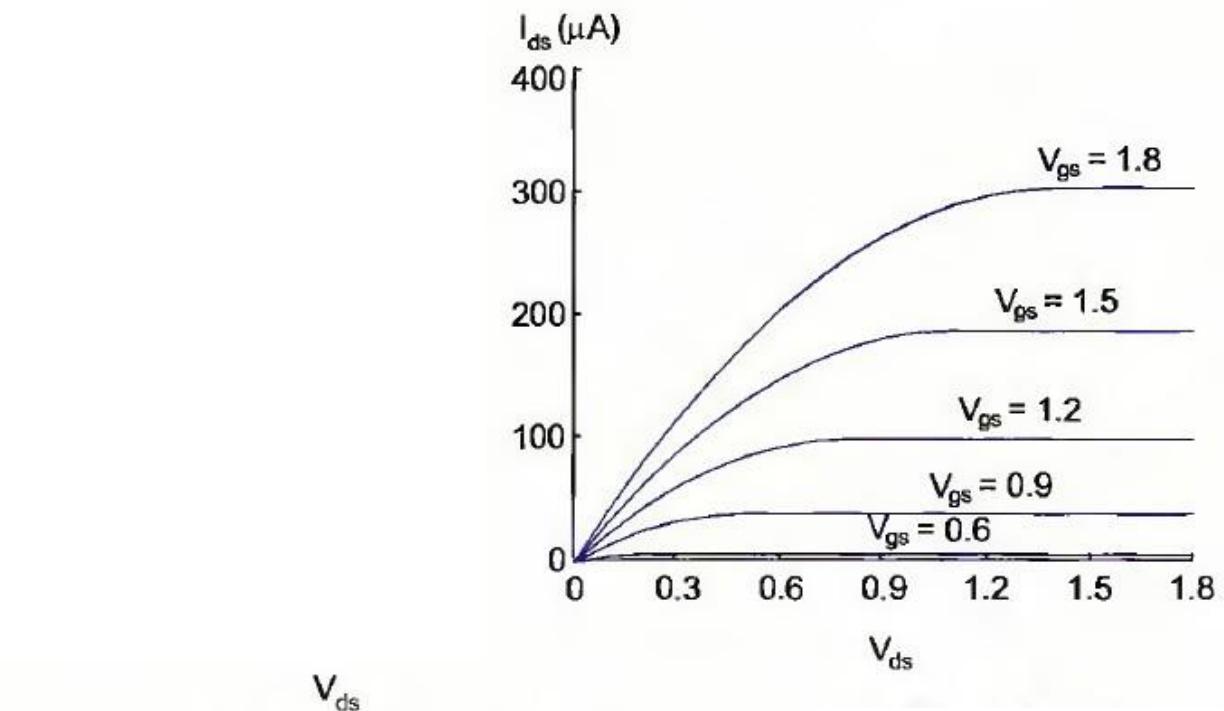
$$V_{out} \geq V_{in} - V_{TO,n}$$

The pMOS operates in the saturation region if $V_{in} < V_{DD} + V_{TO,p}$ and if following conditions are satisfied.

$$V_{DS,P} \leq V_{GS,P} - V_{TO,P}$$

$$V_{out} \geq V_{in} - V_{TO,P}$$





Governing Conditions

	Cutoff	Linear	Saturated
nMOS	$V_{gsn} < V_{tn}$	$V_{gsn} > V_{tn}$	$V_{gsn} > V_{tn}$
	$V_{in} < V_{tn}$	$V_{in} > V_{tn}$	$V_{in} > V_{tn}$
		$V_{dsn} < V_{gsn} - V_{tn}$	$V_{dsn} > V_{gsn} - V_{tn}$
		$V_{out} < V_{in} - V_{tn}$	$V_{out} > V_{in} - V_{tn}$
pMOS	$V_{gsp} > V_{tp}$	$V_{gsp} < V_{tp}$	$V_{gsp} < V_{tp}$
	$V_{in} > V_{tp} + V_{DD}$	$V_{in} < V_{tp} + V_{DD}$	$V_{in} < V_{tp} + V_{DD}$
		$V_{dsp} > V_{gsp} - V_{tp}$	$V_{dsp} < V_{gsp} - V_{tp}$
		$V_{out} > V_{in} - V_{tp}$	$V_{out} < V_{in} - V_{tp}$

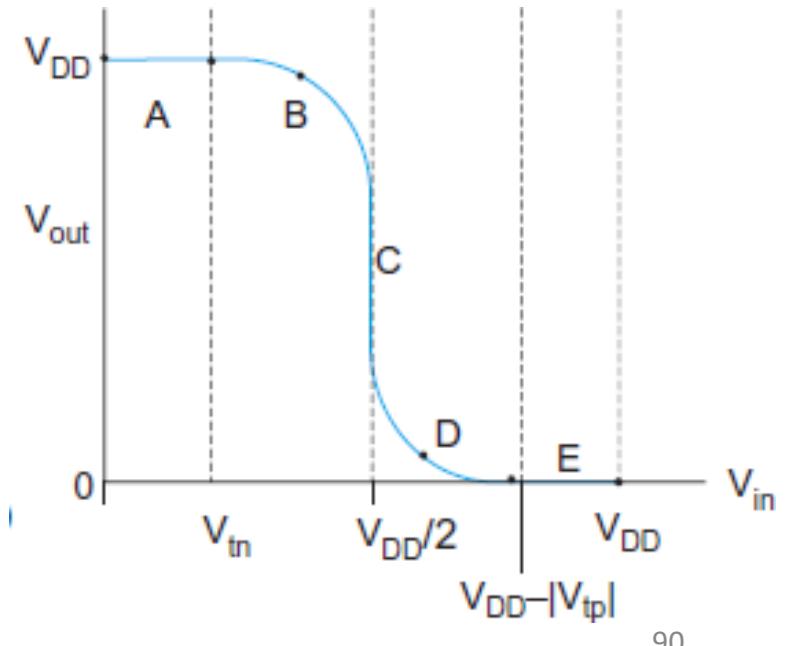
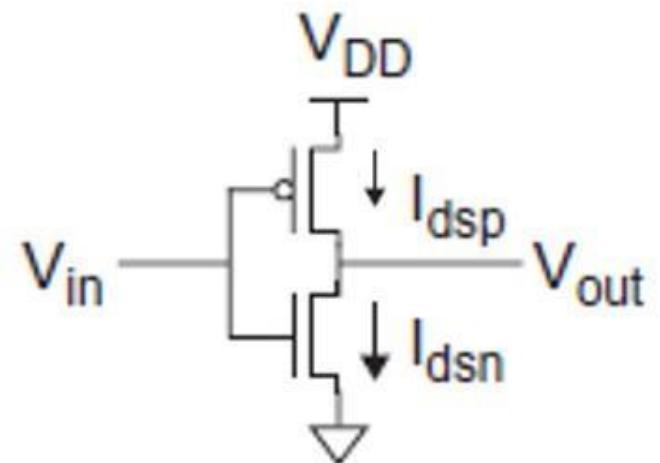
Region A

- PMOS – Non-saturation Region
- NMOS - Cutoff

$$V_{out} = V_{DD} - I_D R_{CPMOS}$$

$$V_{out} = V_{DD}$$

Region	V_{in}	V_{out}	nMOS	pMOS
A	$< V_{T0,n}$	V_{OH}	cut-off	linear
B	V_{IL}	high $\approx V_{OH}$	saturation	linear
C	V_{th}	V_{th}	saturation	saturation
D	V_{IH}	low $\approx V_{OL}$	linear	saturation
E	$> (V_{DD} + V_{T0,p})$	V_{OL}	linear	cut-off



Region B $\rightarrow V_{IL}$

PMOS – Non-saturation Region

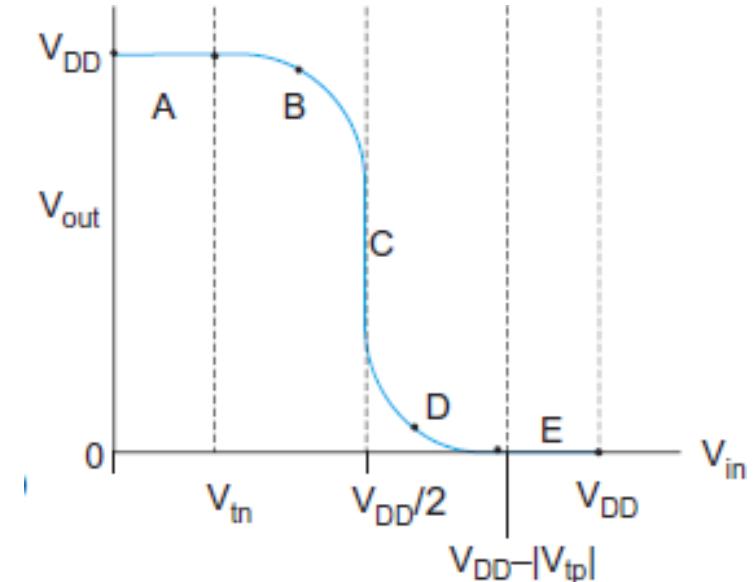
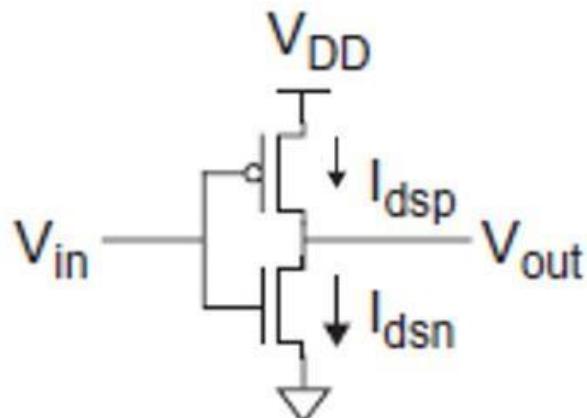
NMOS – Saturation Region

The slope of the VTC is equal to (-1),
when the input voltage is $V = V_{IL}$.

$$\frac{k_n}{2} \cdot (V_{GS,n} - V_{T0,n})^2 = \frac{k_p}{2} \cdot [2 \cdot (V_{GS,p} - V_{T0,p}) \cdot V_{DS,p} - V_{DS,p}^2]$$

$$\frac{k_n}{2} \cdot (V_{in} - V_{T0,n})^2 = \frac{k_p}{2} \cdot [2 \cdot (V_{in} - V_{DD} - V_{T0,p})$$

$$\cdot (V_{out} - V_{DD}) - (V_{out} - V_{DD})^2]$$



$$V_{GS,n} = V_{in}$$

$$V_{DS,n} = V_{OUT}$$

$$V_{GS,p} = V_{in} - V_{DD}$$

$$V_{DS,p} = V_{out} - V_{DD}$$

Region	V_{in}	V_{out}	nMOS	pMOS
A	$< V_{T0,n}$	V_{OH}	cut-off	linear
B	V_{IL}	high $\approx V_{OH}$	saturation	linear
C	V_{th}	V_{th}	saturation	saturation
D	V_{IH}	low $\approx V_{OL}$	linear	saturation
E	$> (V_{DD} + V_{T0,p})$	V_{OL}	linear	cut-off

To satisfy the derivative condition at V_{IL} we differentiate both sides with respect to V_{in} .

$$k_n \cdot (V_{in} - V_{T0,n}) = k_p \cdot \left[(V_{in} - V_{DD} - V_{T0,p}) \cdot \left(\frac{dV_{out}}{dV_{in}} \right) + (V_{out} - V_{DD}) \cdot \left(\frac{dV_{out}}{dV_{in}} \right) \right]$$

Substituting $V_{in} = V_{IL}$ and $(dV_{out}/dV_{in}) = -1$ in (5.60), we obtain

$$k_n \cdot (V_{IL} - V_{T0,n}) = k_p \cdot (2V_{out} - V_{IL} + V_{T0,p} - V_{DD})$$

$$V_{IL} = \frac{2V_{out} + V_{T0,p} - V_{DD} + k_R V_{T0,n}}{1 + k_R}$$

Where

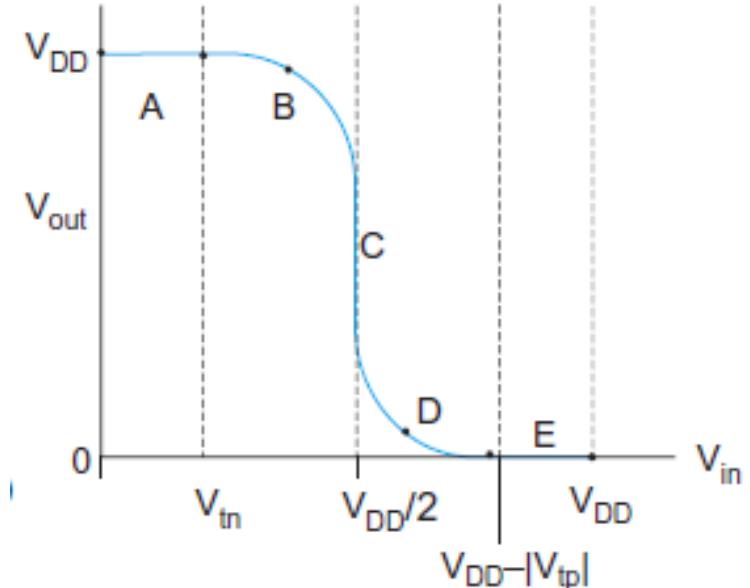
$$k_R = \frac{k_n}{k_p}$$

Region-C $\rightarrow V_{th}$

PMOS – Saturation Region

NMOS – Saturation Region

$$\frac{k_n}{2} \cdot (V_{GS,n} - V_{T0,n})^2 = \frac{k_p}{2} \cdot (V_{GS,p} - V_{T0,p})^2$$



Replacing $V_{GS,n}$ and $V_{GS,p}$ in (5.68) according to (5.51) and (5.52), we obtain

$$\frac{k_n}{2} \cdot (V_{in} - V_{T0,n})^2 = \frac{k_p}{2} \cdot (V_{in} - V_{DD} - V_{T0,p})^2$$

$$V_{GS,n} = V_{in}$$

$$V_{DS,n} = V_{OUT}$$

$$V_{GS,p} = V_{in} - V_{DD}$$

$$V_{DS,p} = V_{out} - V_{DD}$$

The correct solution for V_{in} for this equation is

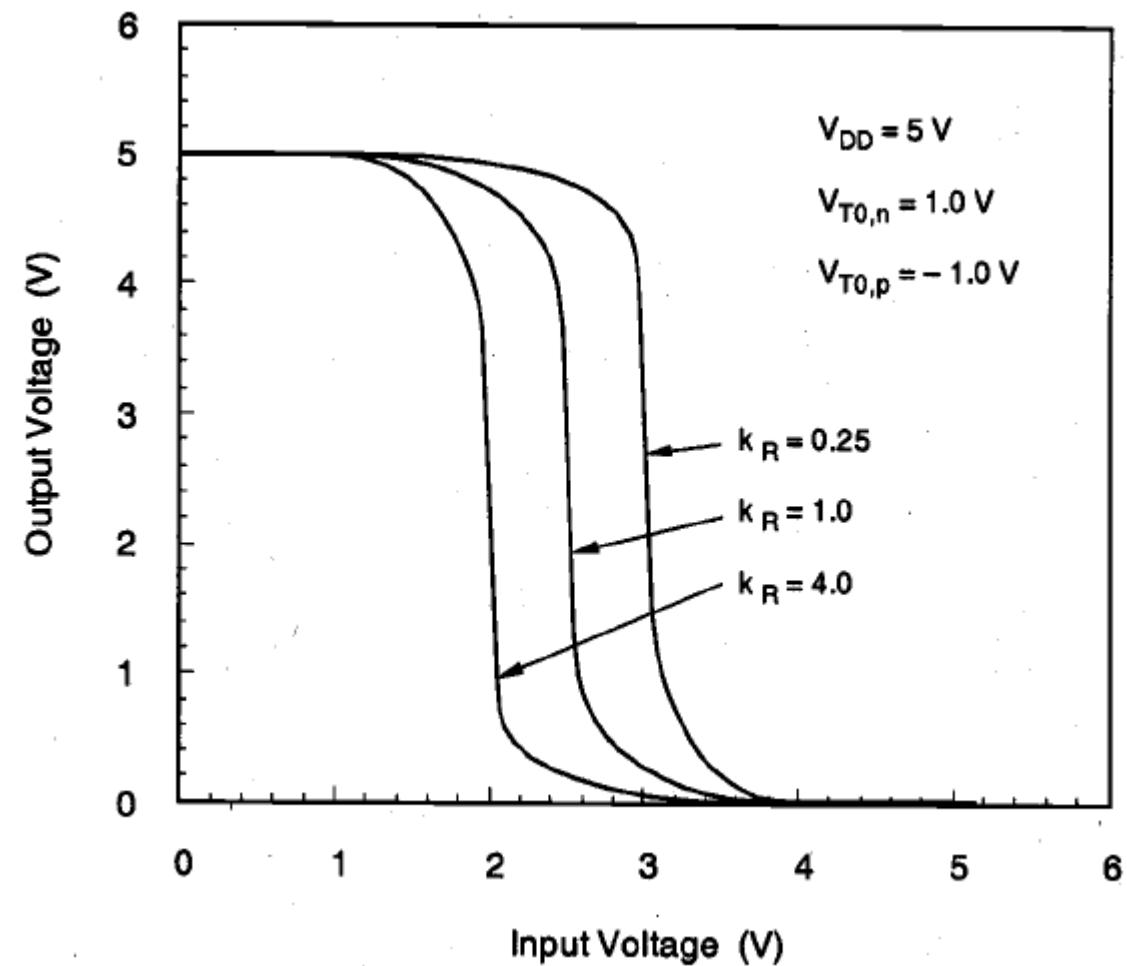
$$V_{in} \cdot \left(1 + \sqrt{\frac{k_p}{k_n}}\right) = V_{T0,n} + \sqrt{\frac{k_p}{k_n}} \cdot (V_{DD} + V_{T0,p})$$

Region	V_{in}	V_{out}	nMOS	pMOS
A	$< V_{T0,n}$	V_{OH}	cut-off	linear
B	V_{IL}	high $\approx V_{OH}$	saturation	linear
C	V_{th}	V_{th}	saturation	saturation
D	V_{IH}	low $\approx V_{OL}$	linear	saturation
E	$> (V_{DD} + V_{T0,p})$	V_{OL}	linear	cut-off

$$V_{th} = \frac{V_{T0,n} + \sqrt{\frac{1}{k_R} \cdot (V_{DD} + V_{T0,p})}}{\left(1 + \sqrt{\frac{1}{k_R}}\right)}$$

Where

$$k_R = \frac{k_n}{k_p}$$



Region-D \rightarrow V_{IH}

PMOS – Saturation Region

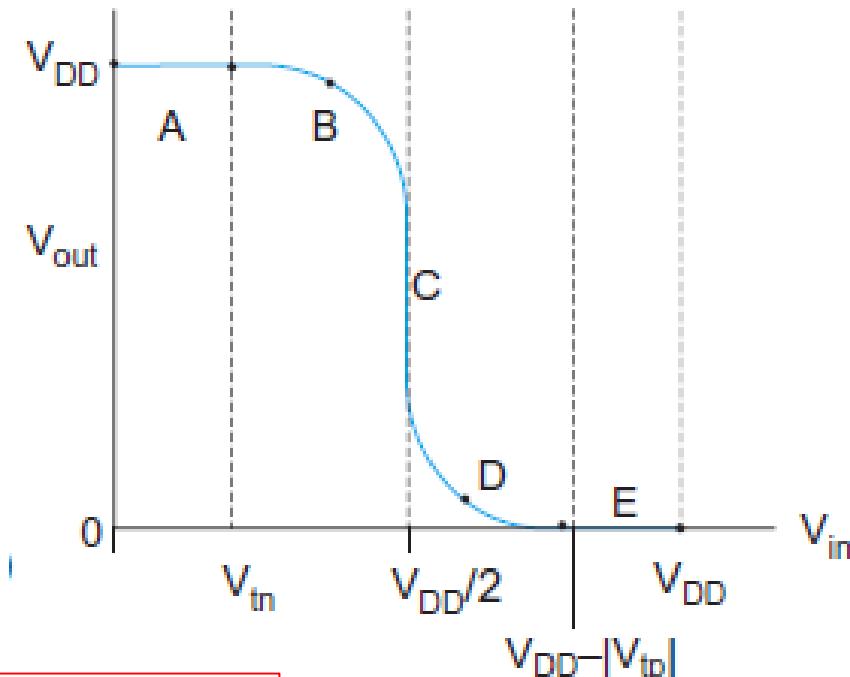
NMOS – Non-saturation Region

$$\frac{k_n}{2} \cdot [2 \cdot (V_{GS,n} - V_{T0,n}) \cdot V_{DS,n} - V_{DS,n}^2] = \frac{k_p}{2} \cdot (V_{GS,p} - V_{T0,p})^2$$

$$\frac{k_n}{2} \cdot [2 \cdot (V_{in} - V_{T0,n}) \cdot V_{out} - V_{out}^2] = \frac{k_p}{2} \cdot (V_{in} - V_{DD} - V_{T0,p})^2$$

Now, differentiate both sides of (5.64) with respect to V_{in} .

$$k_n \cdot \left[(V_{in} - V_{T0,n}) \cdot \left(\frac{dV_{out}}{dV_{in}} \right) + V_{out} - V_{out} \cdot \left(\frac{dV_{out}}{dV_{in}} \right) \right] = k_p \cdot (V_{in} - V_{DD} - V_{T0,p})$$



$$V_{GS,n} = V_{in}$$

$$V_{DS,n} = V_{OUT}$$

$$V_{GS,p} = V_{in} - V_{DD}$$

$$V_{DS,p} = V_{out} - V_{DD}$$

Region	V_{in}	V_{out}	nMOS	pMOS
A	$< V_{T0,n}$	V_{OH}	cut-off	linear
B	V_{IL}	high $\approx V_{OH}$	saturation	linear
C	V_{th}	V_{th}	saturation	saturation
D	V_{IH}	low $\approx V_{OL}$	linear	saturation
E	$> (V_{DD} + V_{T0,p})$	V_{OL}	linear	cut-off

Substituting $V_{in} = V_{IH}$ and $(dV_{out}/dV_{in}) = -1$

$$k_n \cdot (-V_{IH} + V_{T0,n} + 2V_{out}) = k_p \cdot (V_{IH} - V_{DD} - V_{T0,p})$$

$$V_{IH} = \frac{V_{DD} + V_{T0,p} + k_R \cdot (2V_{out} + V_{T0,n})}{1 + k_R}$$

Where

$$k_R = \frac{k_n}{k_p}$$

Region-E $\rightarrow V_{OL}$

when the input voltage exceeds VDD the pMOS transistor is turned off. In this case, the nMOS transistor is operating in the linear region, but its drain to- source voltage is equal to zero because

$$I_{D,n} = I_{D,p} = 0$$

The output voltage of the circuit is

$$V_{out} = V_{OL} = 0$$

Region	V_{in}	V_{out}	nMOS	pMOS
A	$< V_{T0,n}$	V_{OH}	cut-off	linear
B	V_{IL}	high $\approx V_{OH}$	saturation	linear
C	V_{th}	V_{th}	saturation	saturation
D	V_{IH}	low $\approx V_{OL}$	linear	saturation
E	$> (V_{DD} + V_{T0,p})$	V_{OL}	linear	cut-off

Non-ideal I-V effects

1. Velocity saturation and mobility degradation.
2. Channel length modulation
3. Body effect
4. Subthreshold conduction
5. Junction leakage.
6. Tunneling
7. Temperature dependence
8. Geometry dependence

1. Velocity saturation and mobility degradation

- At high lateral field strengths (V_{ds}/L), carrier velocity ceases to increase linearly with field strength.
- This is called **velocity saturation**.

$$v = \begin{cases} \frac{\mu_{\text{eff}} E}{1 + \frac{E}{E_c}} & E < E_c \\ v_{\text{sat}} & E \geq E_c \end{cases}$$

Where E is electric field between the drain and source.
The ***critical electric*** field is

$$E_c = \frac{2v_{\text{sat}}}{\mu_{\text{eff}}}$$

- At strong vertical electric fields resulting from large V_{gs} causes the carrier to scatter against the surface and also reduce the carrier mobility.
- This effect is called **mobility degradation**.

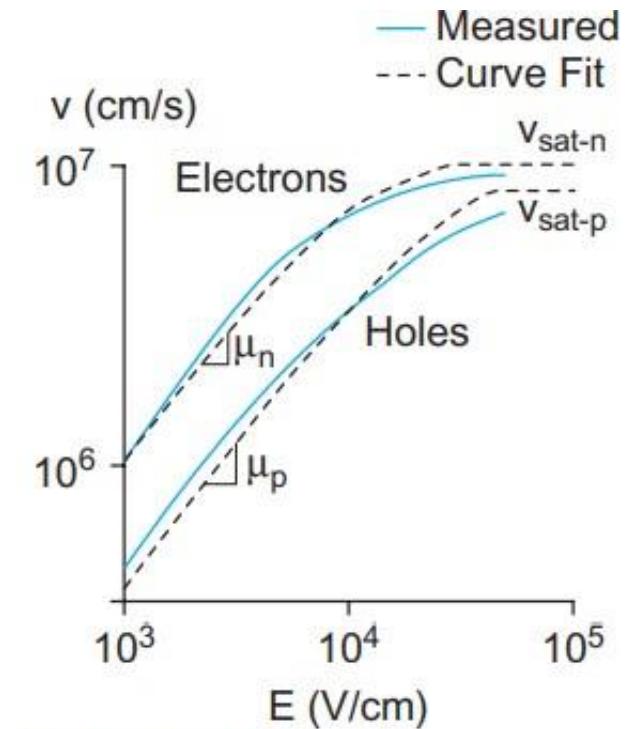


FIGURE 2.15 Carrier velocity vs. electric field at 300 K, adapted from [Jacoboni77]. Velocity saturates at high fields.

$$\mu_{\text{eff}-n} = \frac{540 \frac{\text{cm}^2}{\text{V} \cdot \text{s}}}{1 + \left(\frac{V_{gs} + V_t}{0.54 \frac{\text{V}}{\text{nm}} t_{\text{ox}}} \right)^{1.85}}$$

$$\mu_{\text{eff}-p} = \frac{185 \frac{\text{cm}^2}{\text{V} \cdot \text{s}}}{1 + \left| \frac{V_{gs} + 1.5V_t}{0.338 \frac{\text{V}}{\text{nm}} t_{\text{ox}}} \right|^{1.85}}$$

2. Channel length modulation

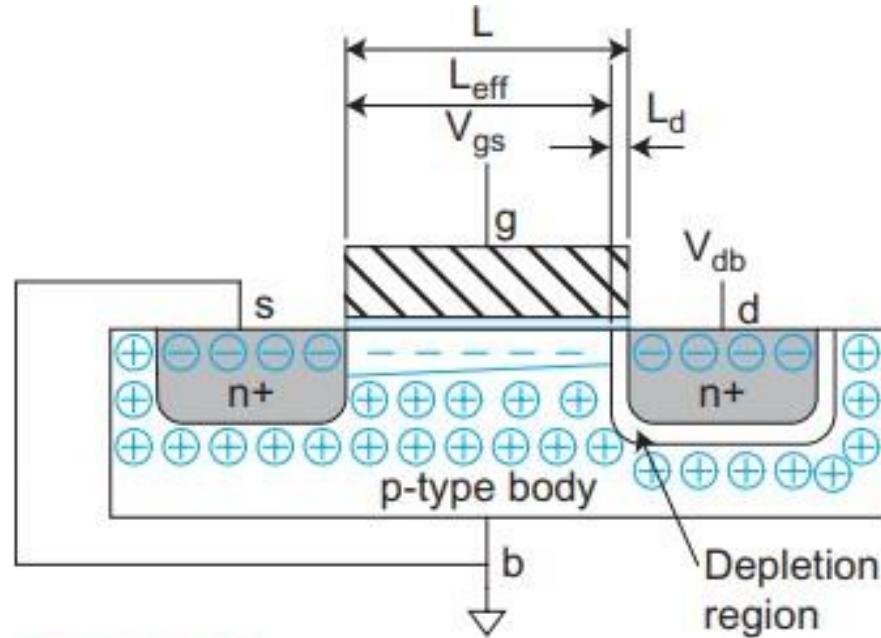


FIGURE 2.18 Depletion region shortens effective channel length

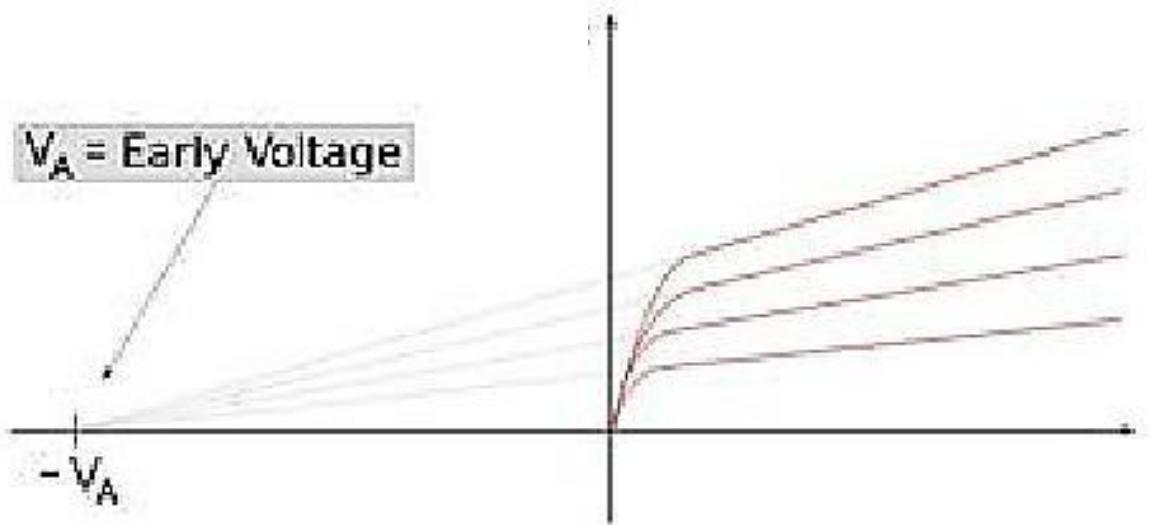
Increasing V_{ds} decreases the effective channel length. Shorter channel length results in higher current; thus, I_{ds} increases with V_{ds} in saturation

$$L_{\text{eff}} = L - L_d$$

$$I_{ds} = \frac{\beta}{2} V_{GT}^2 \left(1 + \frac{V_{ds}}{V_A} \right)$$

where V_A is called the Early voltage

The term $V_{gs} - V_t$ arises so often that it is convenient to abbreviate it as V_{GT}



3. Body effect

- When a voltage V_{sb} is applied between the source and body, it increases the amount of charge required to invert the channel, hence, it increases the threshold voltage.
- The potential difference between source and body V_{sb} affects the threshold voltage.
- The threshold voltage is given by,

$$V_t = V_{t0} + \gamma \left(\sqrt{\phi_s + V_{sb}} - \sqrt{\phi_s} \right) \quad (2.30)$$

$$\phi_s = 2v_T \ln \frac{N_A}{n_i} \quad (2.31)$$

$$\gamma = \frac{t_{ox}}{\epsilon_{ox}} \sqrt{2q\epsilon_{si} N_A} = \frac{\sqrt{2q\epsilon_{si} N_A}}{C_{ox}} \quad (2.32)$$

Where V_{t0} is threshold voltage when S is at body potential

ϕ_s is the surface potential at threshold.

γ is the *body effect coefficient*, typically in the range 0.4 to 1 V^{1/2}.

N_A is doping level of the substrate, n_i is the doping level of the wells

At room temperature the thermal voltage $v_T = kT/q = 26$ mV :

4. Sub threshold conduction

- Ideally nMOS will not allow the current to flow through it when $V_{gs} < V_t$.
- But in real transistors, the current doesn't cut off below threshold.
- This conduction is also known as leakage and flows even though the MOS is OFF.

$$I_{ds} = I_{ds0} e^{\frac{V_{gs}-V_{t0}+\eta V_{ds}-k\gamma V_{sb}}{nv_T} \left(1 - e^{\frac{-V_{ds}}{v_T}} \right)}$$

- I_{ds} is Subthreshold leakage current
 - I_{ds0} is the current at threshold
 - n is a process-dependent term affected by the depletion region characteristics typically in the range of 1.3–1.7 for CMOS processes
 - The final term indicates that leakage is 0 if $V_{ds} = 0$,
 - v_T is thermal voltage
- $$I_{ds0} = \beta v_T^2 e^{1.8}$$
- $$\beta = \mu C_{ox} \frac{W}{L}$$

5. Junction leakage

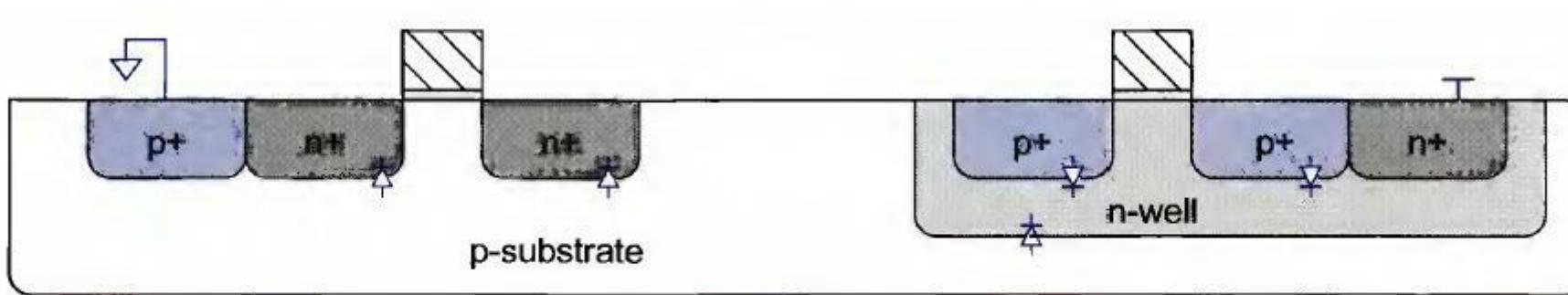


FIG 2.19 Reverse-biased diodes in CMOS circuits

diodes still conduct a small amount of current I_D .

$$I_D = I_s \left(e^{\frac{V_D}{v_T}} - 1 \right)$$

V_D is the diode voltage (e.g., $-V_{sb}$ or $-V_{db}$)

I_s depends on doping levels and on the area and perimeter of the diffusion region

6. Tunnelling

- Two physical mechanisms for gate tunneling are called **Fowler-Nordheim (FN) tunneling and direct tunneling**.
- FN tunneling is most important at high voltage and moderate oxide thickness.
- Direct tunneling is most important at lower voltage with thin oxides and is the dominant leakage component. It is given by the formula,

$$I_{\text{gate}} = WA \left(\frac{V_{DD}}{t_{\text{ox}}} \right)^2 e^{-B \frac{t_{\text{ox}}}{V_{DD}}}$$

where A and B are technology constants.

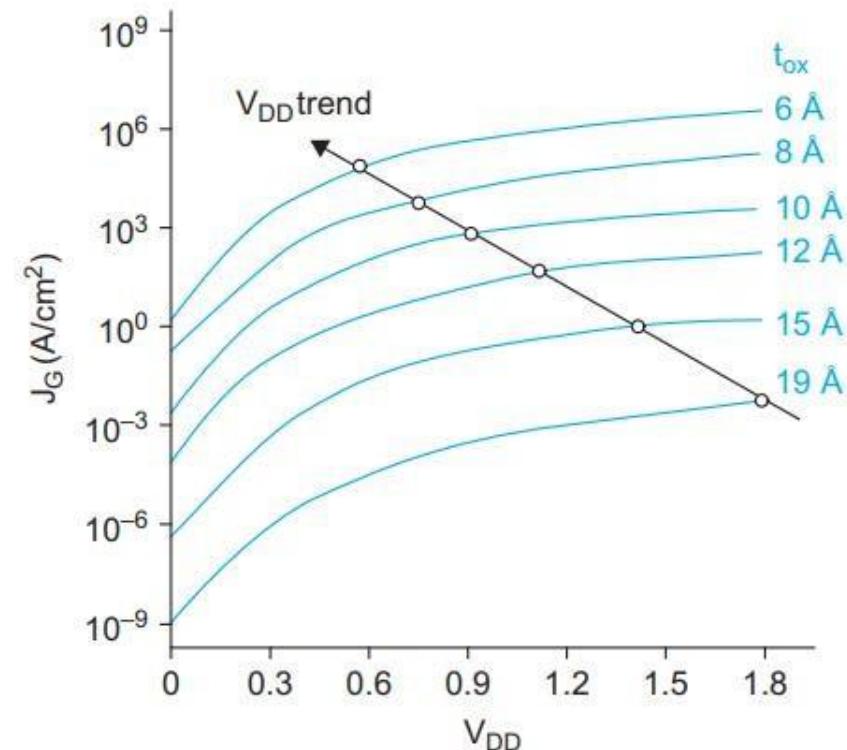


FIGURE 2.21 Gate leakage current from [Song01]

7. Temperature dependence

- At low V_{gs} , the current has a positive temperature coefficient.
- At high V_{gs} , the current has a negative temperature coefficient; i.e., it decreases with temperature.

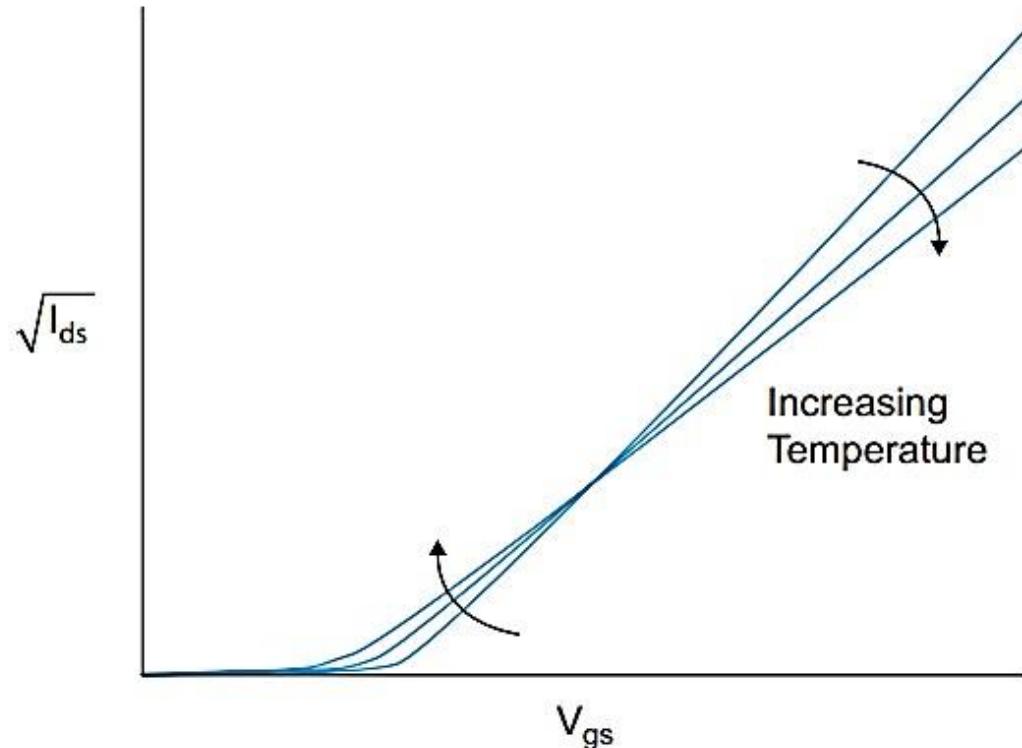


FIGURE 2.23 I-V characteristics of nMOS transistor in saturation at various temperatures

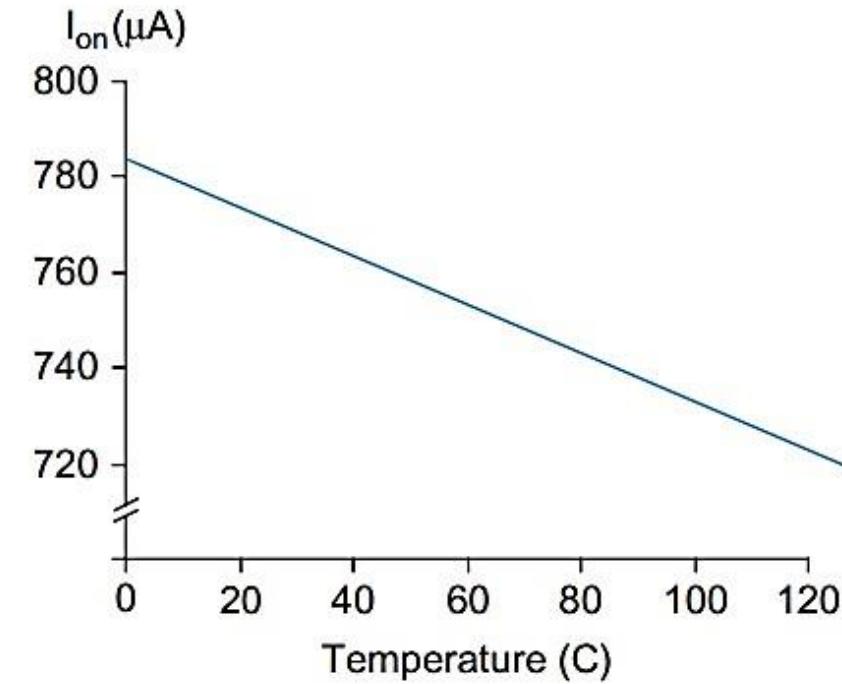


FIGURE 2.24 I_{dsat} vs. temperature

8. Geometry dependence

- The layout designer draws transistors with width and length W_{drawn} and L_{drawn} .
- The actual gate dimensions may differ by some factors X_W and X_L .
- Moreover, the source and drain tend to diffuse laterally under the gate by L_D , producing a shorter effective channel length.
- Similarly, W_D accounts for other effects that shrink the transistor width.
- Putting these factors together, we can compute effective transistor lengths and widths, as:

$$L_{\text{eff}} = L_{\text{drawn}} + X_L - 2L_D$$

$$W_{\text{eff}} = W_{\text{drawn}} + X_W - 2W_D$$

THE nMOS INVERTER

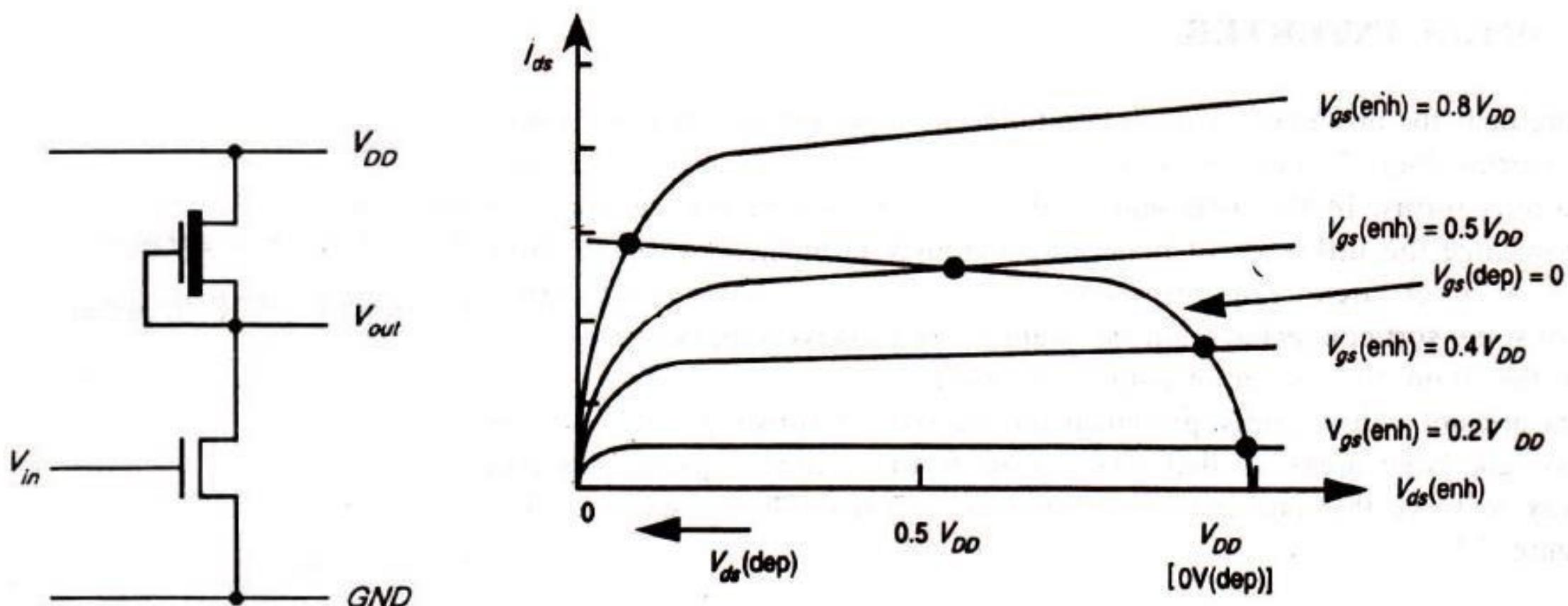


FIGURE 2.5 nMOS inverter.

$$V_{ds}(\text{enh}) = V_{DD} - V_{ds}(\text{dep}) = V_{out}$$

$V_{gs}(\text{enh}) = V_{in} \dots$ intersection points give transfer characteristic

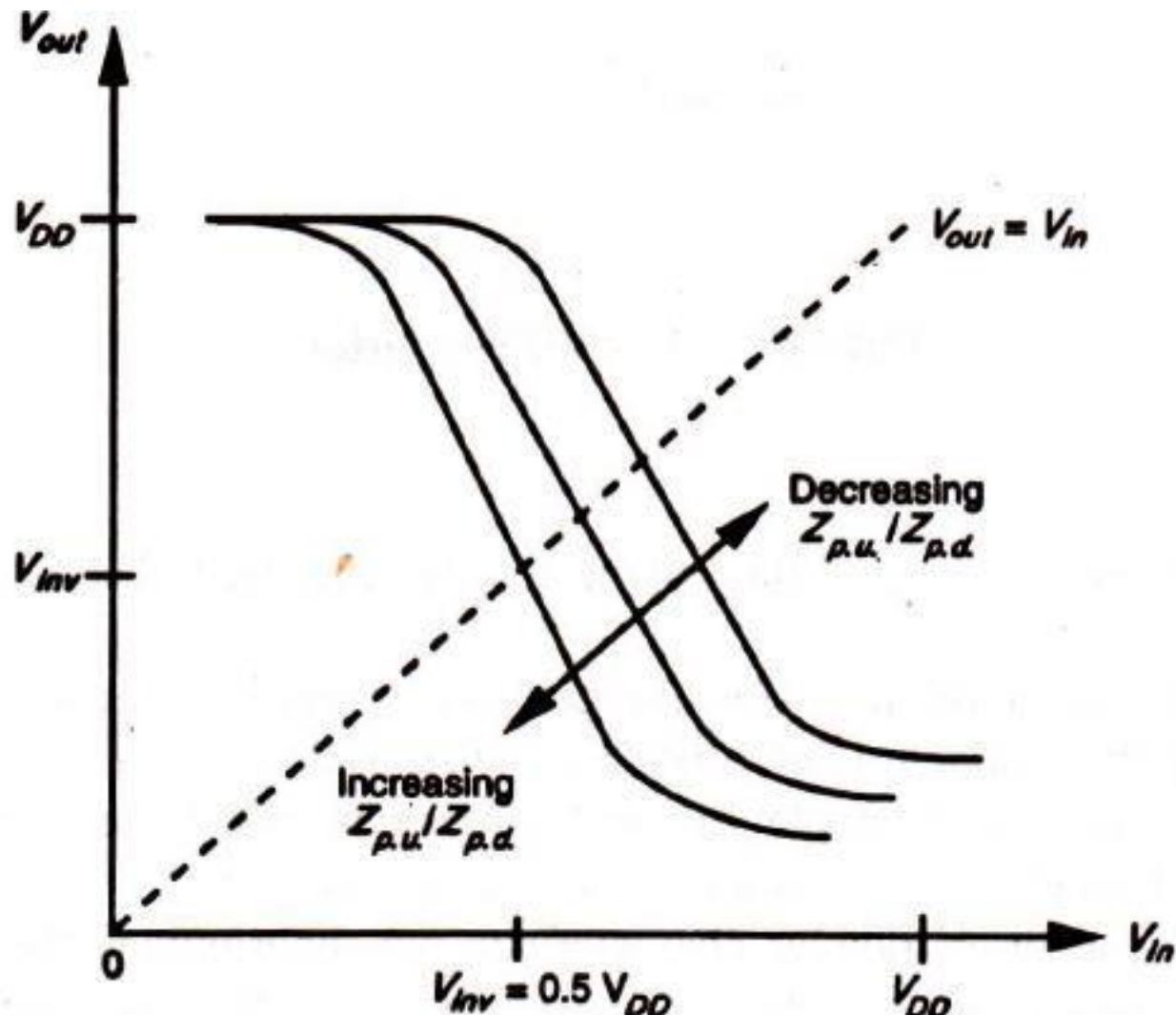


FIGURE 2.7 nMOS Inverter transfer characteristic.

$$\text{Gain} = \frac{\delta V_{out}}{\delta V_{in}}$$

Determination of PU and PD ratio (Z_{pu}/Z_{pd})

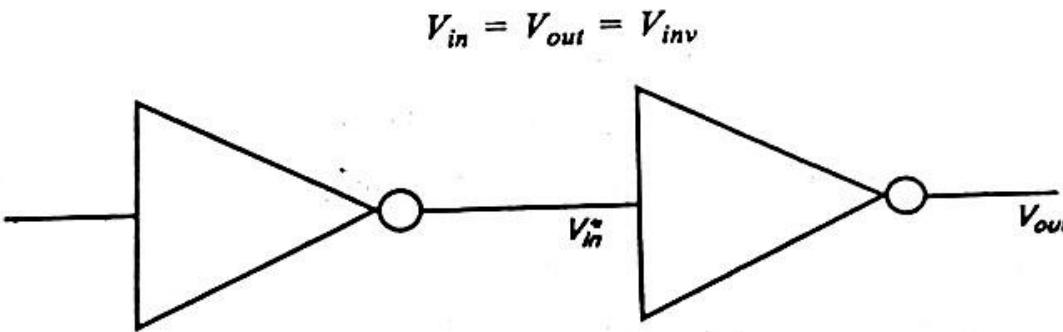


FIGURE 2.8 nMOS inverter driven directly by another inverter.

we set $V_{inv} = 0.5V_{DD}$

At this point both transistors are in saturation and

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

In the depletion mode

$$I_{ds} = K \frac{W_{p.u.}}{L_{p.u.}} \frac{(-V_{td})^2}{2} \text{ since } V_{gs} = 0$$

and in the enhancement mode

$$I_{ds} = K \frac{W_{p.d.}}{L_{p.d.}} \frac{(V_{inv} - V_t)^2}{2} \text{ since } V_{gs} = V_{inv}$$

Equating (since currents are the same) we have

$$\frac{W_{p.d.}}{L_{p.d.}} (V_{inv} - V_t)^2 = \frac{W_{p.u.}}{L_{p.u.}} (-V_{td})^2$$

Now write

$$Z_{p.d.} = \frac{L_{p.d.}}{W_{p.d.}}; Z_{p.u.} = \frac{L_{p.u.}}{W_{p.u.}}$$

we have

$$\frac{1}{Z_{p.d.}} (V_{inv} - V_t)^2 = \frac{1}{Z_{p.u.}} (-V_{td})^2$$

Now we can substitute typical values as follows:

$$V_{inv} = V_t - \frac{V_{td}}{\sqrt{Z_{p.u.}/Z_{p.d.}}}$$

thus, from equation (2.9)

$$V_t = 0.2V_{DD}; V_{td} = -0.6V_{DD}$$
$$V_{inv} = 0.5V_{DD} \text{ (for equal margins)}$$

$$0.5 = 0.2 + \frac{0.6}{\sqrt{Z_{p.u.}/Z_{p.d.}}}$$

whence

and thus

$$\sqrt{Z_{p.u.}/Z_{p.d.}} = 2$$

for an inverter directly driven by an inverter.

Z_{pu}/Z_{pd} for nMOS inverter driven through one or more pass transistors

DC transfer characteristics

1. Static CMOS inverter DC characteristics

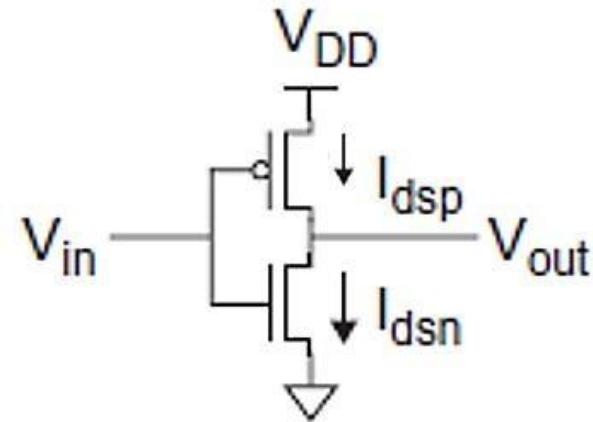
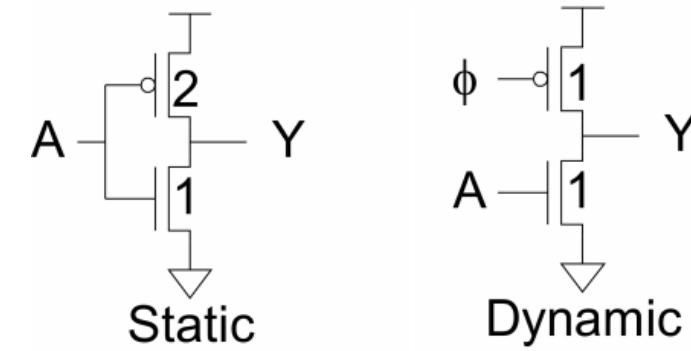


TABLE 2.2 Relationships between voltages for the three regions of operation of a CMOS inverter

	Cutoff	Linear	Saturated
nMOS	$V_{gsn} < V_{tn}$	$V_{gsn} > V_{tn}$	$V_{gsn} > V_{tn}$
	$V_{in} < V_{tn}$	$V_{in} > V_{tn}$	$V_{in} > V_{tn}$
		$V_{dsn} < V_{gsn} - V_{tn}$	$V_{dsn} > V_{gsn} - V_{tn}$
		$V_{out} < V_{in} - V_{tn}$	$V_{out} > V_{in} - V_{tn}$
pMOS	$V_{gsp} > V_{tp}$	$V_{gsp} < V_{tp}$	$V_{gsp} < V_{tp}$
	$V_{in} > V_{tp} + V_{DD}$	$V_{in} < V_{tp} + V_{DD}$	$V_{in} < V_{tp} + V_{DD}$
		$V_{dsp} > V_{gsp} - V_{tp}$	$V_{dsp} < V_{gsp} - V_{tp}$
		$V_{out} > V_{in} - V_{tp}$	$V_{out} < V_{in} - V_{tp}$

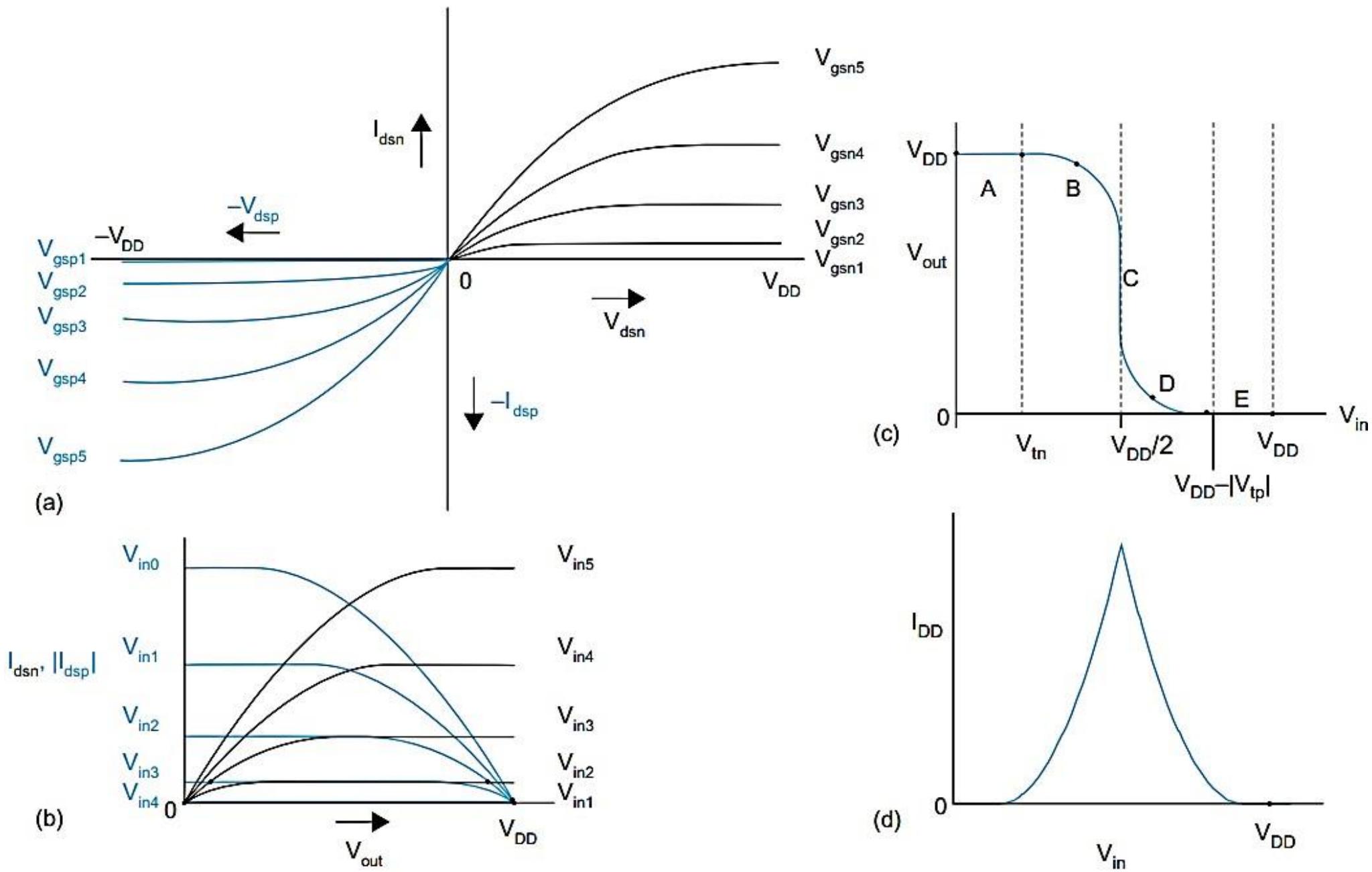


FIGURE 2.26 Graphical derivation of CMOS inverter DC characteristic

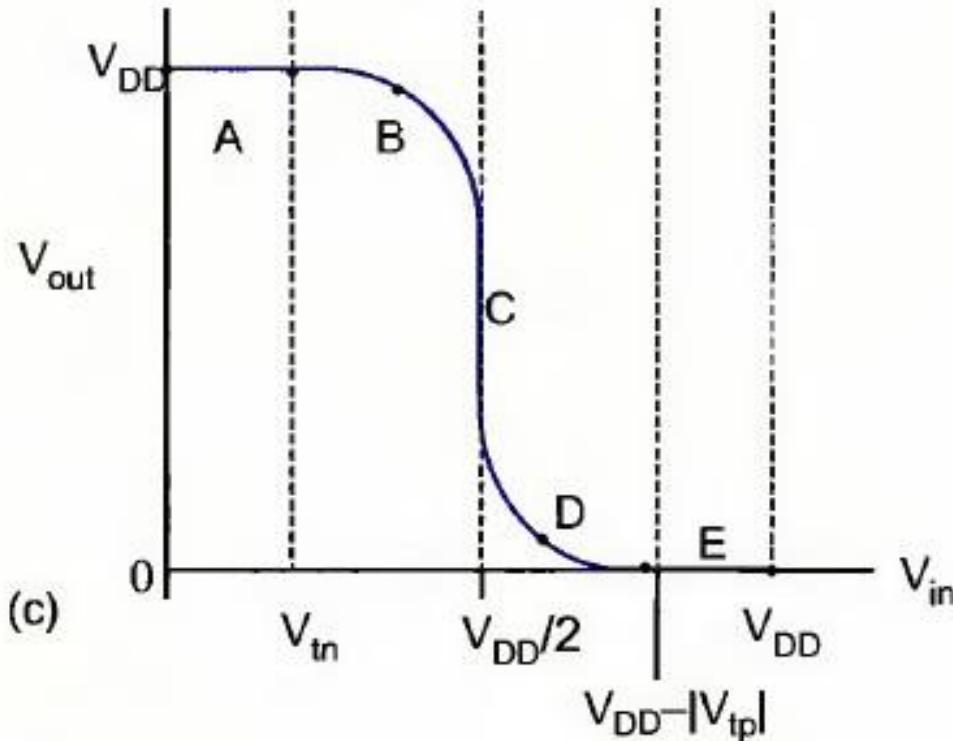


Table 2.3 Summary of CMOS inverter operation

Region	Condition	p-device	n-device	Output
A	$0 \leq V_{in} < V_{tn}$	linear	cutoff	$V_{out} = V_{DD}$
B	$V_{tn} \leq V_{in} < V_{DD}/2$	linear	saturated	$V_{out} > V_{DD}/2$
C	$V_{in} = V_{DD}/2$	saturated	saturated	V_{out} drops sharply
D	$V_{DD}/2 < V_{in} \leq V_{DD} - V_{tp} $	saturated	linear	$V_{out} < V_{DD}/2$
E	$V_{in} > V_{DD} - V_{tp} $	cutoff	linear	$V_{out} = 0$

2. Beta Ratio effects

- We have seen that for $\beta_p = \beta_n$, the inverter threshold voltage V_{inv} is $V_{DD}/2$.
- This may be desirable because it maximizes noise margins and allows a capacitive load to charge and discharge in equal times by providing equal current source and sink capabilities.
- Inverters with different beta ratios $r = \beta_p/\beta_n$ are called **skewed inverters**.
- If $r > 1$, the inverter is **HI-skewed** (stronger pMOS and weaker nMOS)
- If $r < 1$, the inverter is **LO-skewed** (weaker pMOS and stronger nMOS).
- If $r = 1$, the inverter has **normal skew** or is **unskewed**.

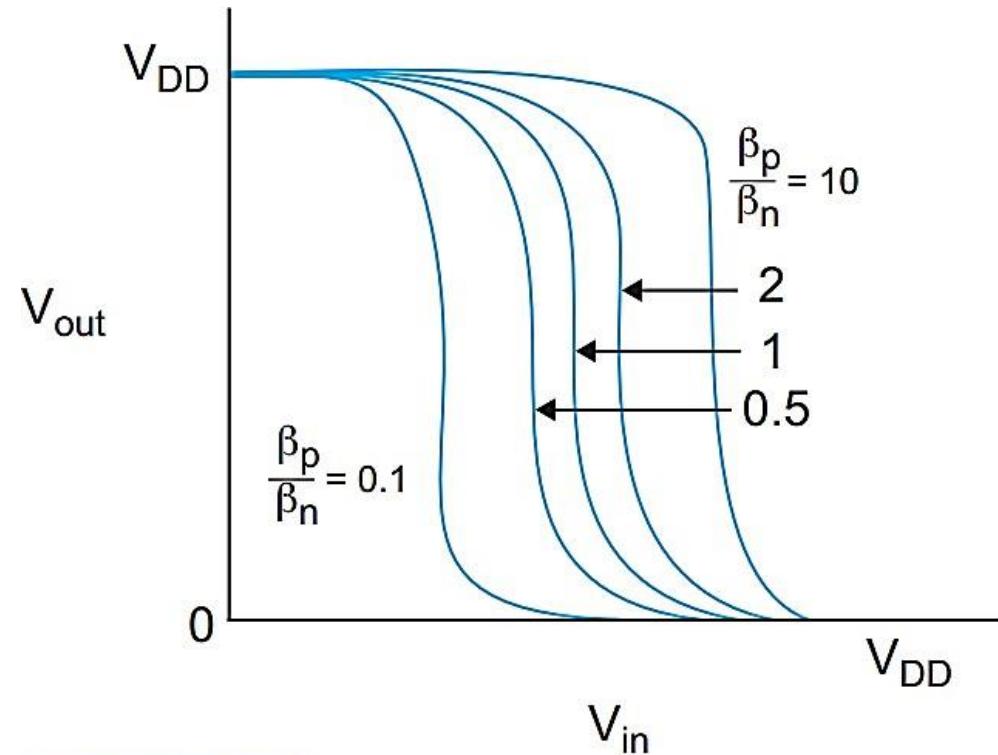


FIGURE 2.28 Transfer characteristics of skewed inverters

3. Noise margin

- Noise margin is closely related to the DC voltage characteristics.
- This parameter allows you to determine the allowable noise voltage on the input of a gate so that the output will not be corrupted.
- The specification most commonly used to describe noise margin (or noise immunity) uses two parameters: the **LOW noise margin, NM_L** , and the **HIGH noise margin, NM_H** .

- NM_L is defined as the difference in maximum LOW input voltage recognized by the receiving gate and the maximum LOW output voltage produced by the driving gate. $NM_L = V_{IL} - V_{OL}$
- The value of NM_H is the difference between the minimum HIGH output voltage of the driving gate and the minimum HIGH input voltage recognized by the receiving gate. $NM_H = V_{OH} - V_{IH}$

V_{OH} : Maximum output voltage when the output level is logic "1"

V_{OL} : Minimum output voltage when the output level is logic "0"

V_{IL} : Maximum input voltage which can be interpreted as logic "0"

V_{IH} : Minimum input voltage which can be interpreted as logic "1"

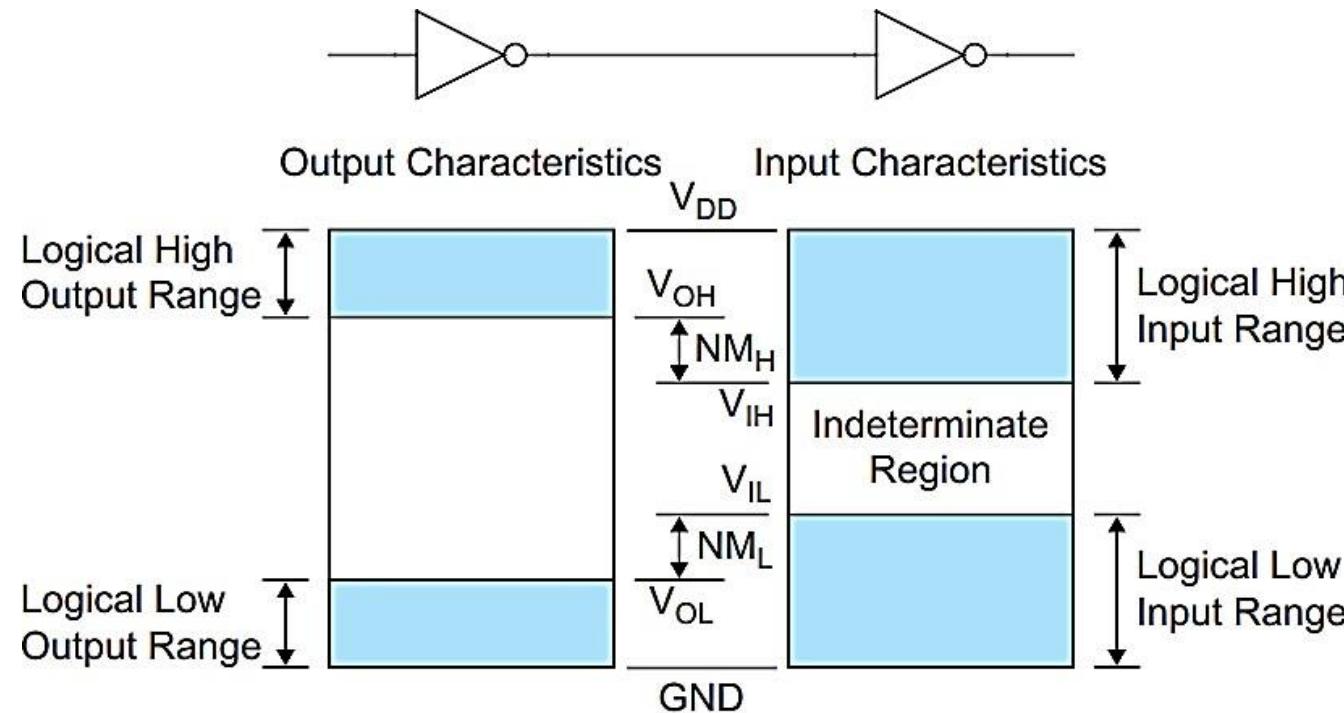


FIGURE 2.29 Noise margin definitions

4. Pass transistor DC characteristics

- a. Recall that nMOS passes “0” well, but “1” poorly.
- b. pMOS passes “1” well, but “0” poorly.
- c. Imagine that the source is initially at $V_s = 0$. $V_{gs} > V_{tn}$, so the transistor is ON and current flows. If the voltage on the source rises to $V_s = V_{DD} - V_{tn}$, V_{gs} falls to V_{tn} and the transistor cuts itself OFF. Therefore, nMOS transistors attempting to pass a 1 never pull the source above $V_{DD} - V_{tn}$. This loss is sometimes called a ***threshold drop***.

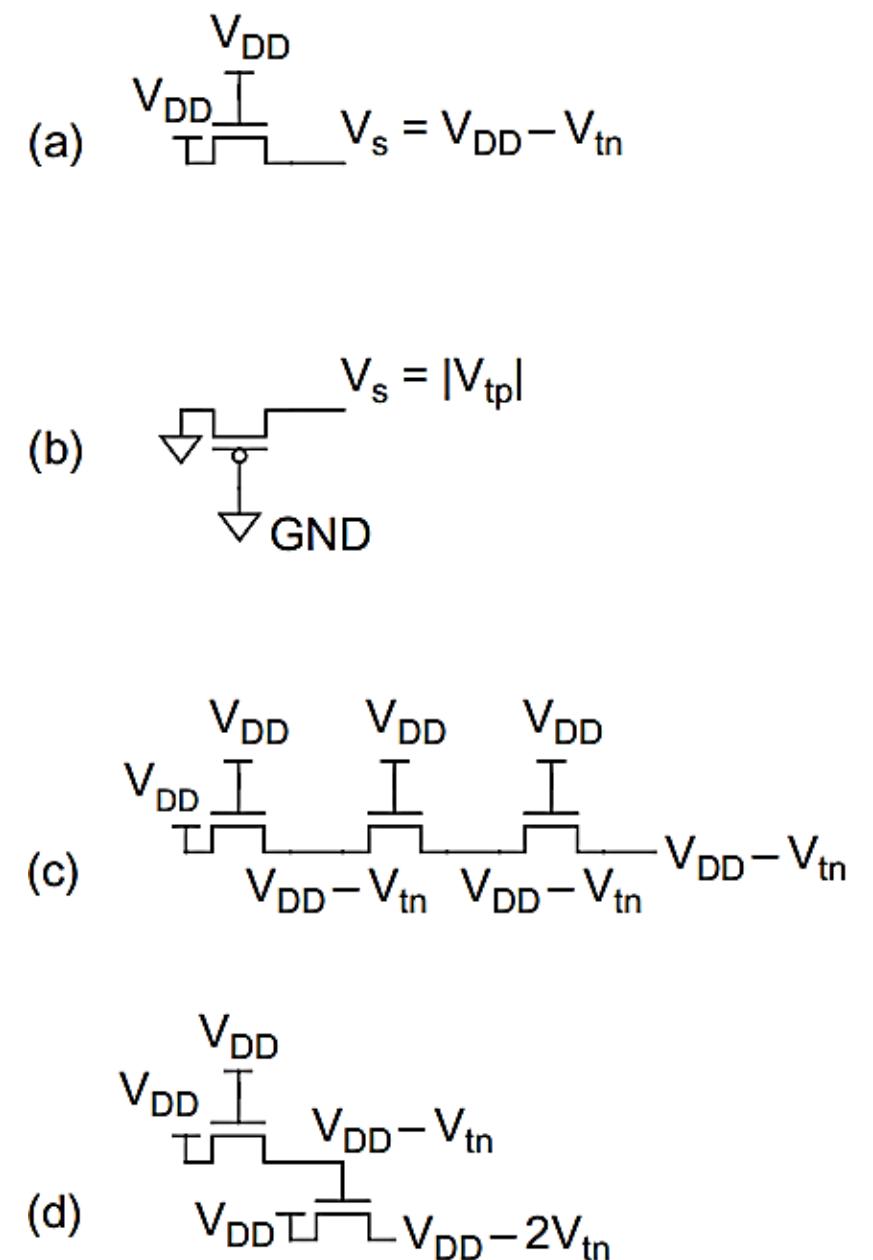


FIGURE 2.31 Pass transistor threshold drops

CMOS logic

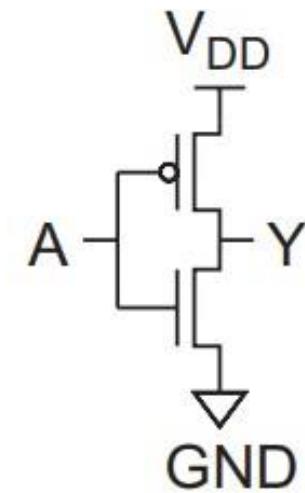
1. The Inverter
2. The NAND Gate
3. CMOS Logic Gates
4. The NOR Gate
5. Compound Gates
6. Pass Transistors and Transmission Gates
7. Tristates
8. Multiplexers
9. Sequential Circuits

CMOS logic

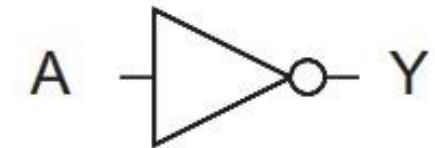
1. Inverter (NOT gate)

TABLE 1.1 Inverter truth table

A	Y
0	1
1	0



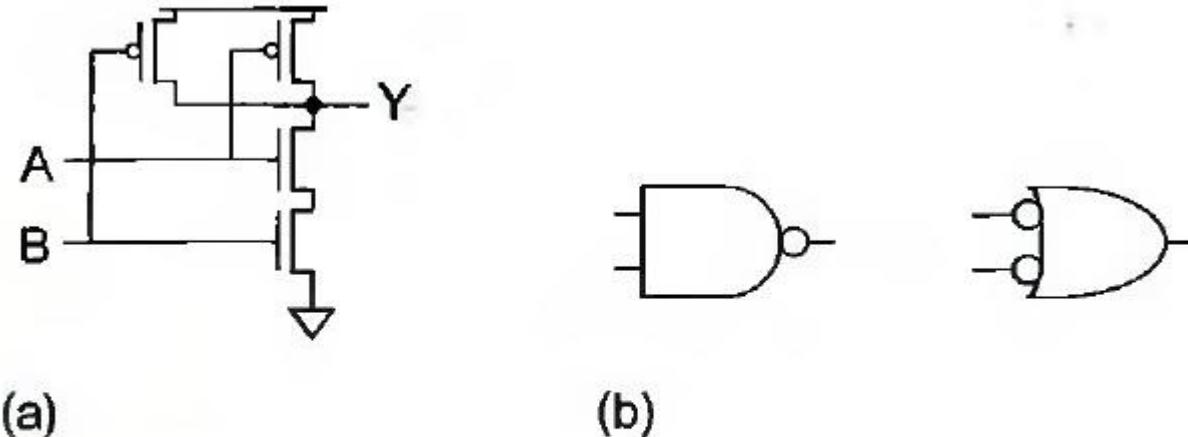
(a)



(b)

FIGURE 1.11
Inverter schematic
(a) and symbol
(b) $Y = \bar{A}$

2. NAND gate



	Gate=0	Gate=1
nMOS	OFF	ON
PMOS	ON	OFF

FIG 1.11 2-input NAND gate schematic (a) and symbol (b) $Y = \bar{A} \cdot \bar{B}$

Operation	AND	OR
nMOS	Series	Parallel
PMOS	Parallel	Series

Table 1.2 NAND gate truth table

A	B	pull-down network	pull-up network	Y
0	0	OFF	ON	1
0	1	OFF	ON	1
1	0	OFF	ON	1
1	1	ON	OFF	0

3. CMOS logic gates

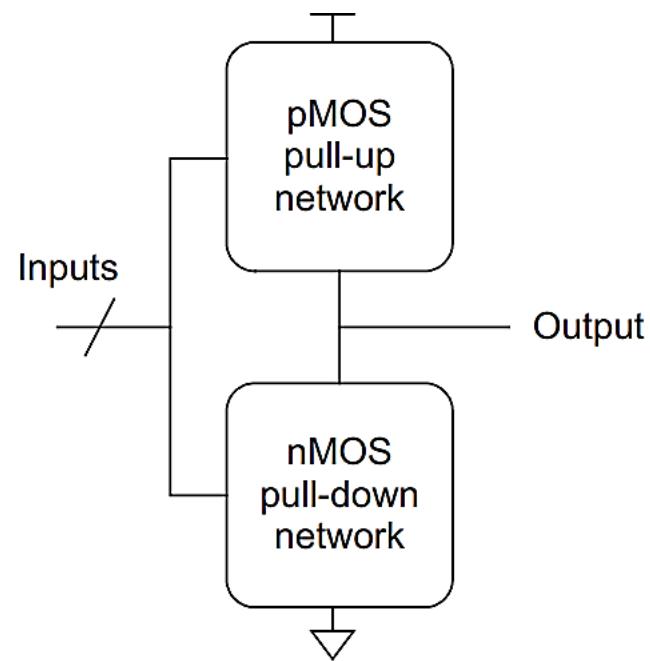


FIGURE 1.14 General logic gate using pull-up and pull-down networks

TABLE 1.3 Output states of CMOS logic gates

	pull-up OFF	pull-up ON
pull-down OFF	Z	1
pull-down ON	0	crowbarred (X)

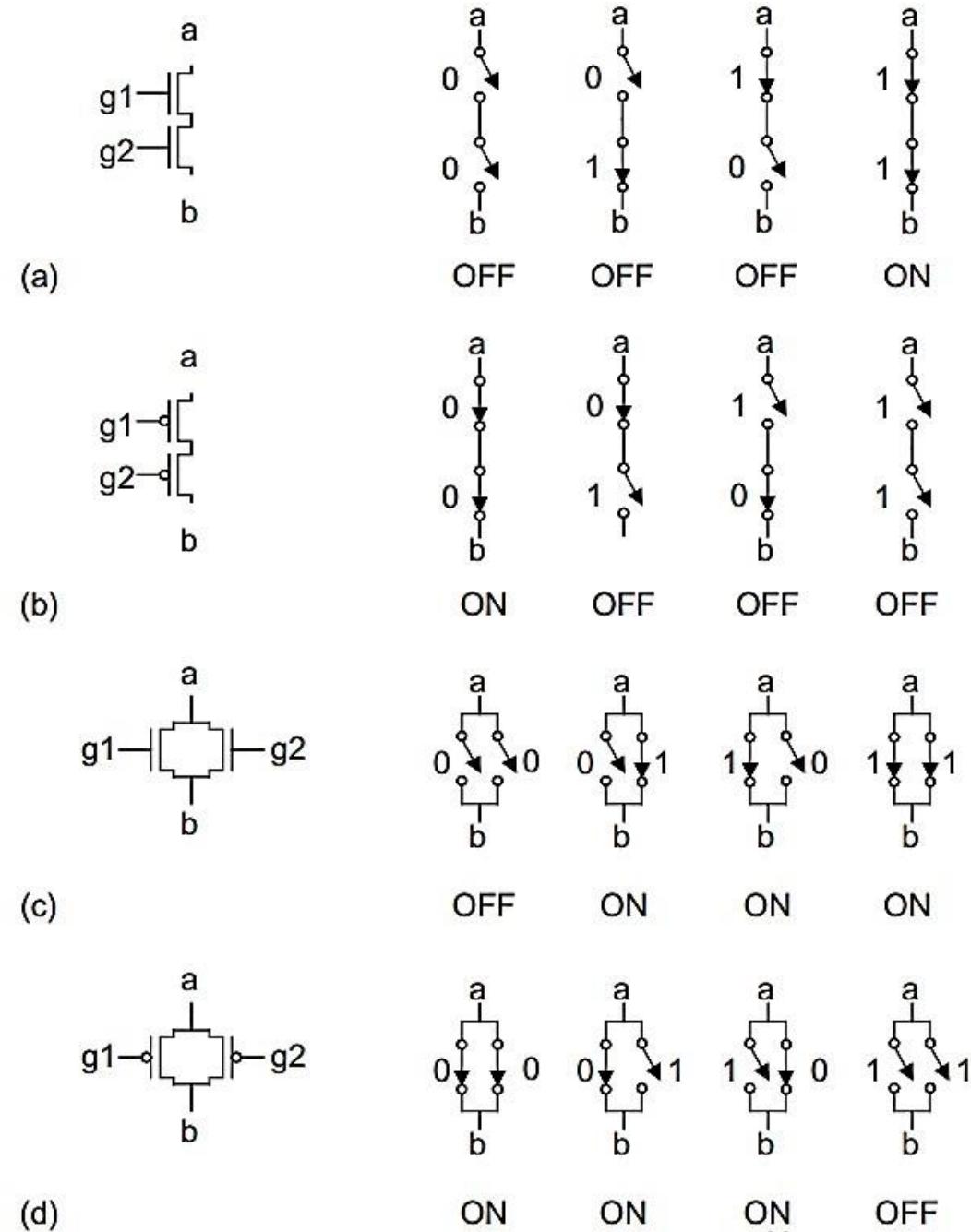
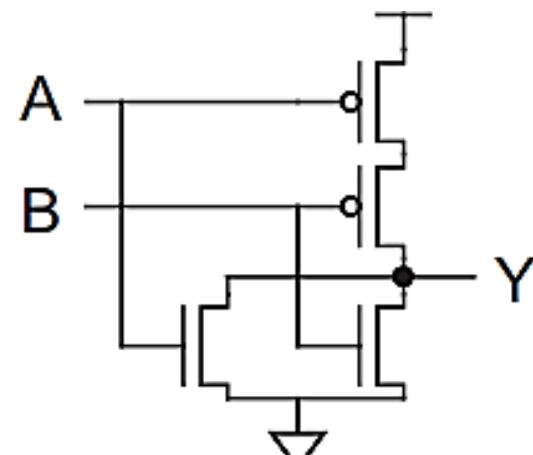
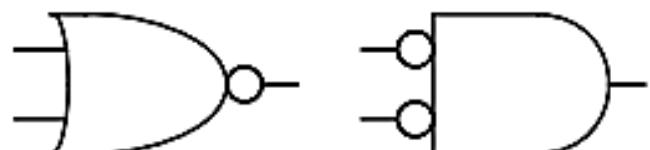


FIGURE 1.15 Connection and behavior of series and parallel transistors

4. NOR gate



(a)



(b)

FIGURE 1.16 2-input NOR gate schematic (a) and symbol
(b) $Y = \overline{A + B}$

Operation	AND	OR
nMOS	Series	Parallel
PMOS	Parallel	Series

	Gate=0	Gate=1
nMOS	OFF	ON
PMOS	ON	OFF

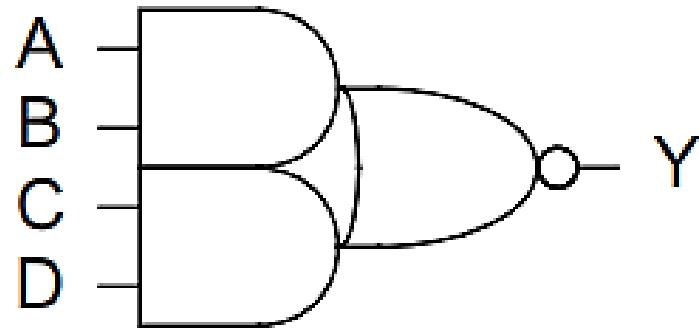
TABLE 1.4 NOR gate truth table

A	B	Y
0	0	1
0	1	0
1	0	0
1	1	0

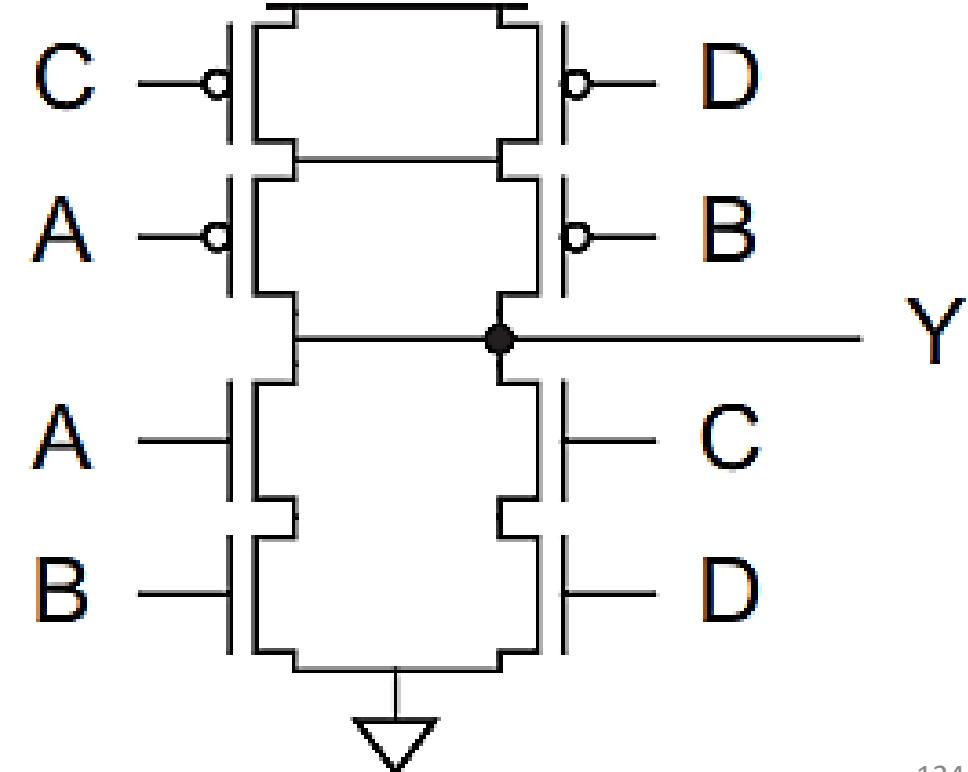
5. Compound gates

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

This function is sometimes called AND-OR-INVERT-22, or AOI22



Operation	AND	OR
nMOS	Series	Parallel
PMOS	Parallel	Series

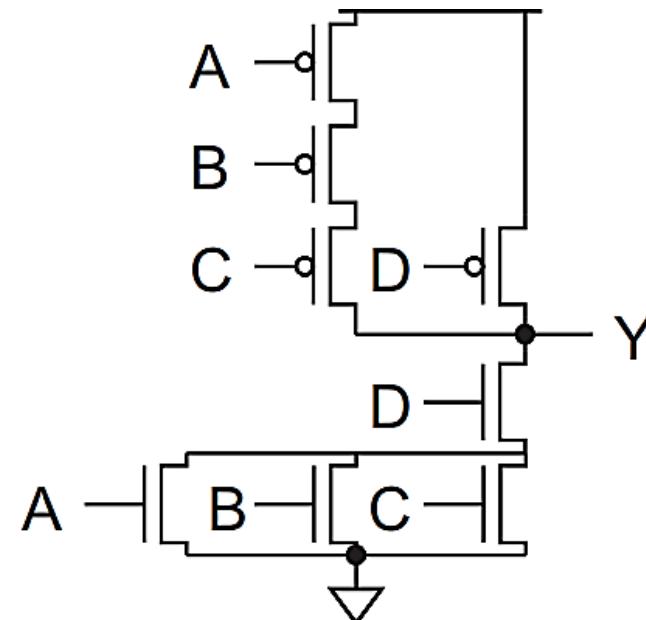


AOI (AND OR INVERT)

- $\overline{AB} + C$ is known as a 2-1 AOI gate.
- $\overline{AB} + \overline{CD}$ is known as a 2-2 AOI gate.
- $\overline{ABC} + \overline{DEF}$ is known as a 3-3 AOI gate.
- $\overline{ABCD} + \overline{EFGH}$ is known as a 4-4 AOI gate.
- $\overline{ABCDE} + \overline{FGH} + \overline{JK}$ is known as a 4-3-2 AOI gate.
- and other variations.

Sketch a static CMOS gate computing $Y = \overline{(A + B + C) \cdot D}$.

Figure 1.19 shows such an OR-AND-INVERT-3-1 (OAI31) gate.



Operation	AND	OR
nMOS	Series	Parallel
pMOS	Parallel	Series

FIGURE 1.19
CMOS compound gate
for function
 $Y = \overline{(A + B + C) \cdot D}$

6. Pass transistors and Transmission gates

- > V_{dd} is the rich source of “strong 1”.
- Gnd is the rich source of “strong 0”.
- An nMOS transistor is an almost perfect switch when passing a '0' and thus we say it passes a strong "0".
- However, the nMOS transistor is imperfect at passing a '1. It passes “weak 1”.
- A pMOS transistor is an almost perfect switch when passing a '1' and thus we say it passes a strong “1”.
- However, the pMOS transistor is imperfect at passing a '0. It passes “weak 0”.

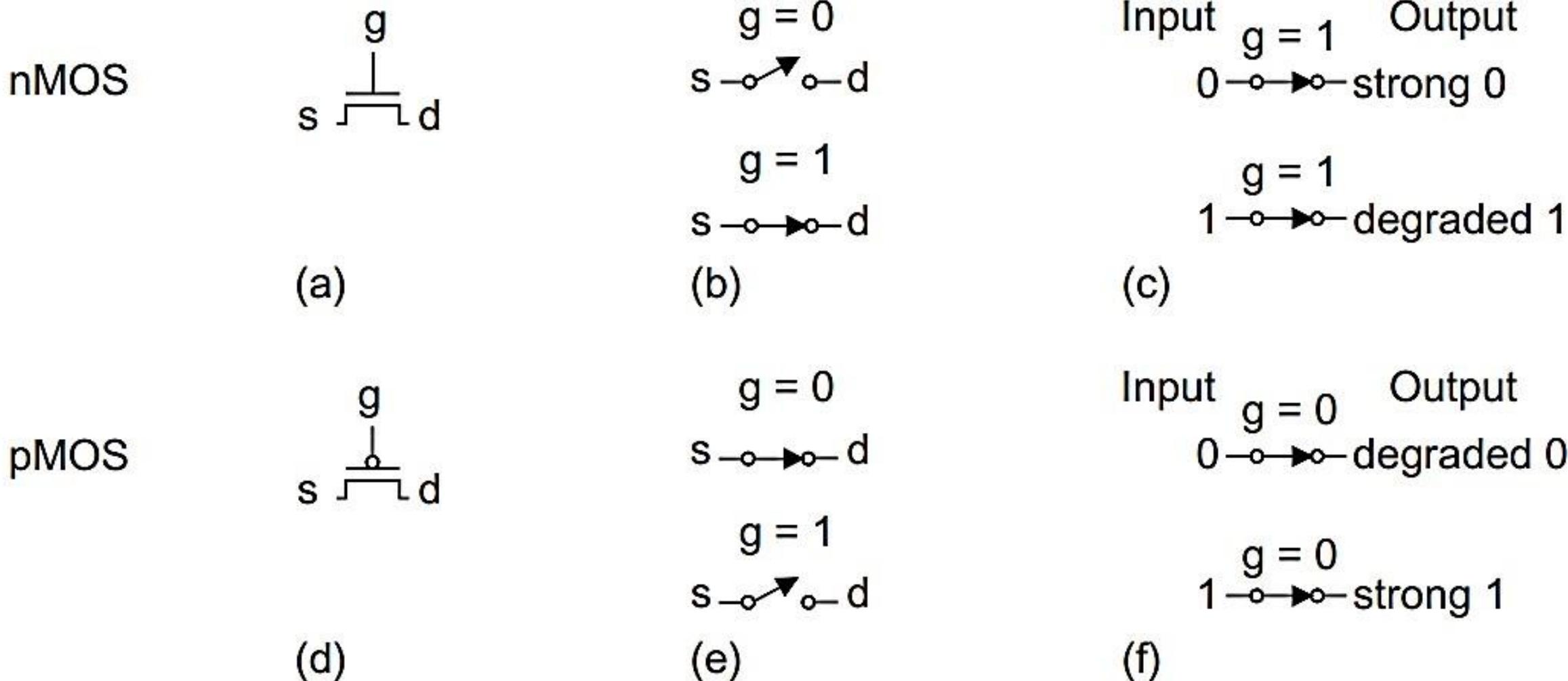
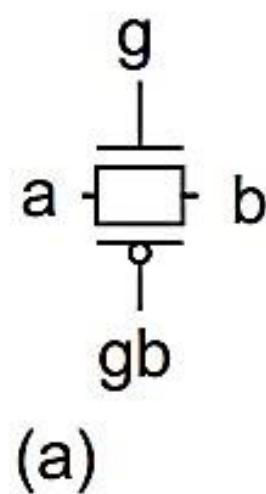


FIGURE 1.20 Pass transistor strong and degraded outputs



$g = 0, gb = 1$
 $a \xrightarrow{\circ} \circ b$

$g = 1, gb = 0$
 $a \xrightarrow{\bullet} \circ b$

(b)

Input	Output
$g = 1, gb = 0$	$g = 1, gb = 0$
$0 \xrightarrow{\bullet} \circ$	strong 0
$1 \xrightarrow{\bullet} \circ$	strong 1

(c)

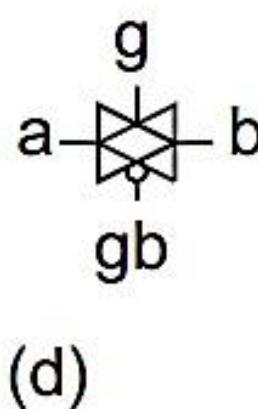


FIGURE 1.21 Transmission gate

7. Tristates

a) Tristate buffer

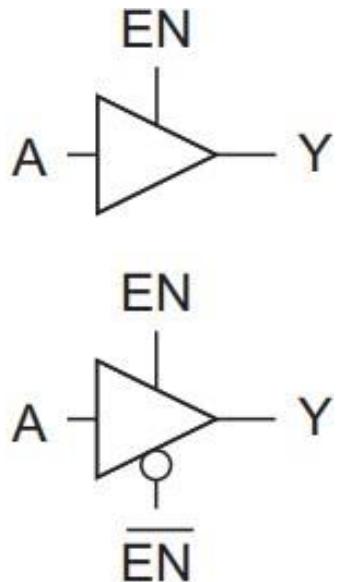


FIGURE 1.25
Tristate buffer
symbol

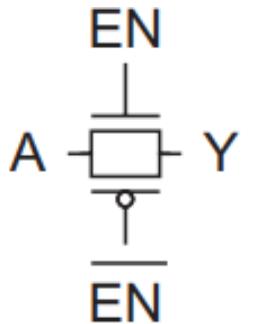
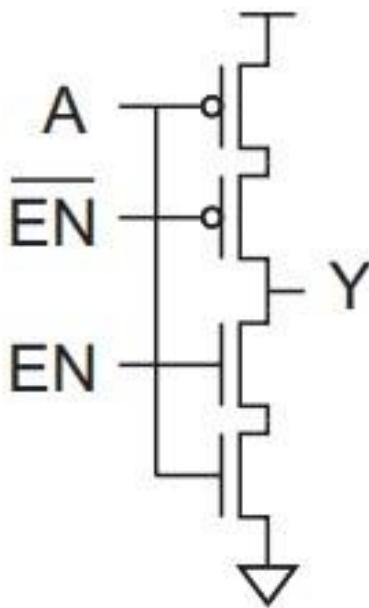


FIGURE 1.26
Transmission gate

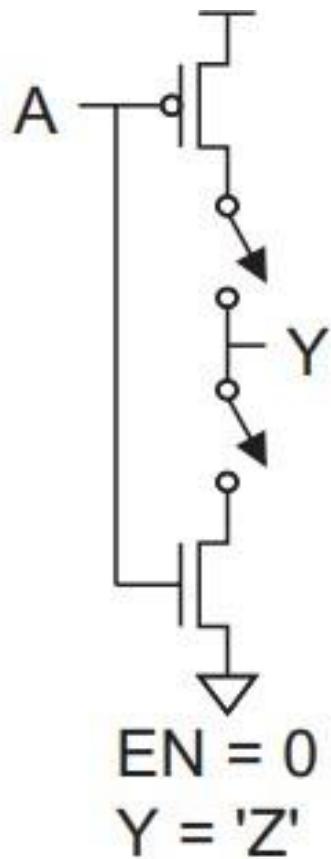
TABLE 1.5 Truth table for tristate

EN / \bar{EN}	A	Y
0 / 1	0	Z
0 / 1	1	Z
1 / 0	0	0
1 / 0	1	1

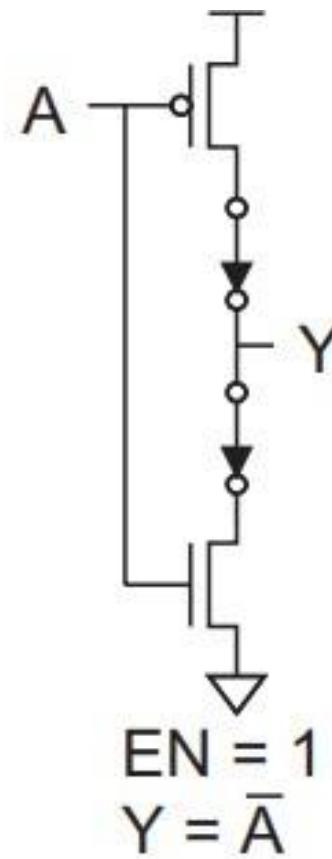
b) Tristate inverter



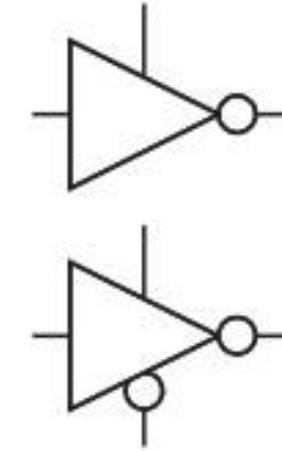
(a)



(b)



(c)

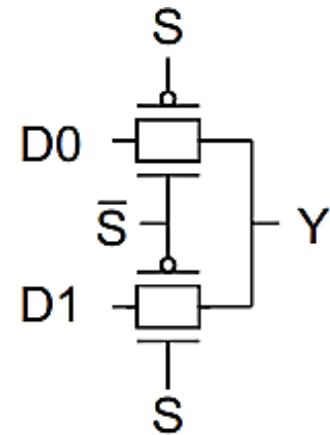


(d)

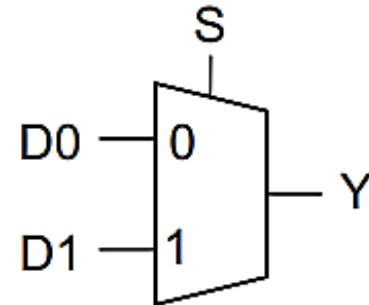
FIGURE 1.27 Tristate Inverter

8. Multiplexers

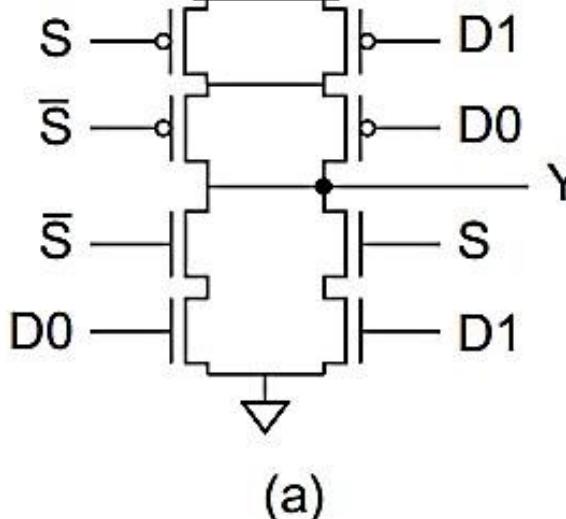
$$Y = \bar{S} \cdot D_0 + S \cdot D_1.$$



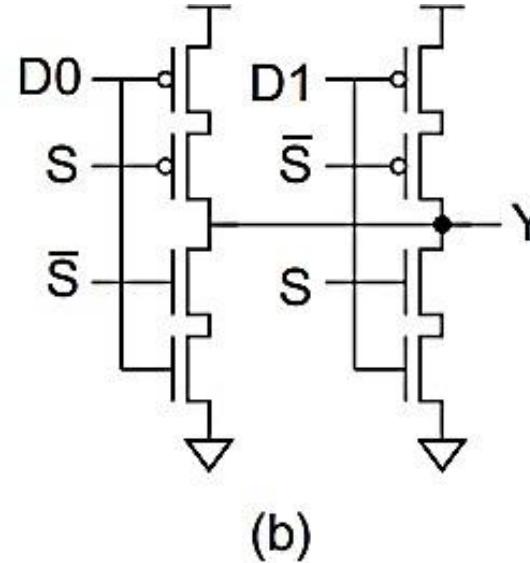
(a)



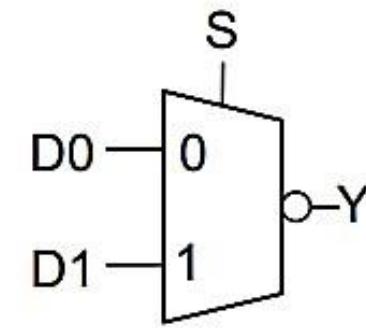
(b)



(a)



(b)



(c)

TABLE 1.6 Multiplexer truth table

S / \bar{S}	D_1	D_0	Y
0 / 1	X	0	0
0 / 1	X	1	1
1 / 0	0	X	0
1 / 0	1	X	1

FIGURE 1.28 Transmission gate multiplexer

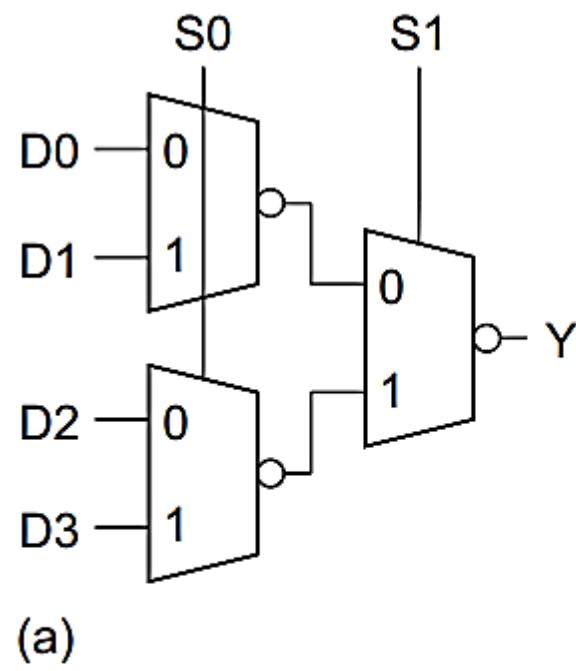
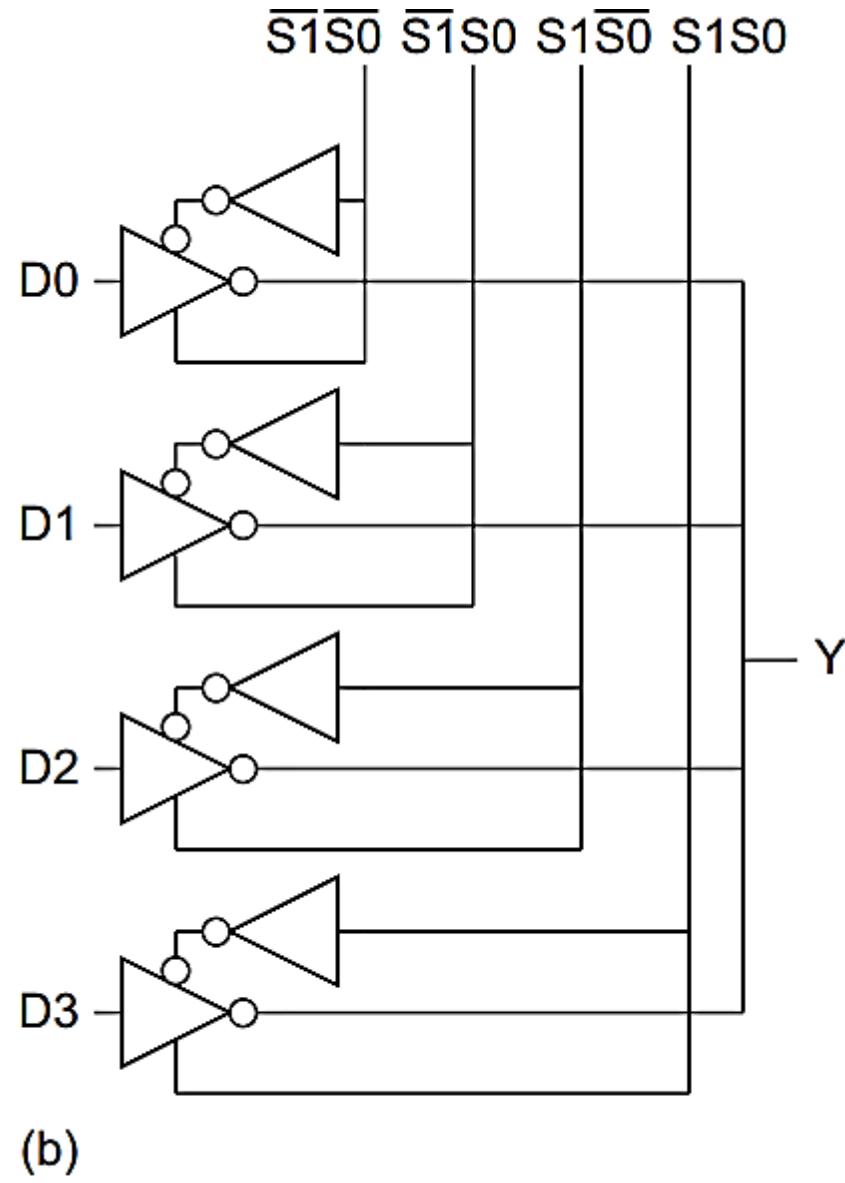


FIGURE 1.30 4:1 multiplexer



9. Latches and flipflops

When $\text{CLK}=1$, latch is transparent

When $\text{CLK}=0$, latch is opaque

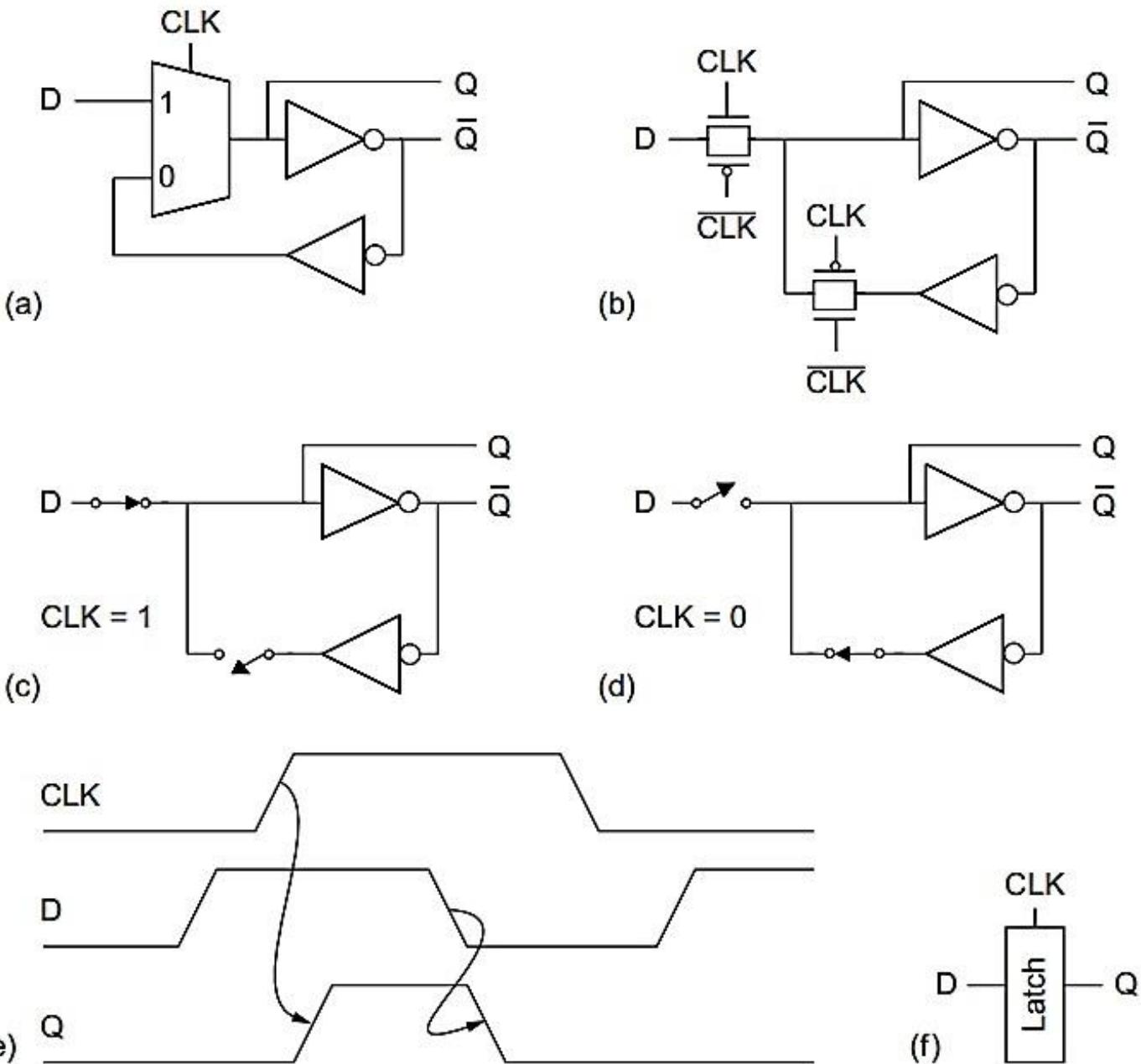
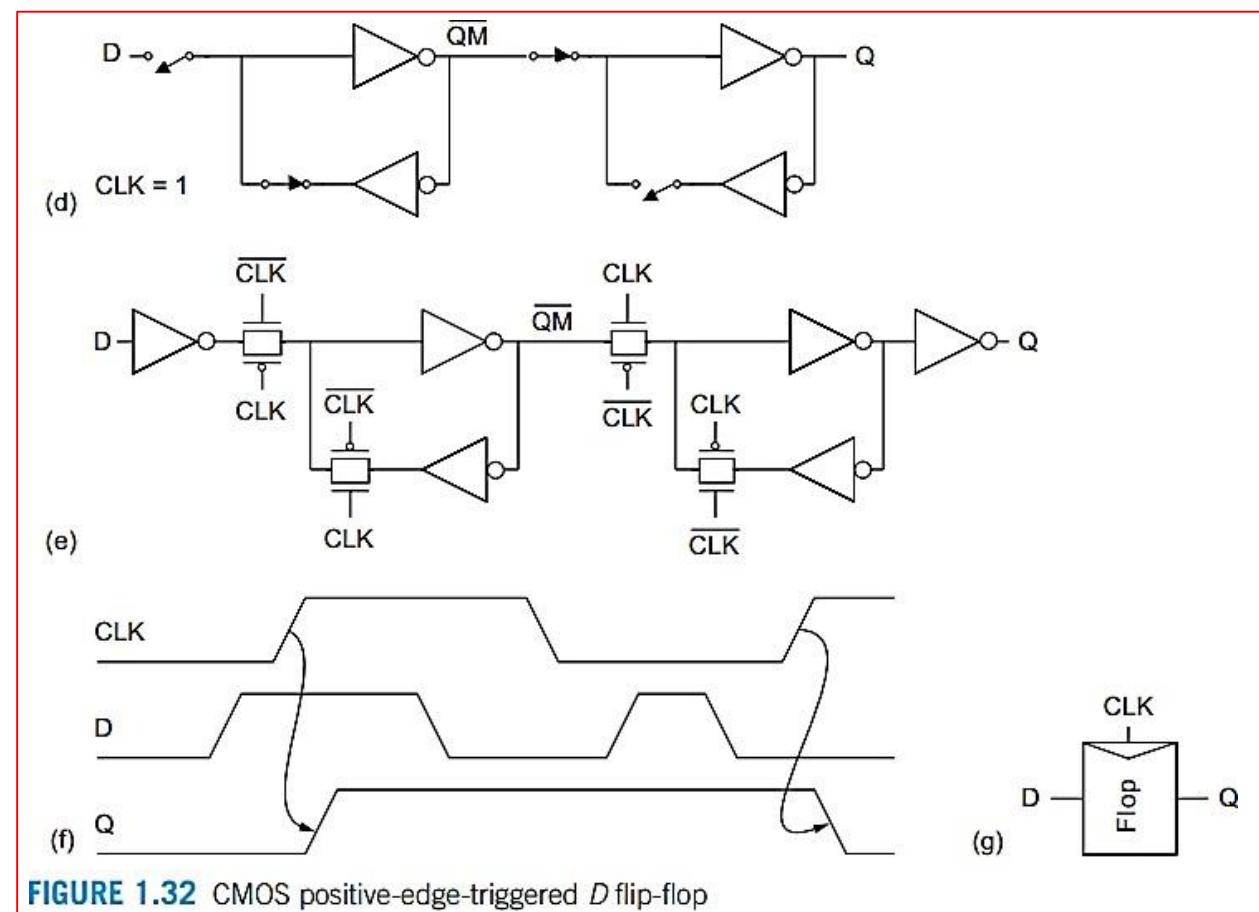
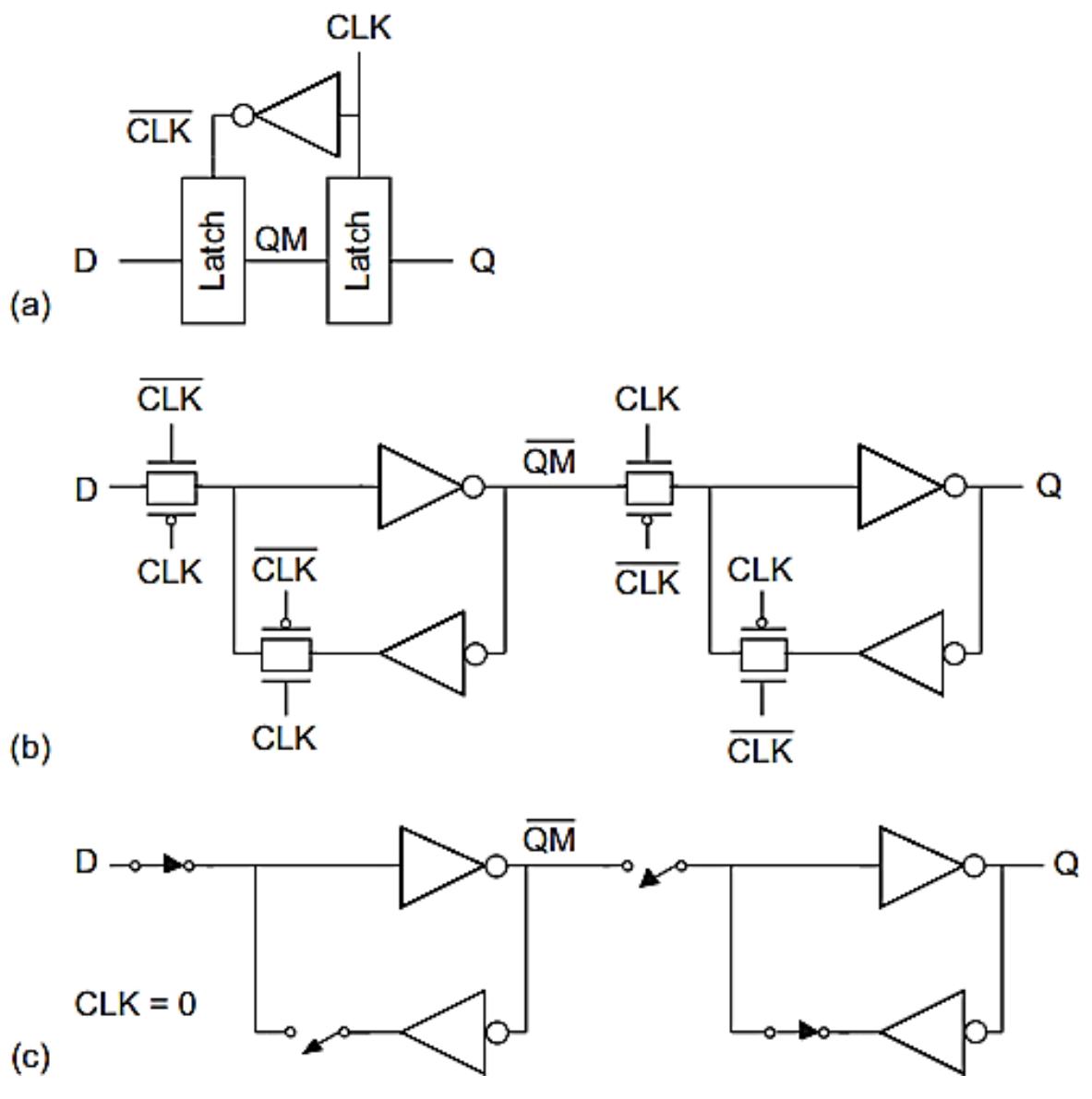
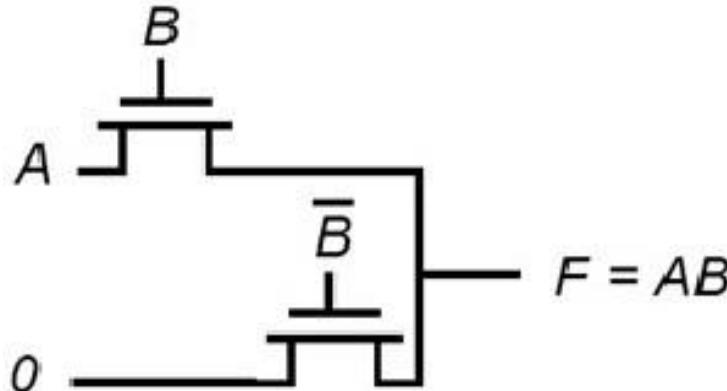


FIGURE 1.31 CMOS positive-level-sensitive D latch



Implementing Boolean expression using Pass transistors and TGs



AND Gate

Truth Table

A	B	$Y = A \cdot B$
0	0	0
0	1	0
1	0	0
1	1	1

Symbol

$Y = A \cdot B$

Circuit

$A \cdot B + 0 \cdot \bar{B}$

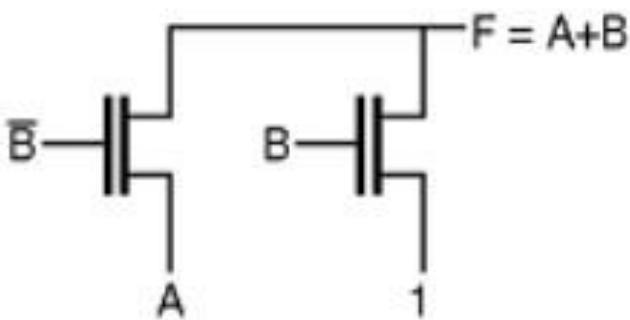
$\bar{A} \cdot B$

136

OR Gate

Truth Table

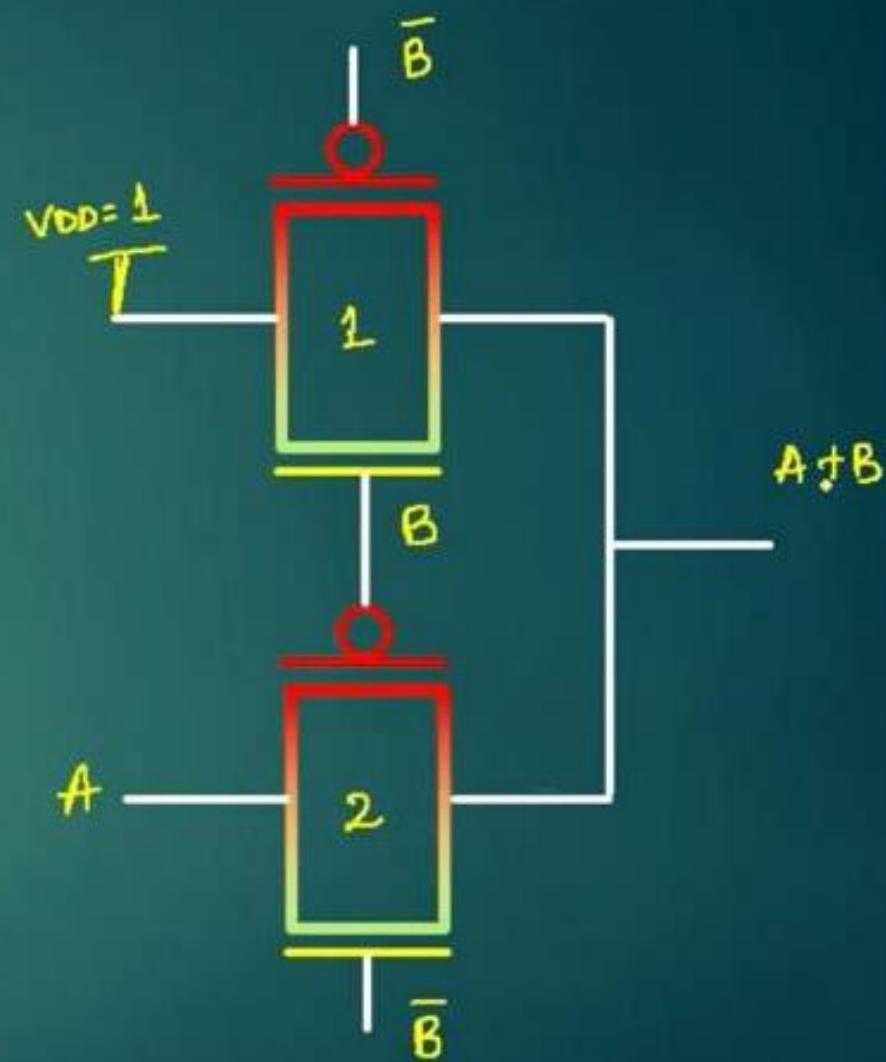
A	B	$Y = A + B$
0 ✓	0 ✓	0 ✓
0	1 ✓	1 ✓
1 ✓	0 ✓	1 ✓
1	1 ✓	1 ✓

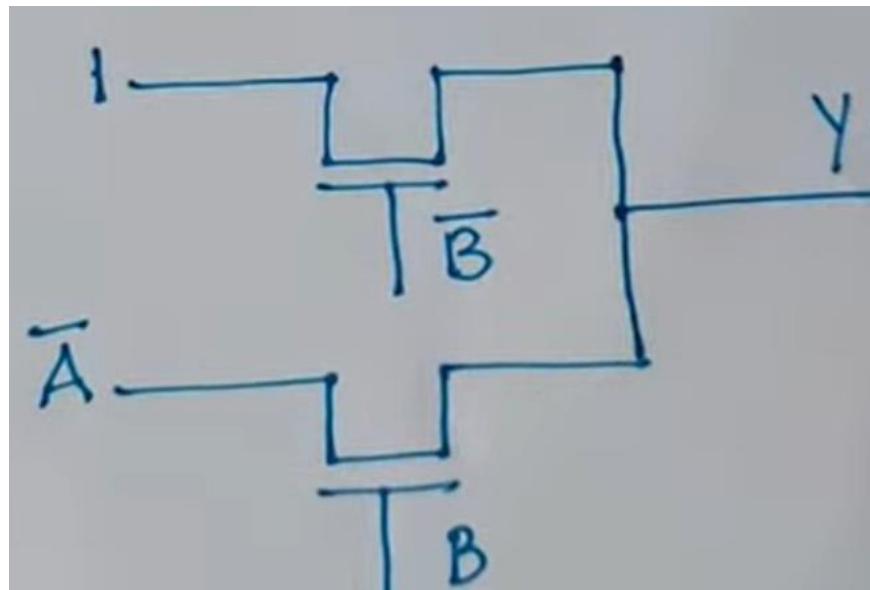


Symbol



Circuit





NAND Gate

Truth Table		
A	B	$Y = \overline{A \cdot B}$
0	0	1
0	1	1
1	0	1
1	1	0

Symbol

The standard logic symbol for a NAND gate is a rectangle with an inverted triangle at the top. Two input lines enter from the left, labeled A and B. One output line exits from the right, labeled $Y = \overline{A \cdot B}$.

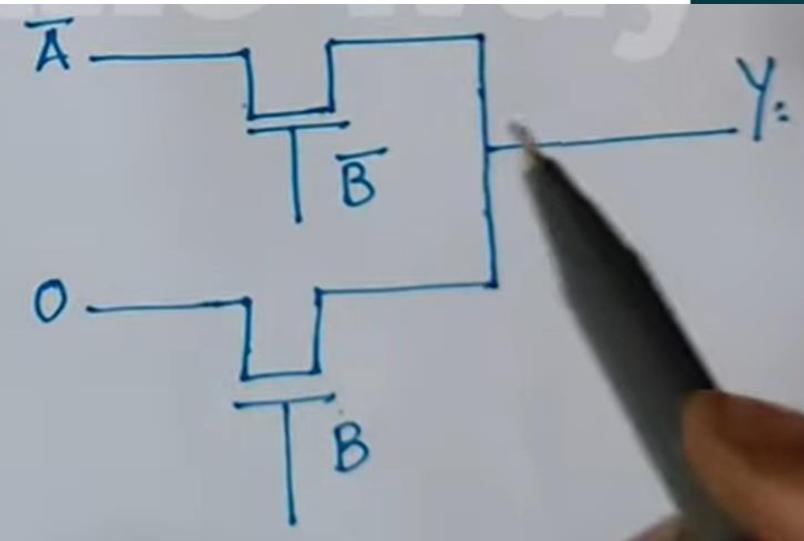
Circuit

A circuit diagram showing how a NAND gate can be implemented using NOR gate primitives. It consists of two NOR gates (represented by rectangles with triangles pointing down) connected in series. The first NOR gate has inputs \bar{A} and B, and its output is labeled \bar{B} . This output is connected to the second NOR gate, which also has input \bar{A} . The other input of the second NOR gate is connected to ground ($V_{DD} = 0$). The output of the second NOR gate is labeled $A \cdot B$.

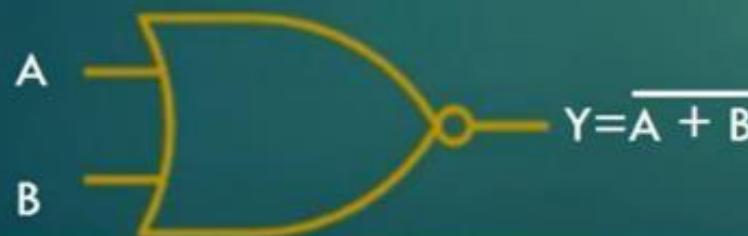
NOR Gate

Truth Table

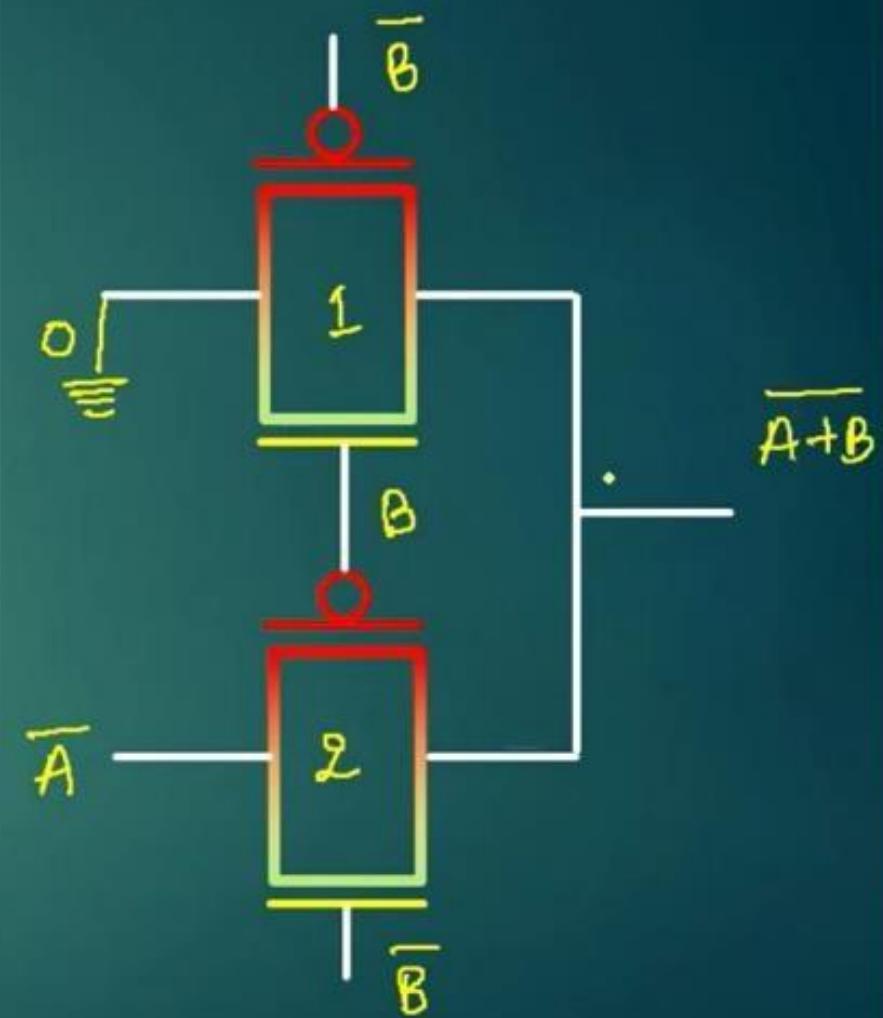
A	B	$Y = \overline{A+B}$
0	0	1
0	1	0
1	0	0
1	1	0



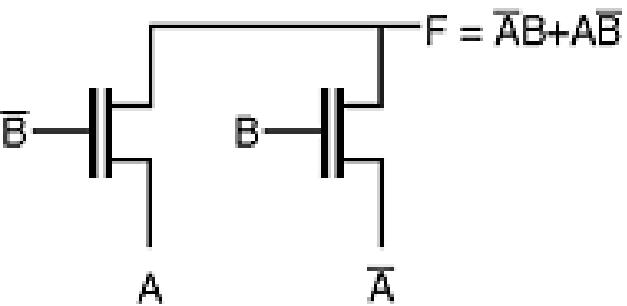
Symbol



Circuit



XOR Gate



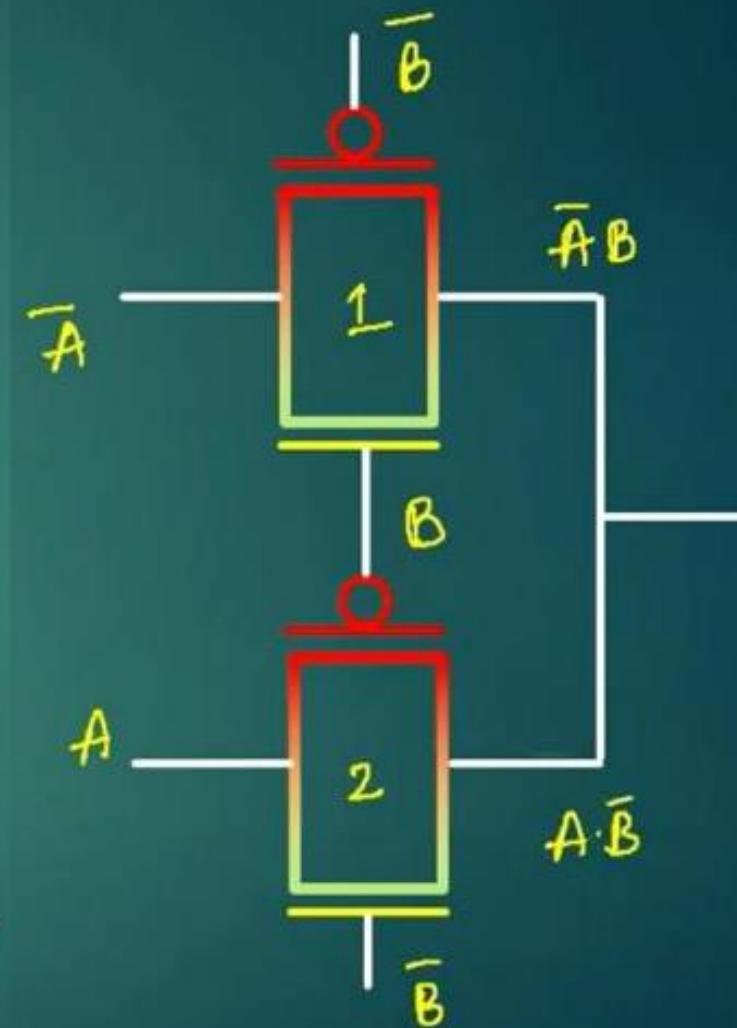
Truth Table

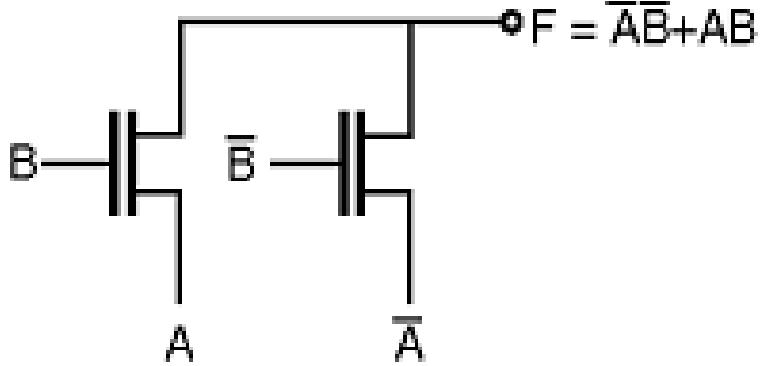
A	B	$Y = A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

Symbol



Circuit





XNOR Gate

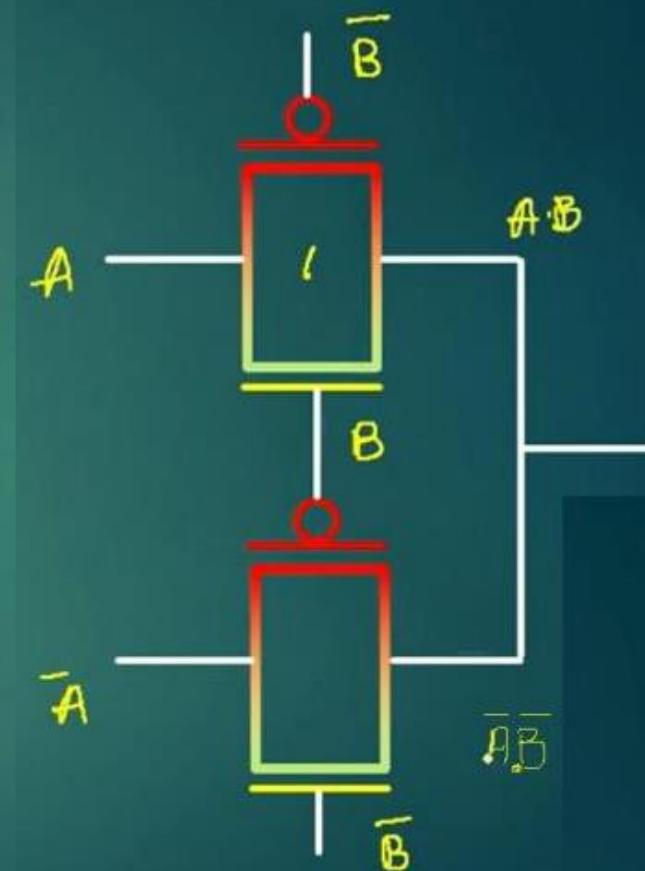
Truth Table

A	B	$Y = \overline{A} \oplus \overline{B}$
0	0	1
0	1	0
1	0	0
1	1	1

Symbol



Circuit



BiCMOS Inverter

- When $V_{in} = 0$, T_3 is off so that T_1 will be non-conducting. But T_4 is ON and supplies current to the base of T_2 which will conduct and act as a current source to charge C_L towards 5V. Output of Inverter will be 5V minus V_{BE} of T2
- With $V_{in} = 5V$, T_4 is OFF so that T_2 will be non conducting. But T_3 will be ON and will supply current to base of T_1 which will conduct and act as a current sink to C_L discharging it to 0. Output of inverter will fall to 0 volts plus V_{CESsat} of T1.
- Owing to the presence of direct path from V_{DD} to Gnd through T_3 and T_1 , this is not good arrangement to implement since there is significant **static current** flow whenever V_{in} is logic 1.

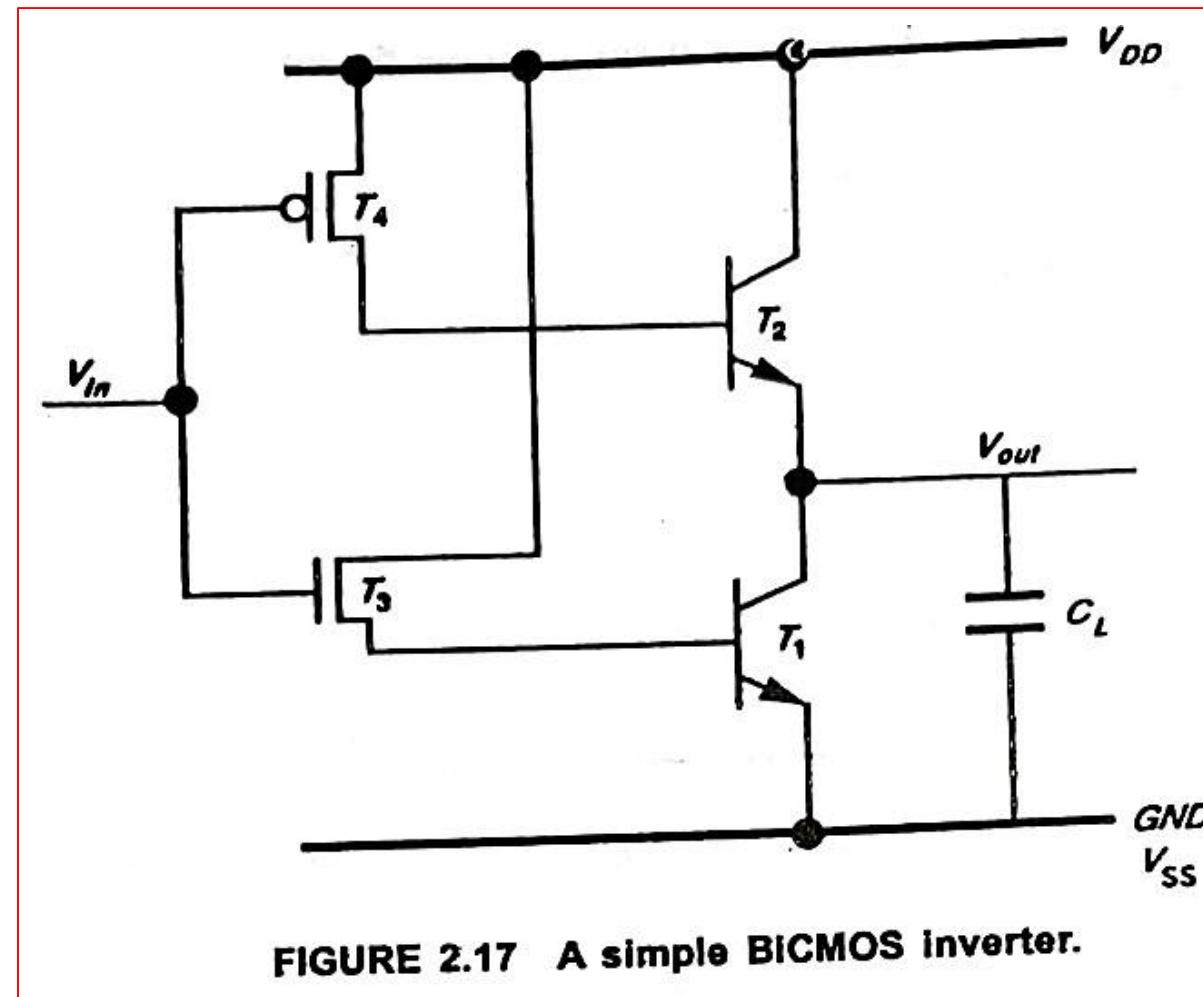


FIGURE 2.17 A simple BiCMOS inverter.

- An improved version of BiCMOS inverter is shown in Figure, in which DC path through T_3 and T_1 is eliminated.
- In both above circuits, there is no discharge path for the current from the base of BJT when it is being turned off.

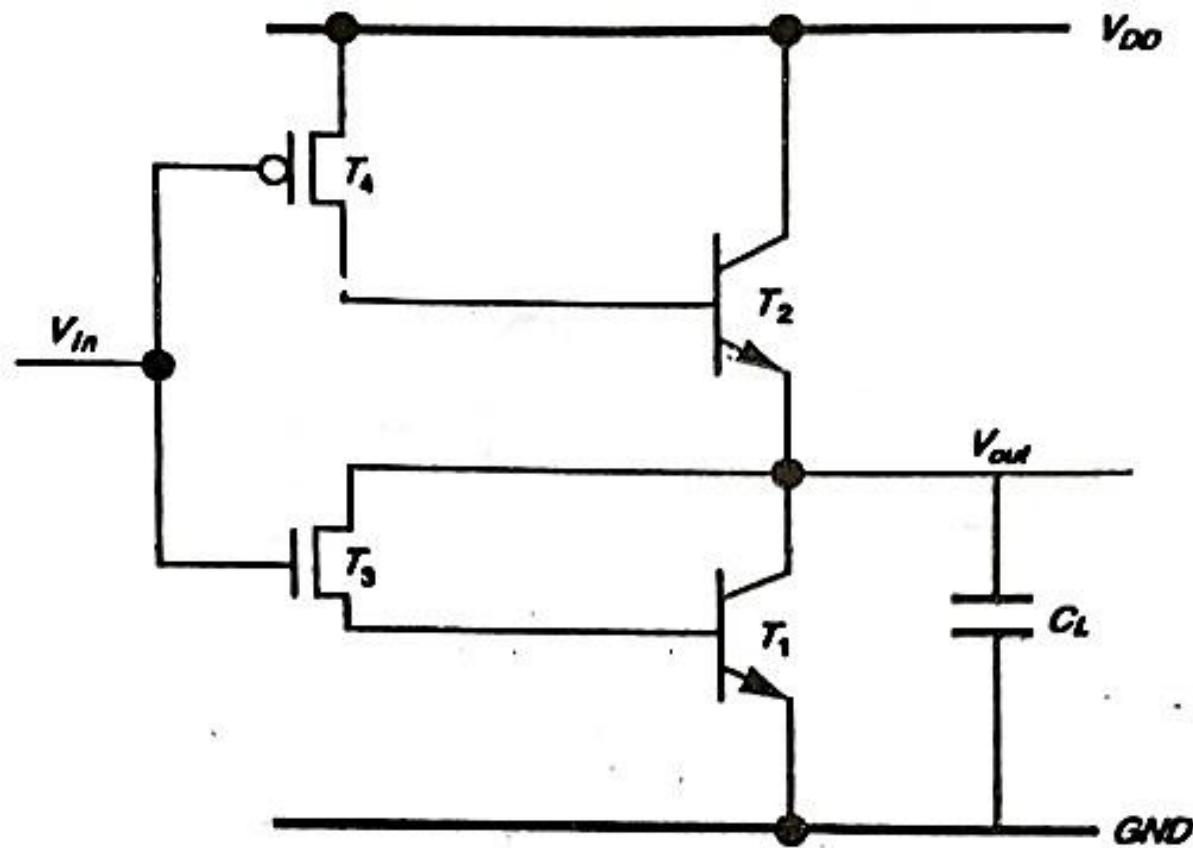


FIGURE 2.18 An alternative BiCMOS inverter with no static current flow.

- An improved inverter arrangement is shown in Figure.
- This provides improved output voltage swing when each BJT is OFF., also provide discharge path for base current during turn-off.

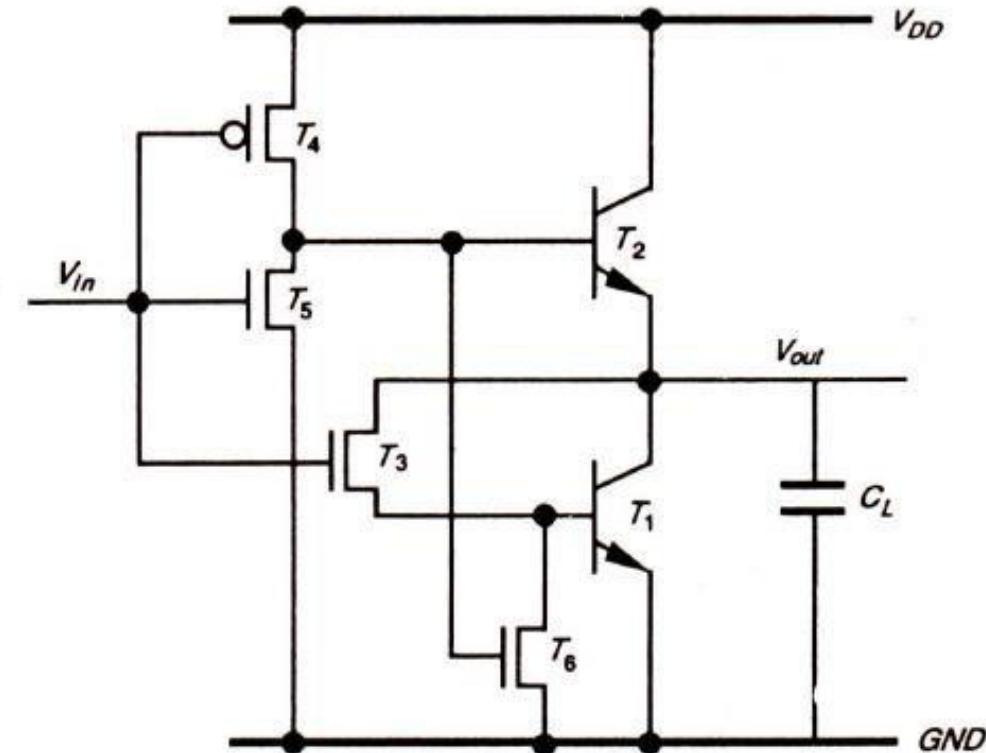
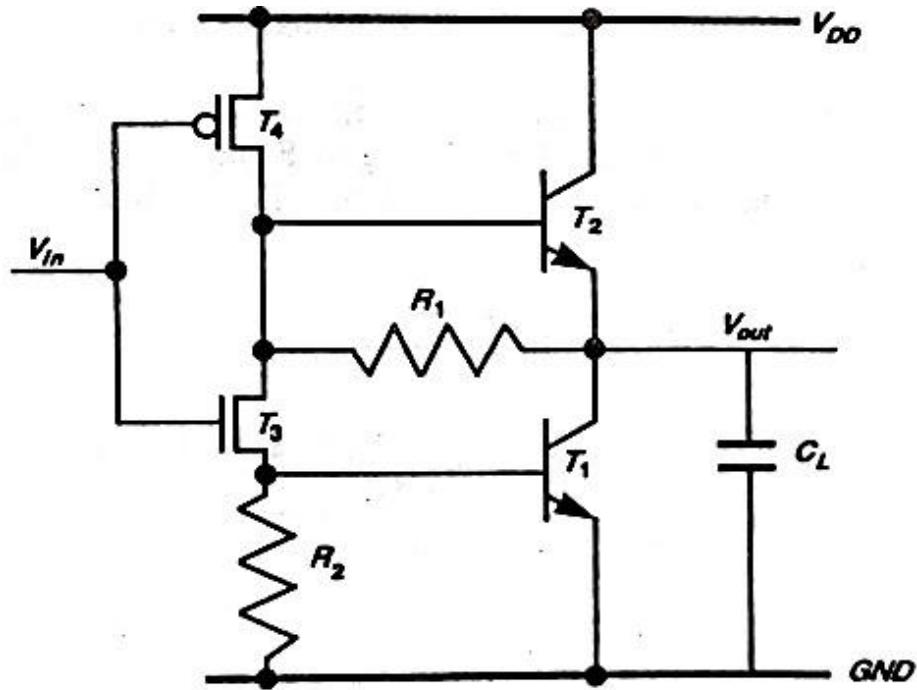
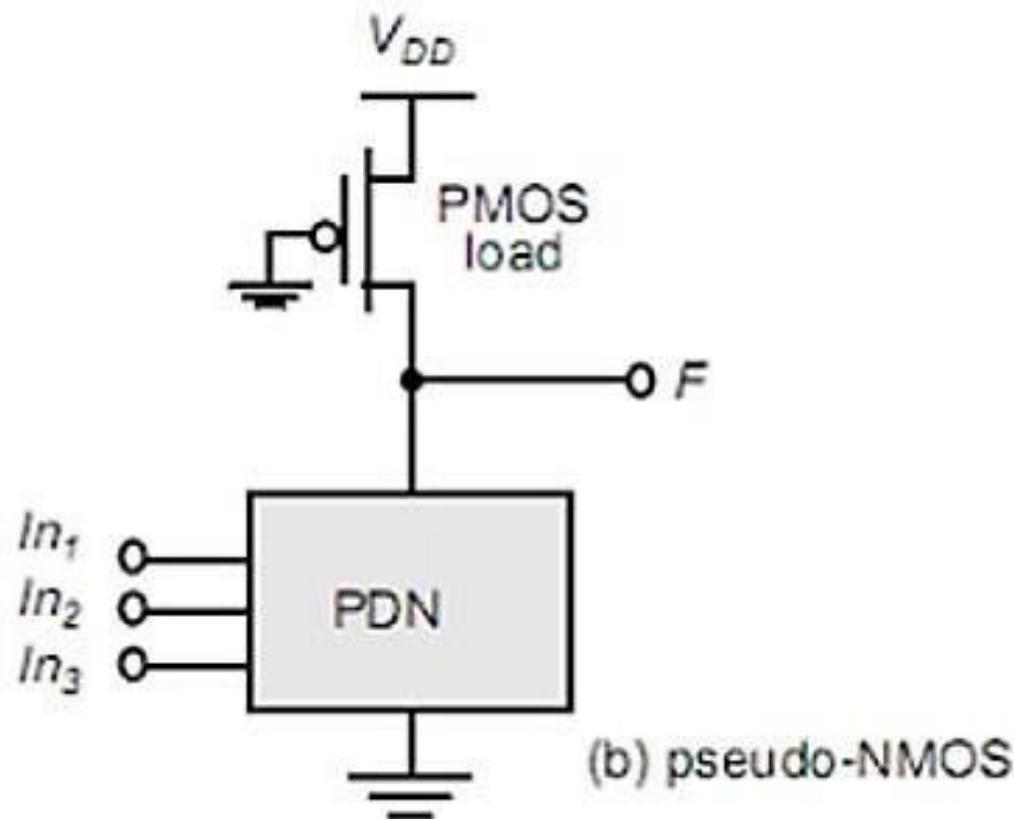


FIGURE 2.19 An Improved BICMOS Inverter with better output logic levels.

Implementation of Logic gates using BiCMOS

Pseudo-nMOS logic with derivation for Z_{pu}/Z_{pd} ratio

- Clearly, if we replace the depletion mode pull-up transistor of the standard nMOS circuits with a p-transistor with gate connected to V_{ss} (Gnd), we have a structure similar to the nMOS equivalent.



(b) pseudo-NMOS

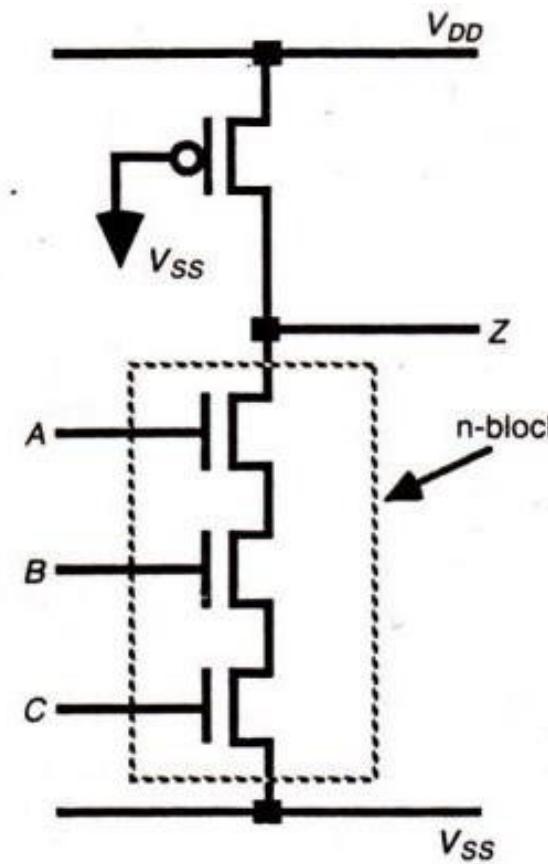


FIGURE 6.9 Pseudo-nMOS Nand gate.

- Consider a situation where on pseudo-nMOS inverter is being driven by another similar inverter.
- As for the nMOS analysis, we consider the conditions for which,

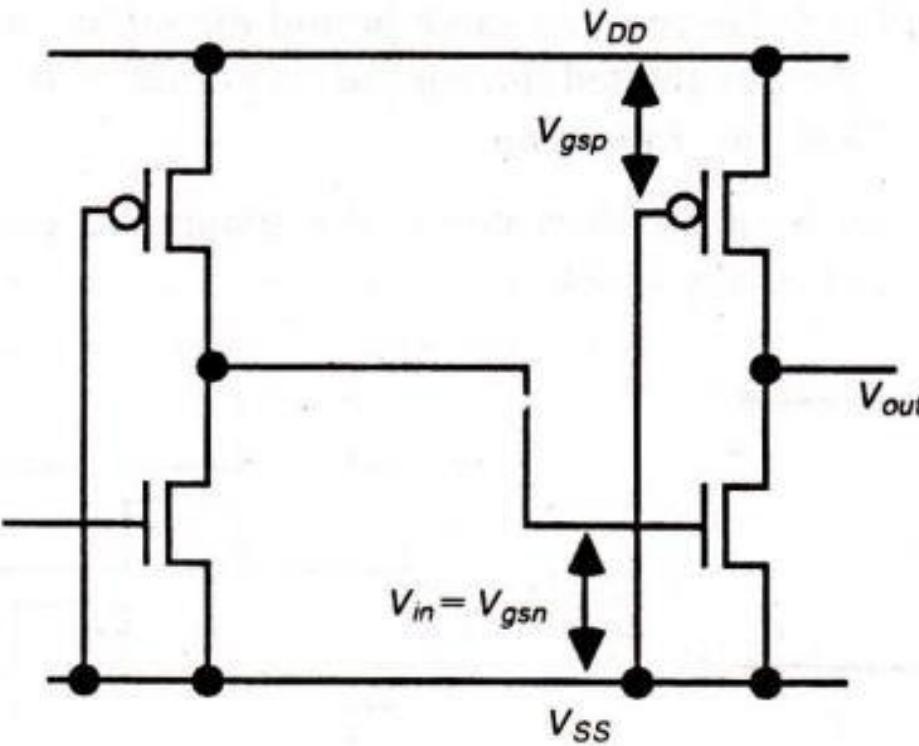
$$V_{inv} = V_{DD}/2$$

- At this point n-device is in saturation,

$$0 < V_{gsn} - V_{tn} < V_{dsn}$$

- p-device is in non-saturation,

$$0 < V_{dsp} < V_{gsp} - V_{tp}$$



Pseudo-nMOS inverter when driven from a similar inverter

- Equating currents of the n-transistor and the p-transistor, and by suitable rearrangement of the resultant expression, we obtain,

$$V_{inv} = V_{tn} + \frac{(2\mu_p/\mu_n)^{1/2} [(-V_{DD} - V_{tn})V_{dsp} - V_{dsp}^2]^{1/2}}{(Z_{p.u.}/Z_{p.d.})^{1/2}}$$

- Where, $Z_{p.u.} = L_p/W_p$ $Z_{p.d.} = L_n/W_n$

- With, $V_{inv} = 0.5V_{DD}$

$$V_{tn} = |V_{tp}| = 0.2V_{DD}$$

$$V_{DD} = 5 \text{ V}$$

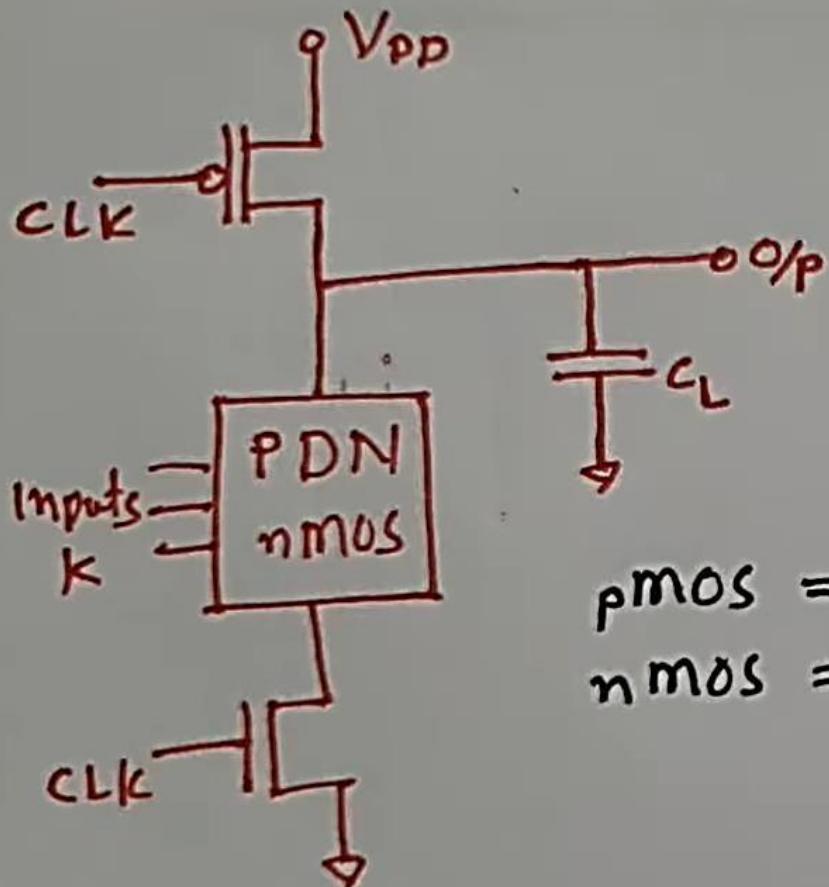
$$\mu_n = 2.5 \mu_p$$

- We obtain,

$$\frac{Z_{p.u.}}{Z_{p.d.}} = \frac{3}{1}$$

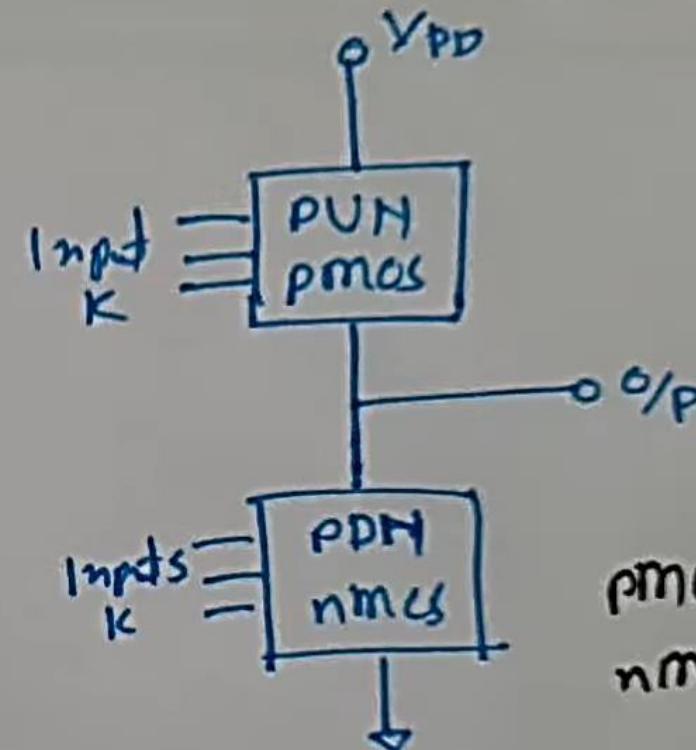
Dynamic and clocked CMOS inverters

Dynamic CMOS



$$\begin{aligned} \text{pmos} &= 1 \\ \text{nmos} &= K+1 \end{aligned}$$

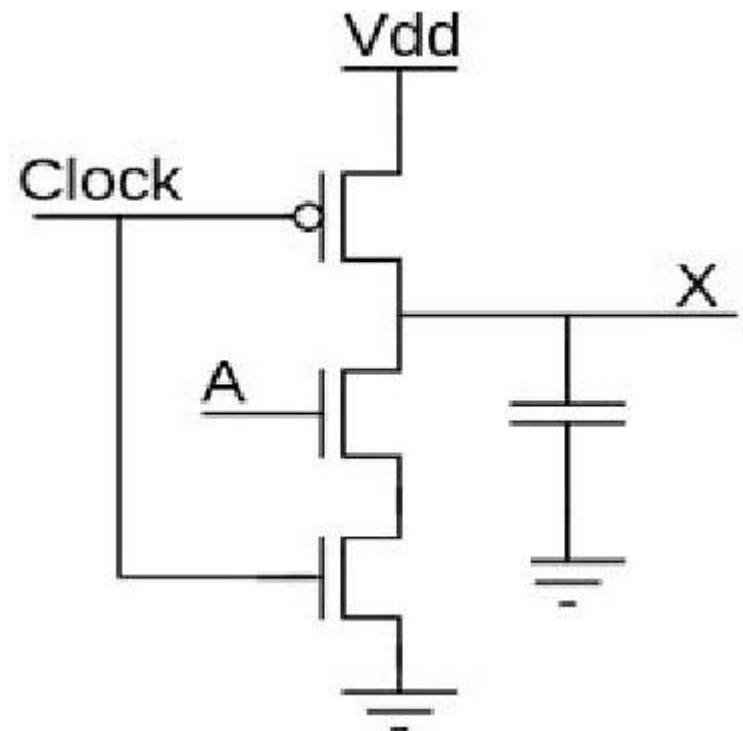
Static CMOS



$$\begin{aligned} \text{pmos} &= K \\ \text{nmos} &= K \end{aligned}$$

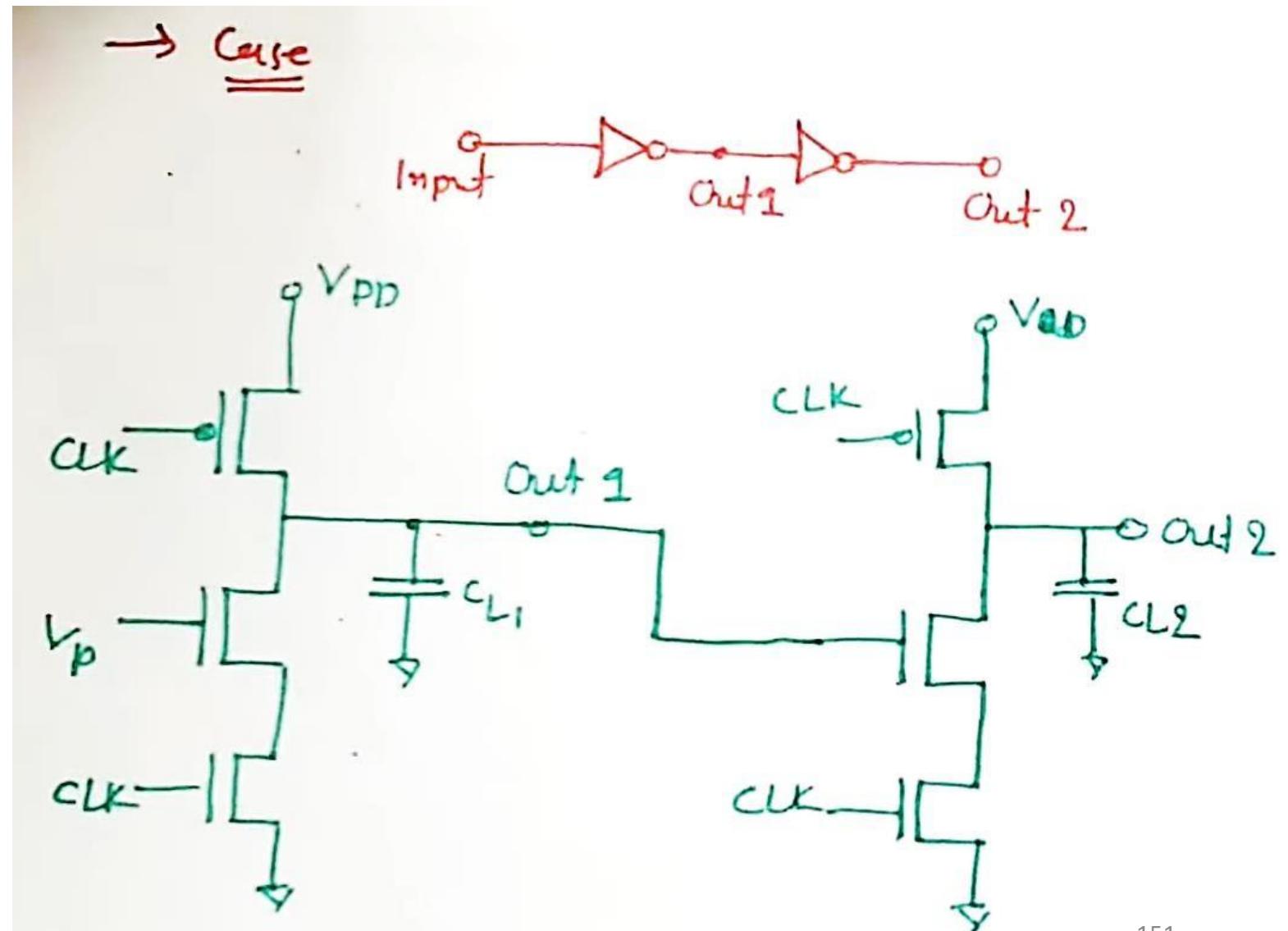
- Circuit works in 2 modes:
 1. Pre-charge mode (occurs when clock = 0)
 2. Evaluation mode (occurs when clock = 1)

Waveform:



Cascading problem in Dynamic CMOS

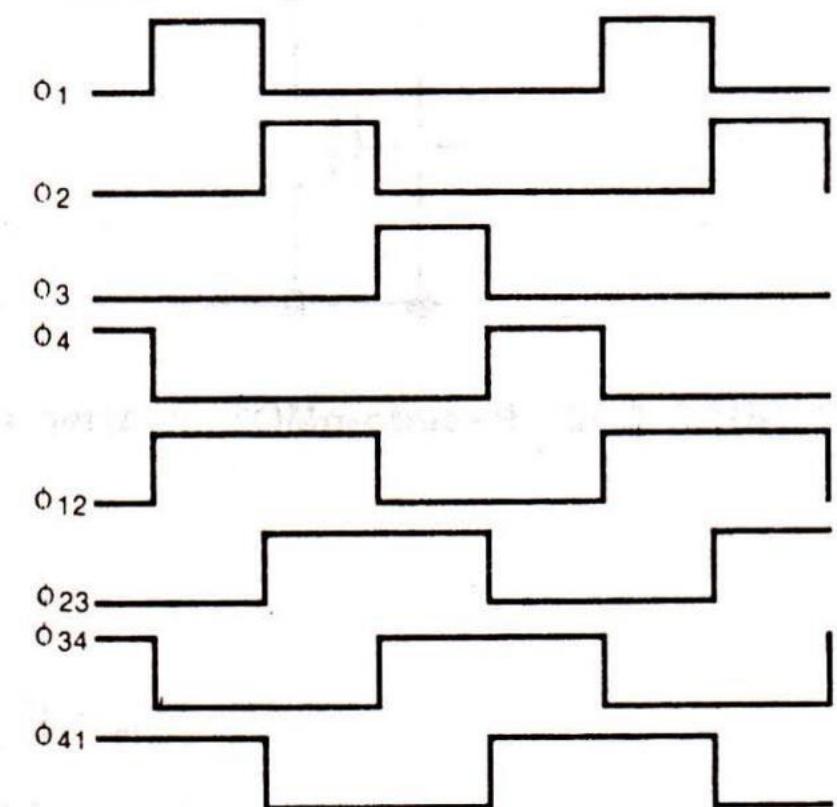
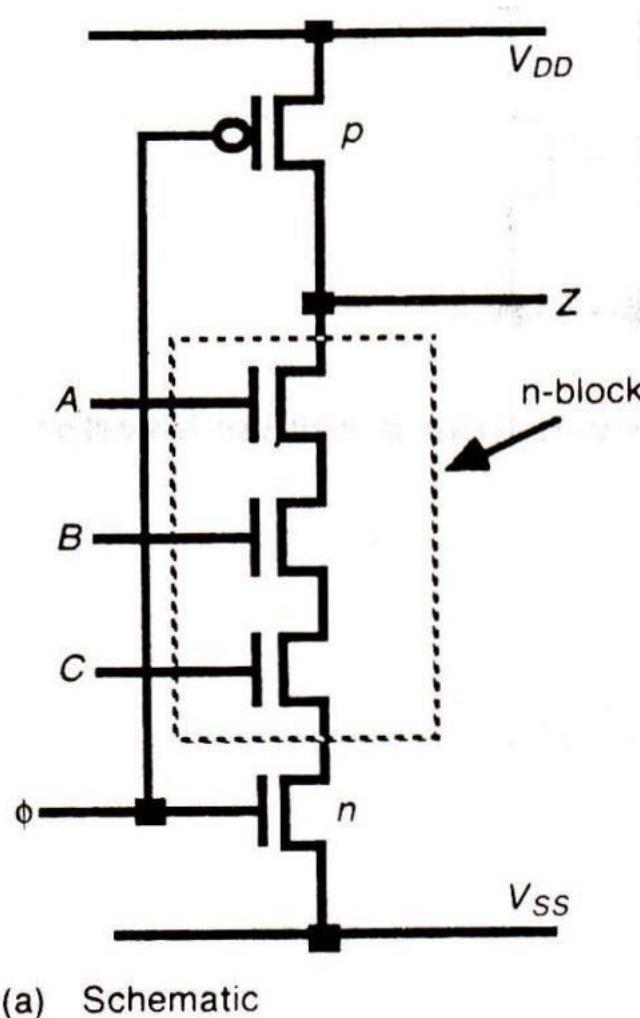
- Waveform:



Dynamic CMOS logic

Note the following:

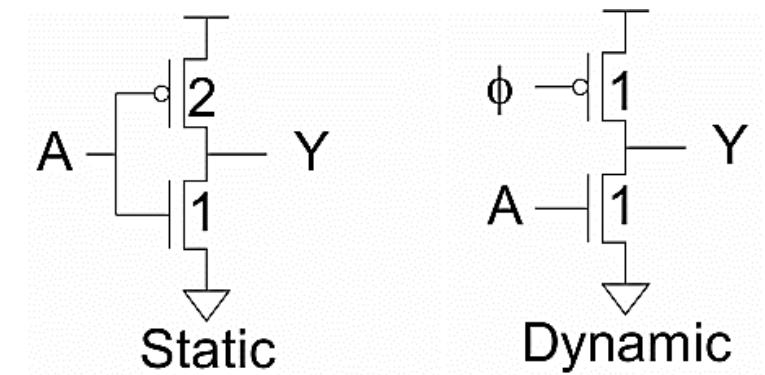
1. Charge sharing may be a problem unless the inputs are constrained' not to change during the on period of the clock.
2. Single phase dynamic logic structures cannot be cascaded since, owing to circuit delays, an incorrect input to the next stage may be present when evaluation begins, so that its output is inadvertently discharged and the wrong output results.



(b) Possible 4o and derived clocks

Advantages:

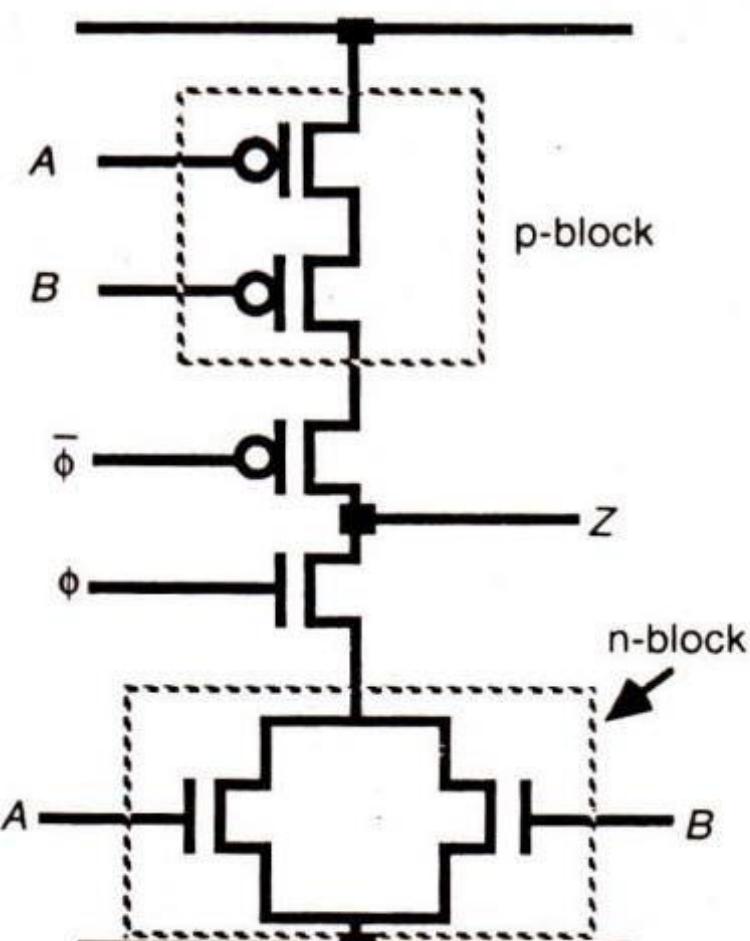
- Max output voltage swing (V_{DD} to Gnd).
- It has 1 pMOS. Fabrication is easier.
- Circuit size is small compared to CMOS circuits.
- Faster switching speed.
- Less capacitance loading.



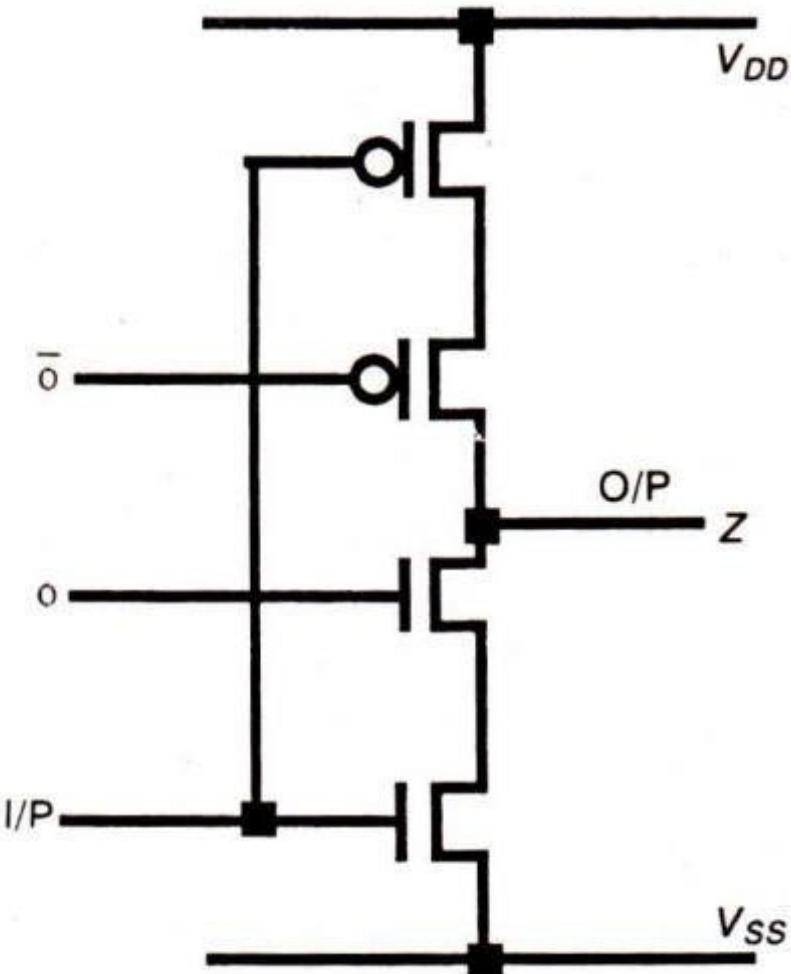
Disadvantages:

- Circuits cannot work without pre-charging.
- When output is logic 1, it stays in high impedance which is prone to noise.

Clocked CMOS (C^2MOS) logic



(a) 2 1/P Nor gate



(b) Inverter

FIGURE 6.12 Clocked CMOS (C^2MOS) logic.

CMOS domino logic

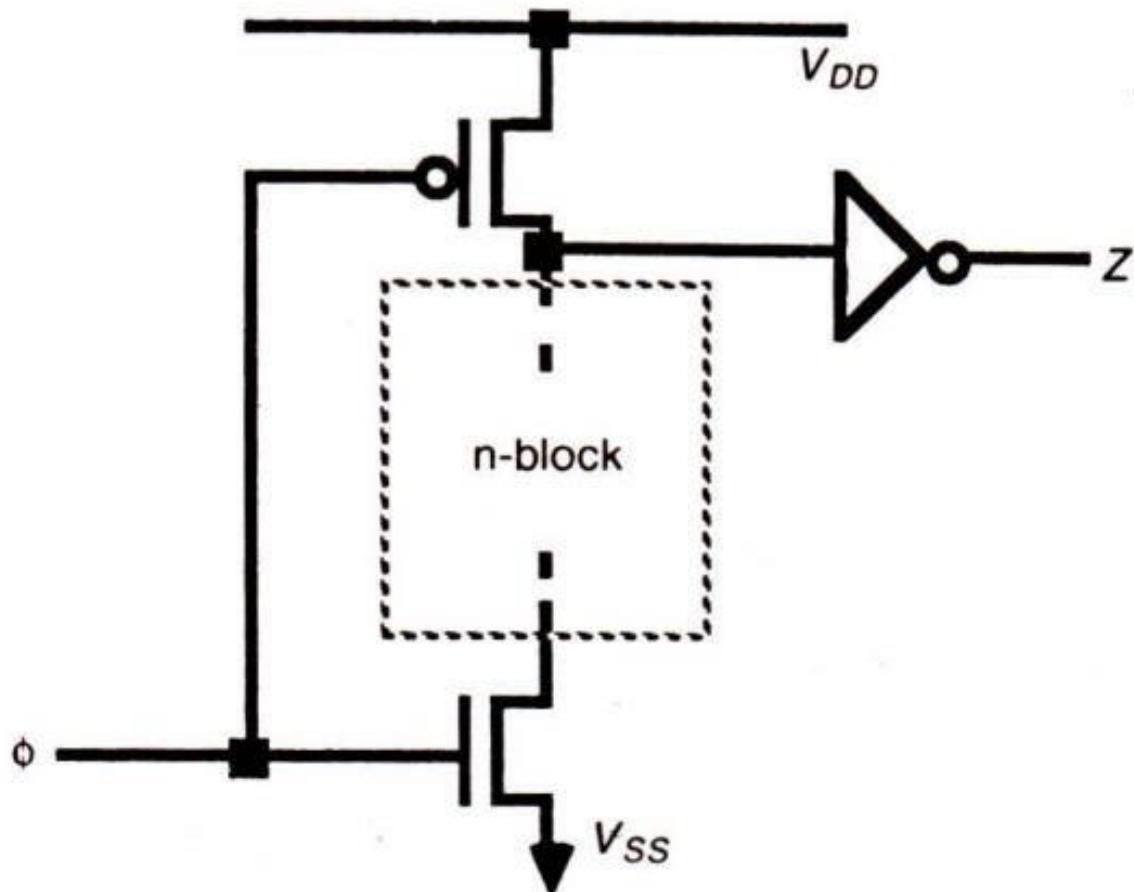
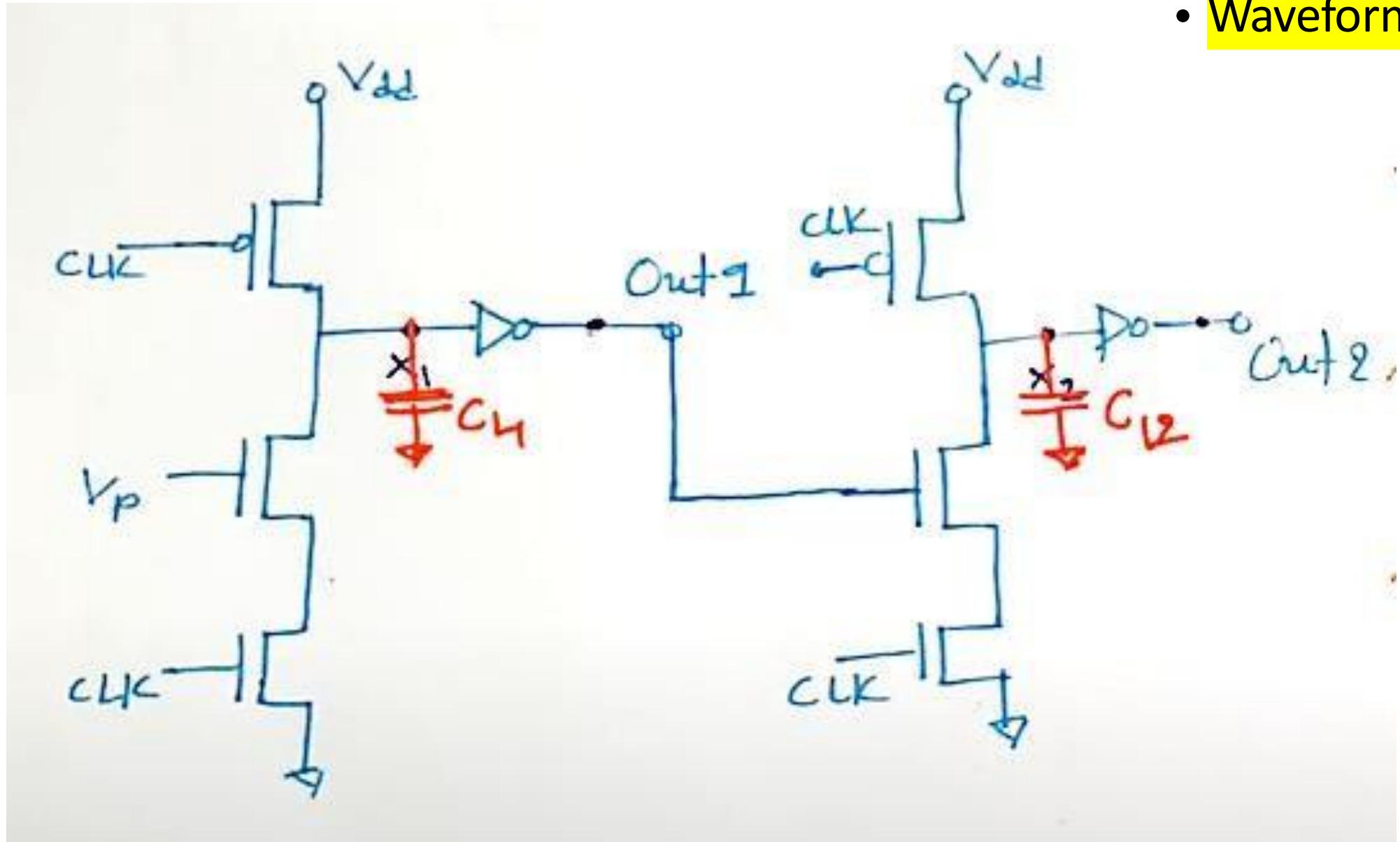


FIGURE 6.13 CMOS domino logic.

• Waveform:



Fabrication of MOSFETs

Mask

- Common material used for masks are Photoresist, Polysilicon, Silicon dioxide, Silicon nitride.
- To create mask:
 - (a) deposit mask material over entire surface
 - (b) cut windows in the mask to create exposed areas
 - (c) deposit dopant
 - (d) remove un-required mask material
- Masks plays important role in process called selective diffusions.
- The selective diffusion involves
 1. Patterning windows in a mask material on the surface of the wafer.
 2. Subjecting the exposed areas to a dopant source.
 3. Removing any un-required mask material.

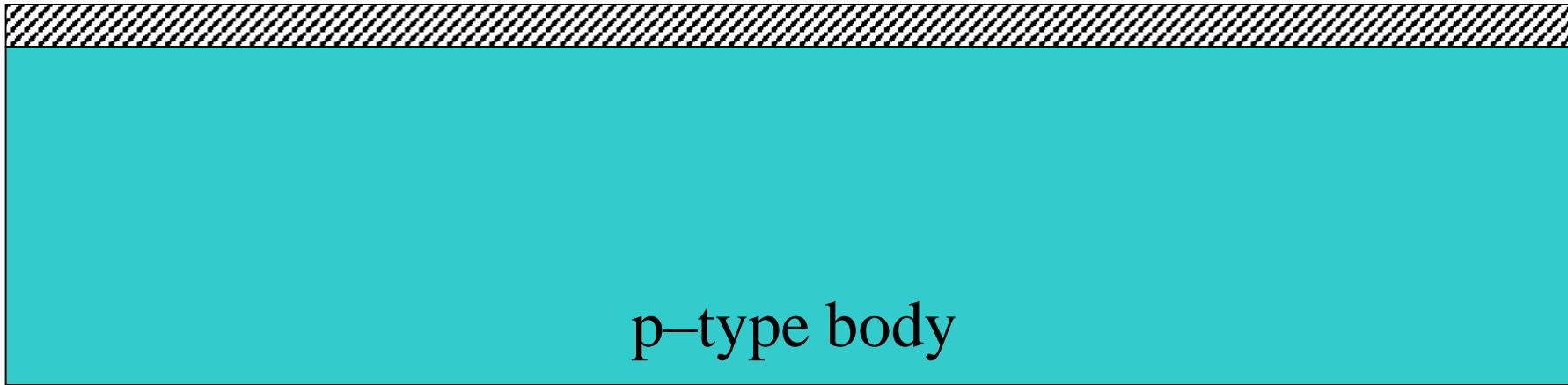
Photolithography

- The Process of using an optical image and a photosensitive film to produce a pattern on a substrate is **photolithography**
- Photolithography depends on a photosensitive film called a photo-resist.

Types of resist

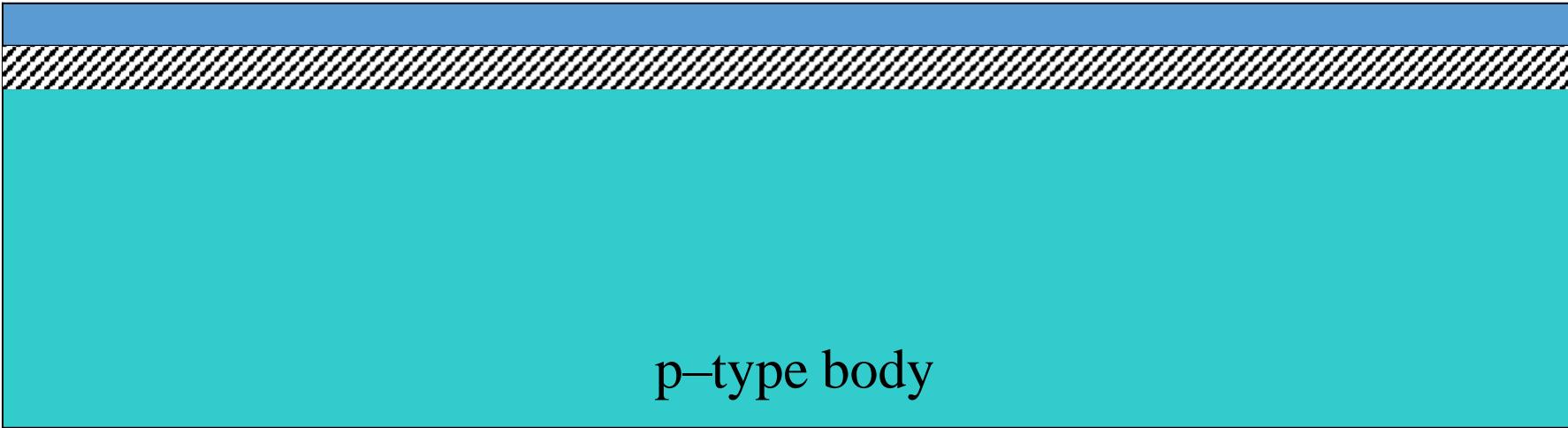
- **Positive resist**, a resist that become soluble when exposed and forms a positive image of the plate.
- **Negative resist**, a resist that lose solubility when illuminated forms a negative image of the plate.

Photolithography



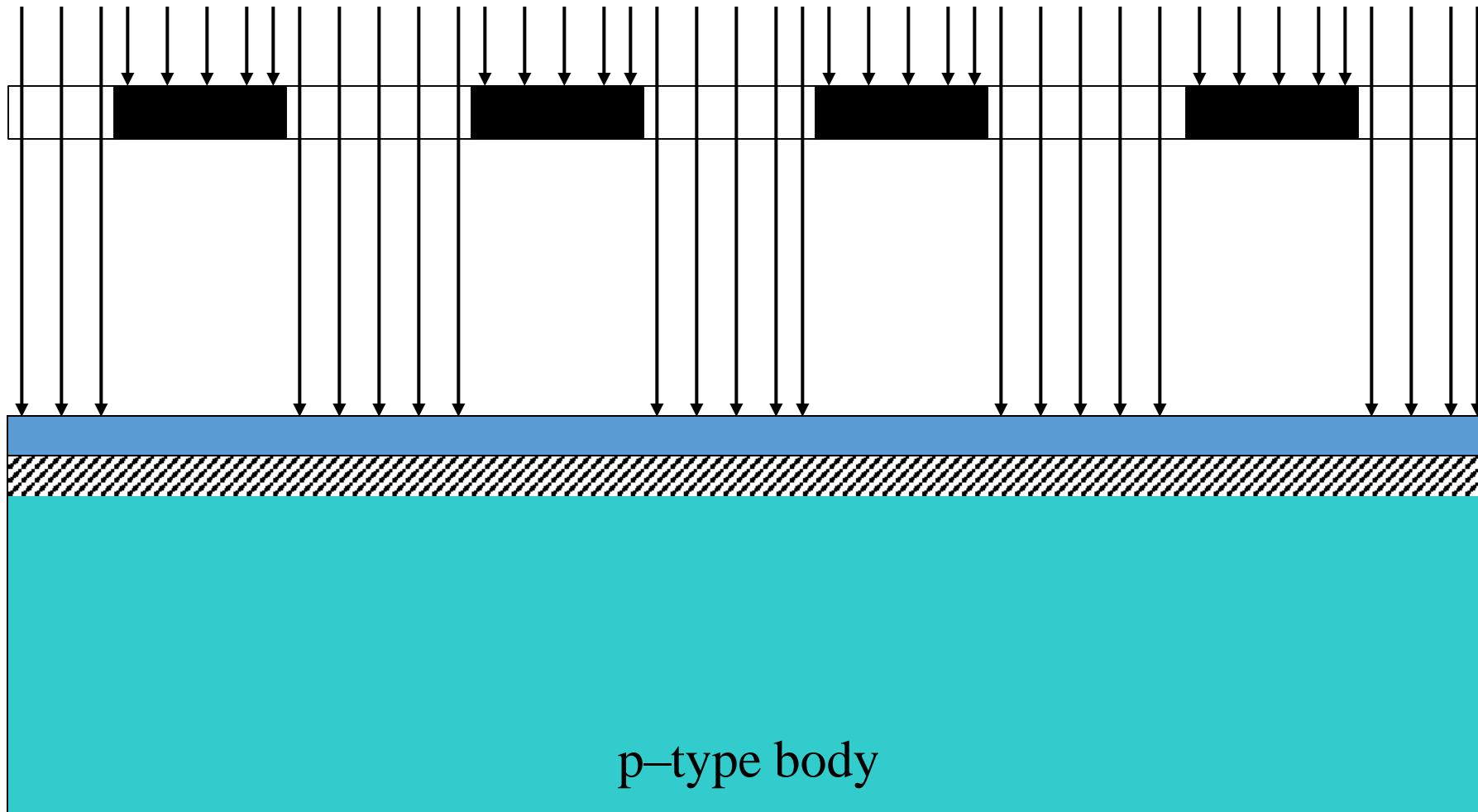
Substrate

Photolithography



Resist application

Photolithography

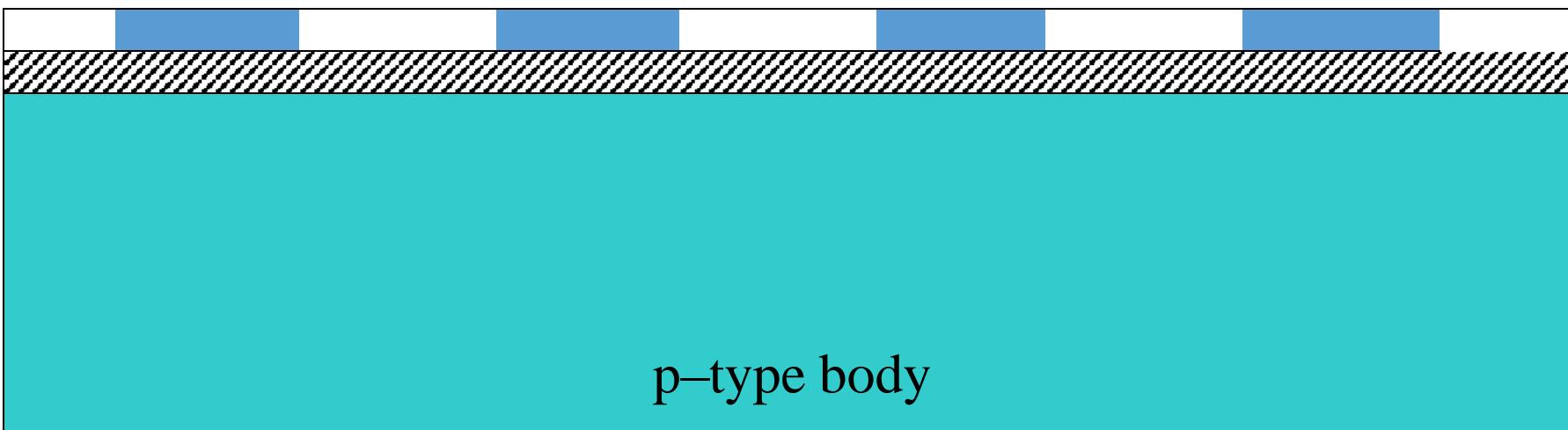


Exposure

Photolithography



Etching

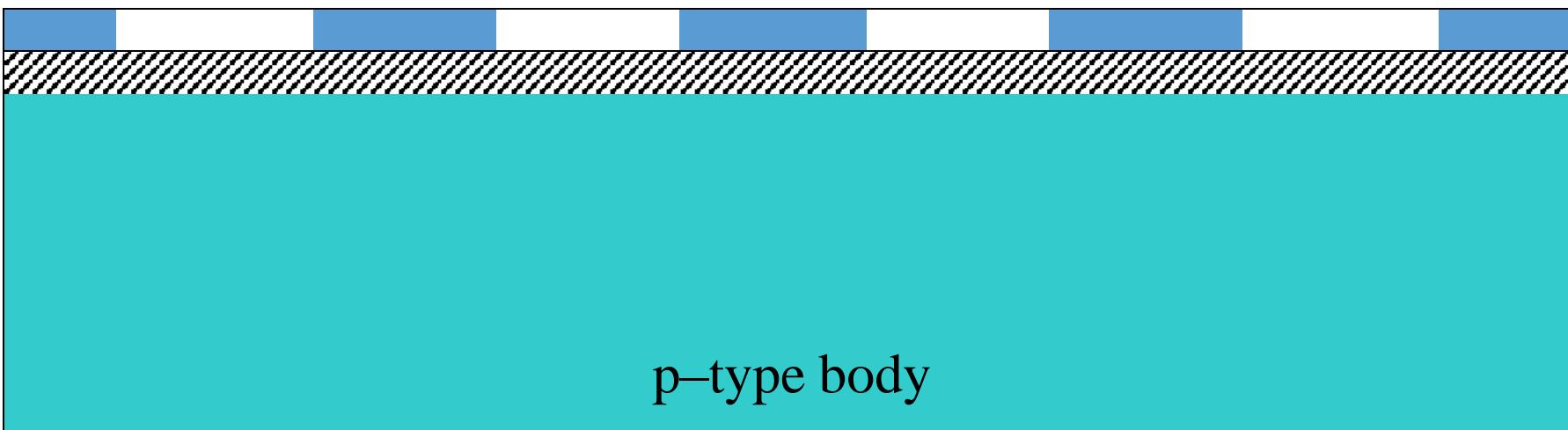


Positive Resist

Photolithography



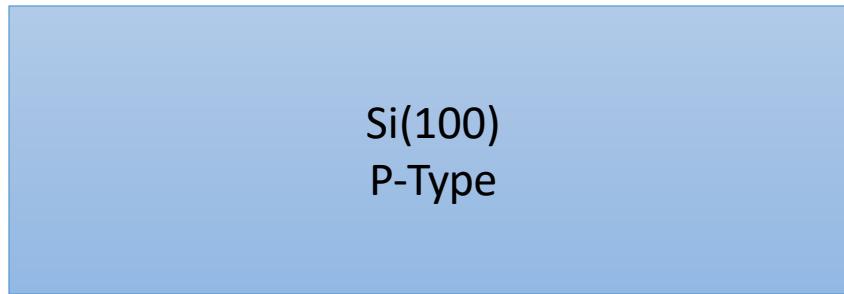
Etching



Negative Resist

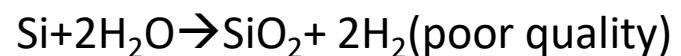
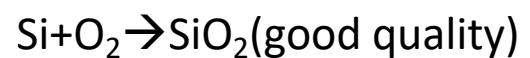
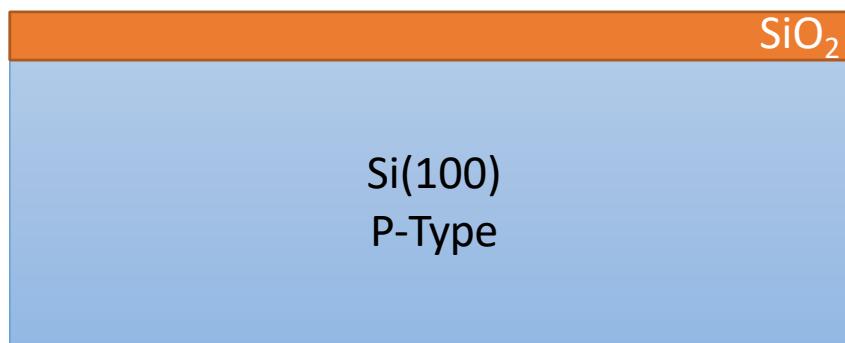
NMOS Fabrication Steps

1. Selection of Substrate

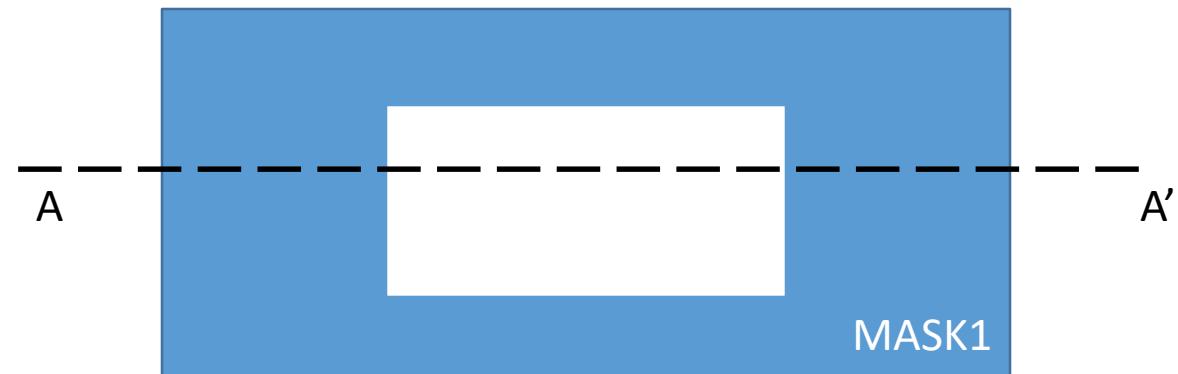
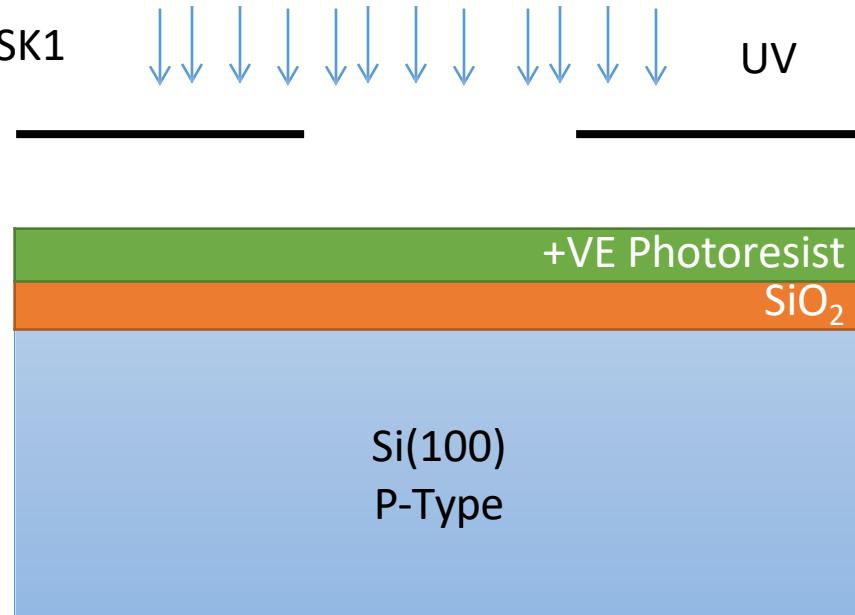


2. Cleaning

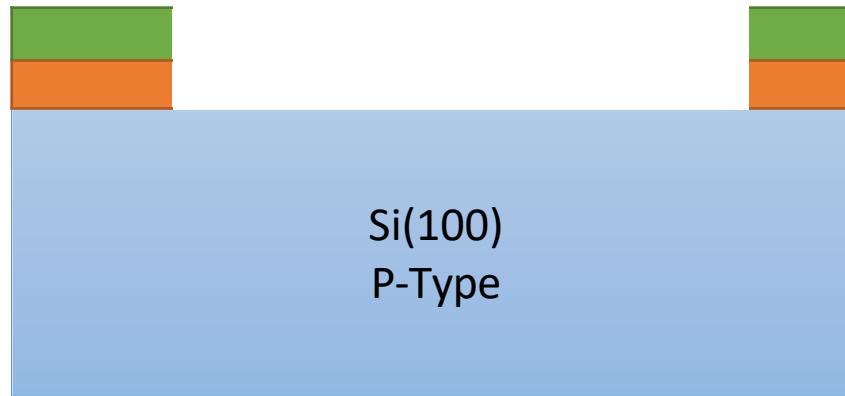
3. Oxidation



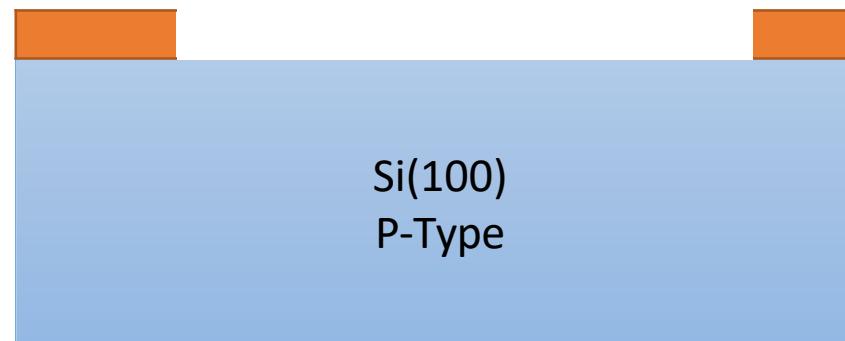
4. Lithography with MASK1



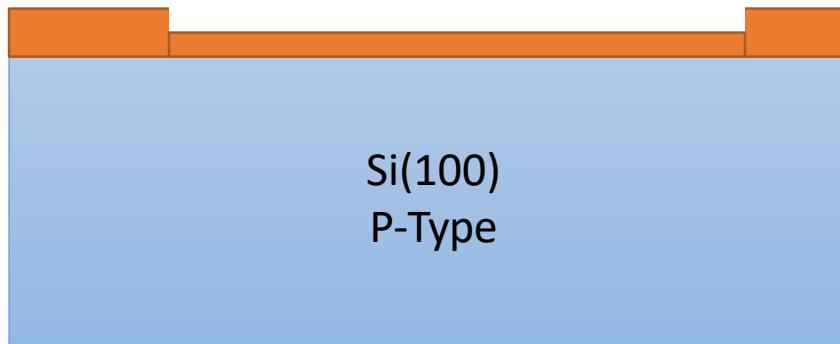
Photoresist development and Oxide Etching



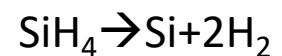
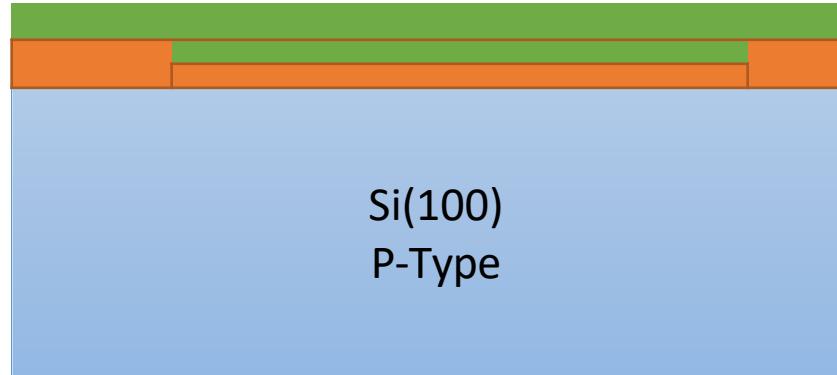
Photoresist Etching



Gate Oxidation

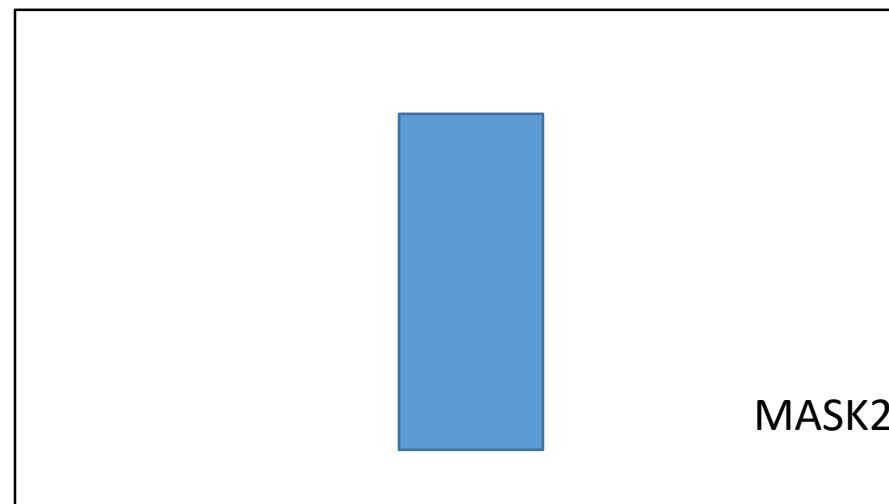
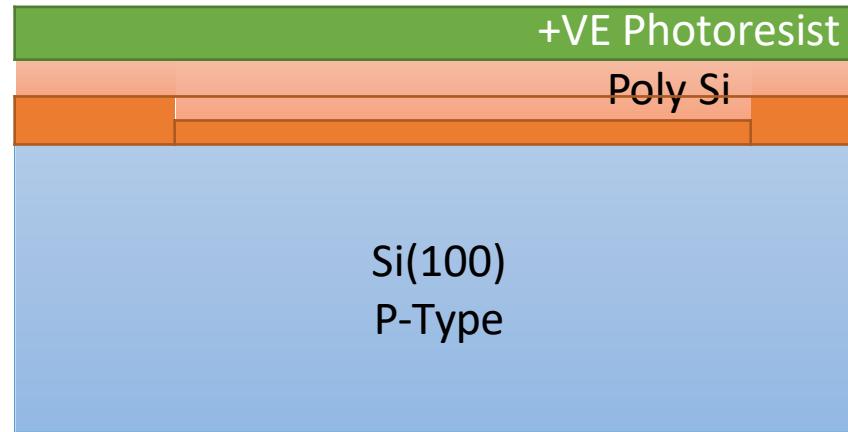


Poly Silicon Deposition

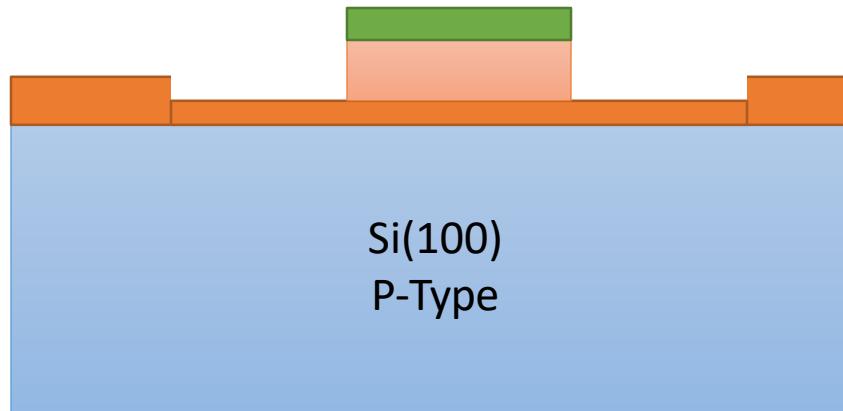


Lithography for Gate Electrode

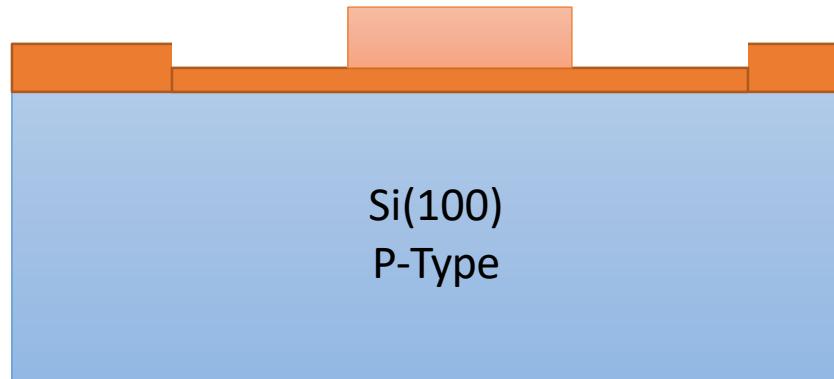
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ UV



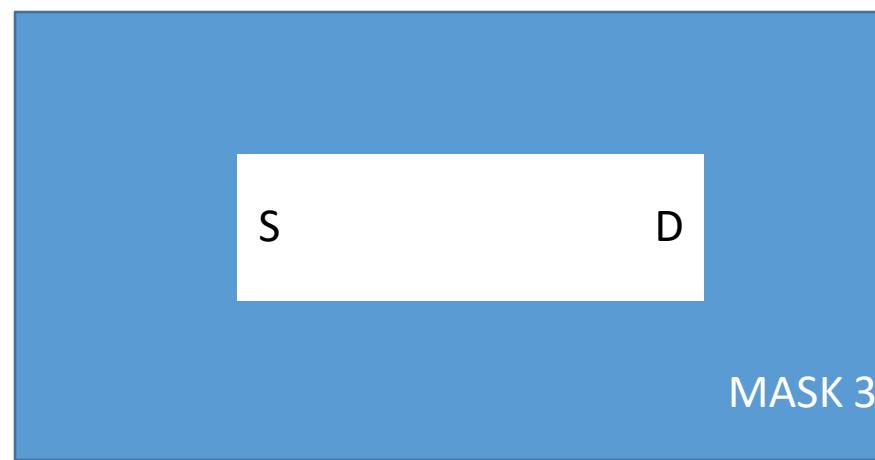
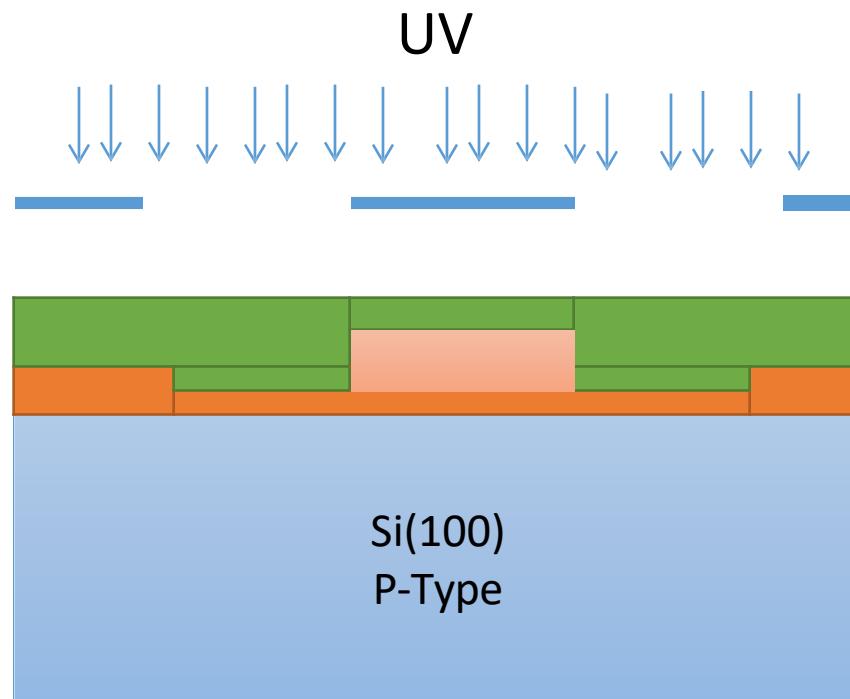
Poly Patterning



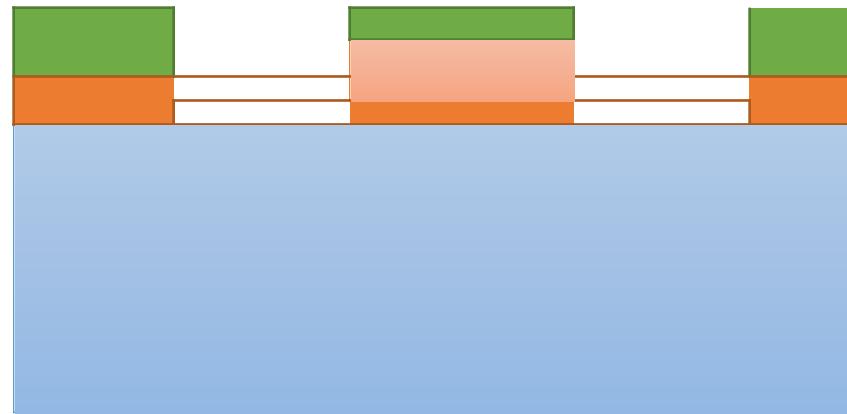
Photoresist Cleaning



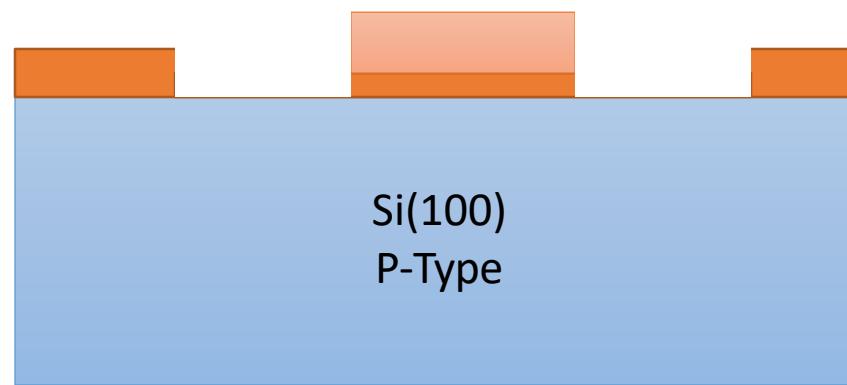
Lithography for Source and Drain region



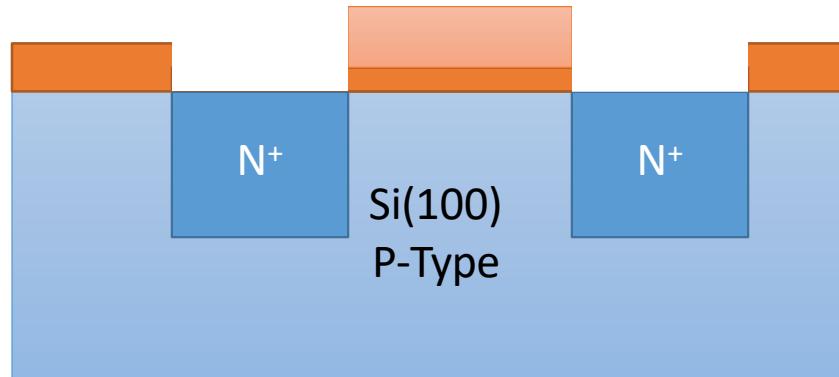
Oxide etching (HF Cleaning)



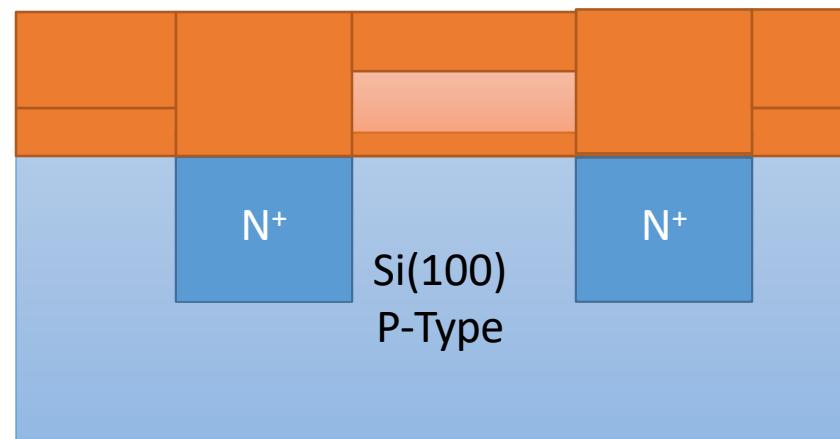
Photoresist cleaning



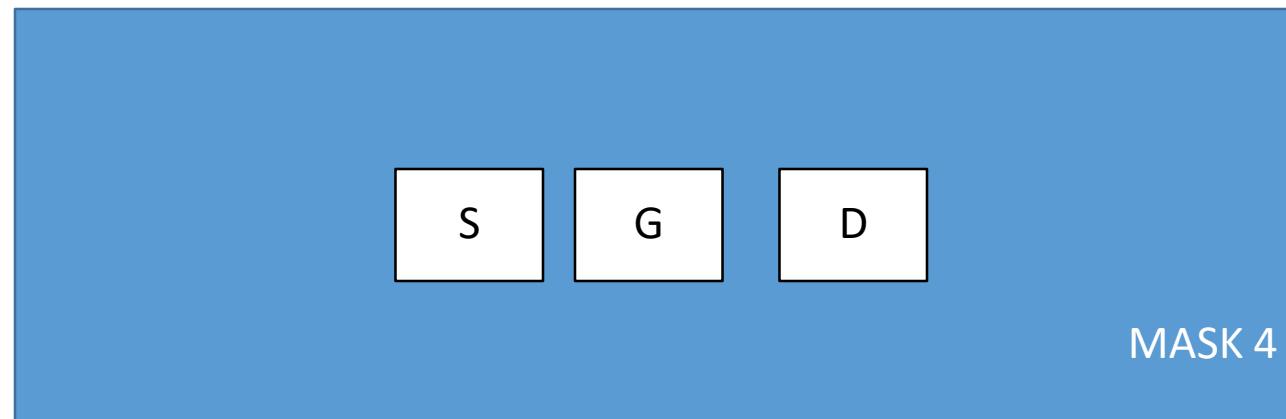
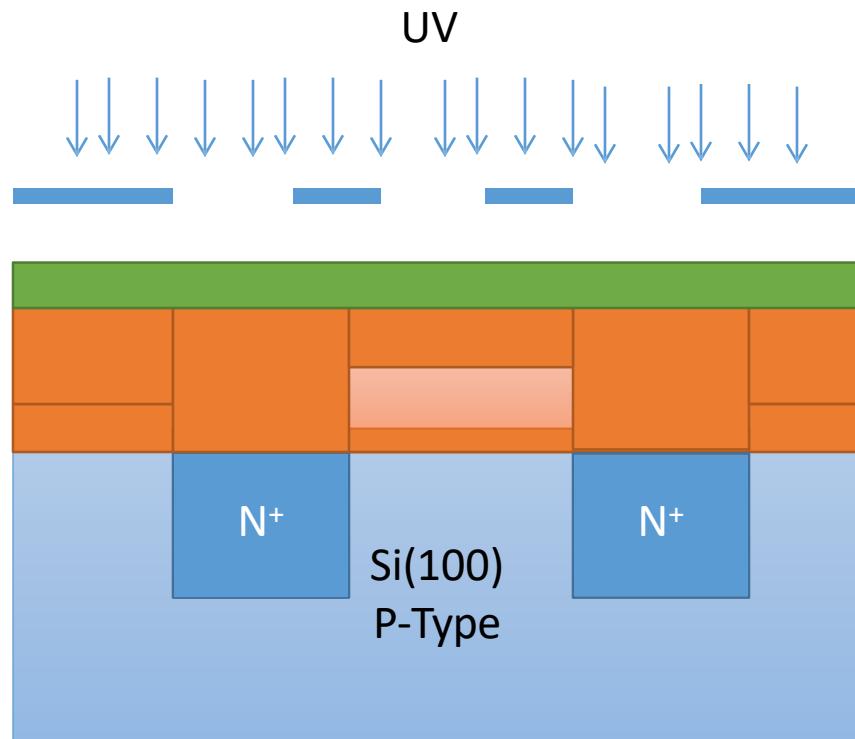
Ion Implantation



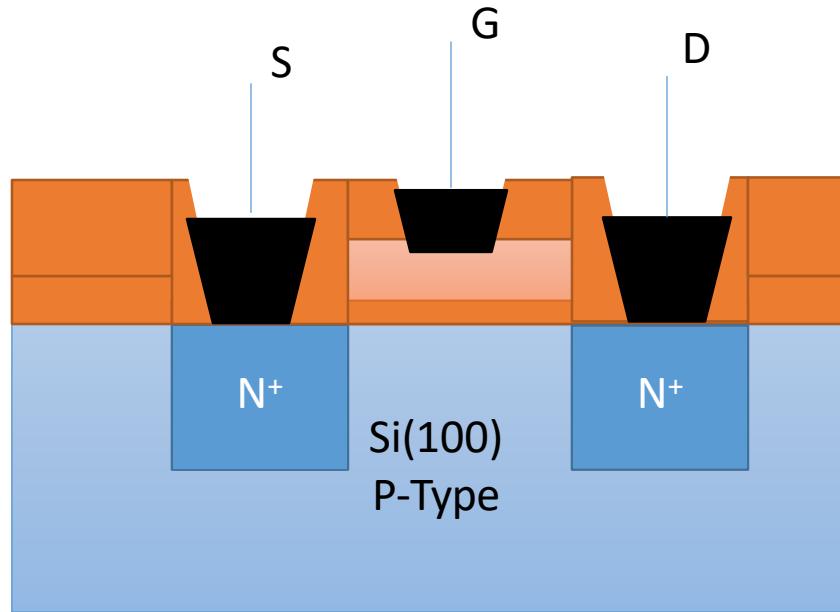
Thick Oxide Deposition



Lithography and Contact Opening



Metallization and Patterning



Body terminal is not shown.....

Fabrication of CMOS Devices

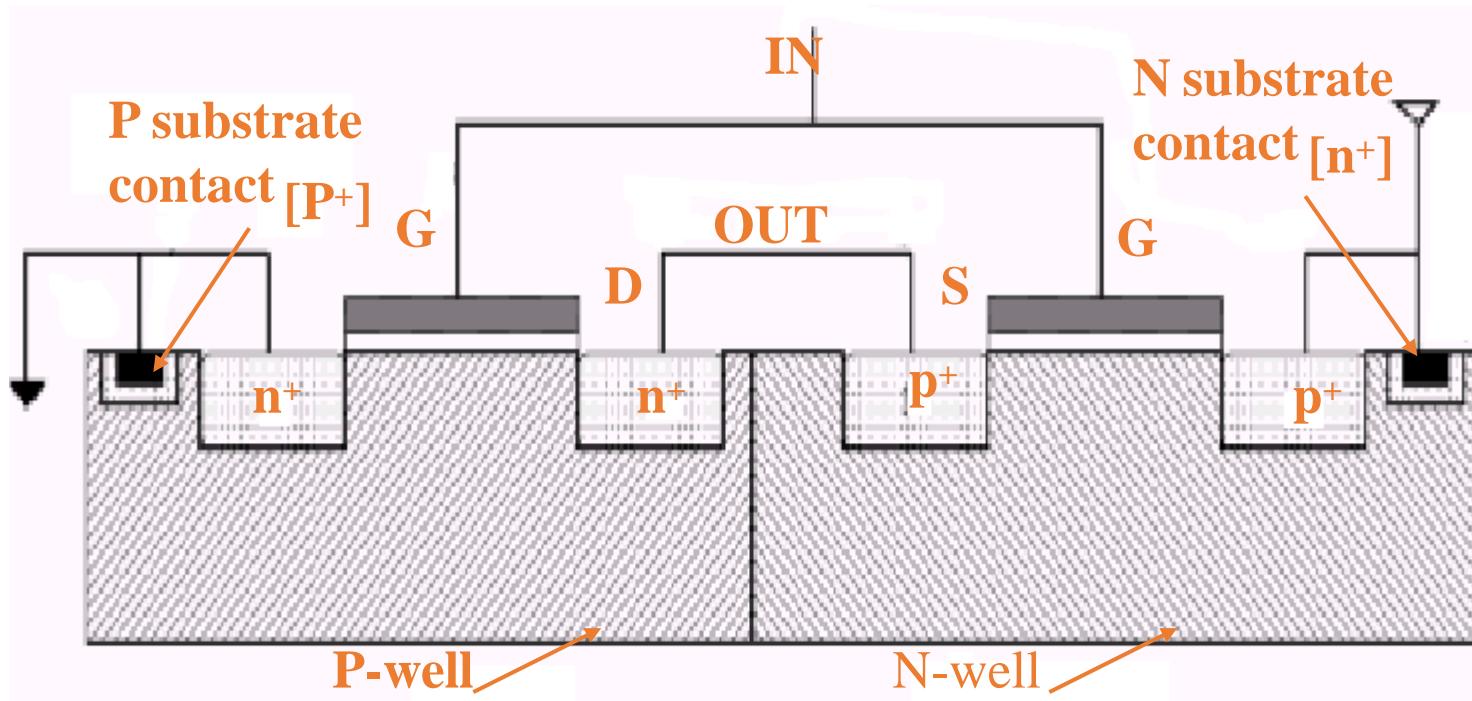
Technologies used for CMOS fabrications include

- **N-well process**
- **P-well process**
- **Twin-tub process**
- **Silicon on insulator.**

P-Wells and N-Wells

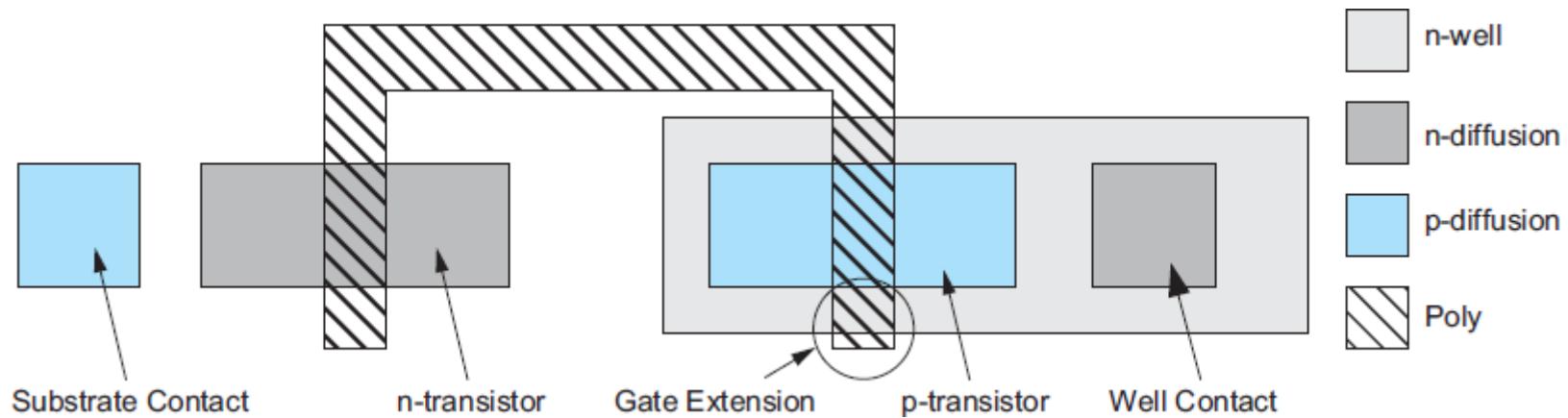
- In order to have both types of transistors on the same substrate, the substrate is divided into “well” regions (*Shaded region in the standard cells*)
- Two types of wells are available - n- well and p- well
- In a p- substrate, an n- well is used to create a local region of n type substrate, wherein the designer can create p- transistors
- In a n- substrate, a p- well creates a local p- type substrate region, to accommodate the n- transistors.
- Hence, every p- device is surrounded by an n- well, that must be connected to V_{DD} via a V_{DD} substrate contact.
- Similarly, n- devices are surrounded by p- well connected to GND using a GND substrate contact.

P-Wells and N-Wells



- A p- transistor is built on an n- substrate and an n- transistor is built on a p-substrate

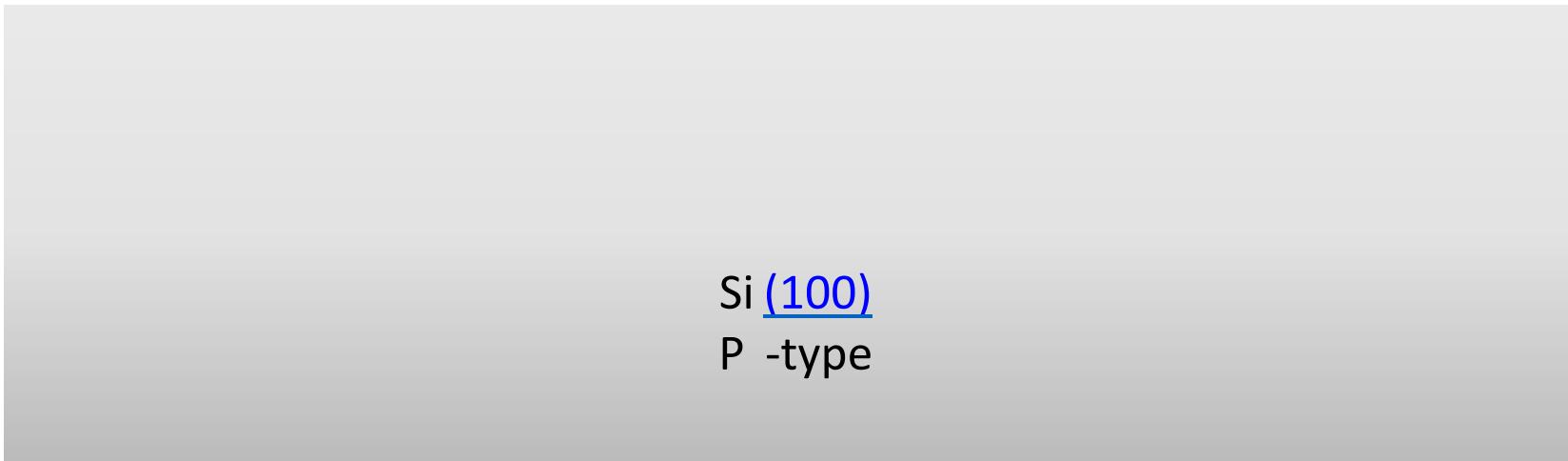
MASK for CMOS Inverter



N-Well CMOS Inverter



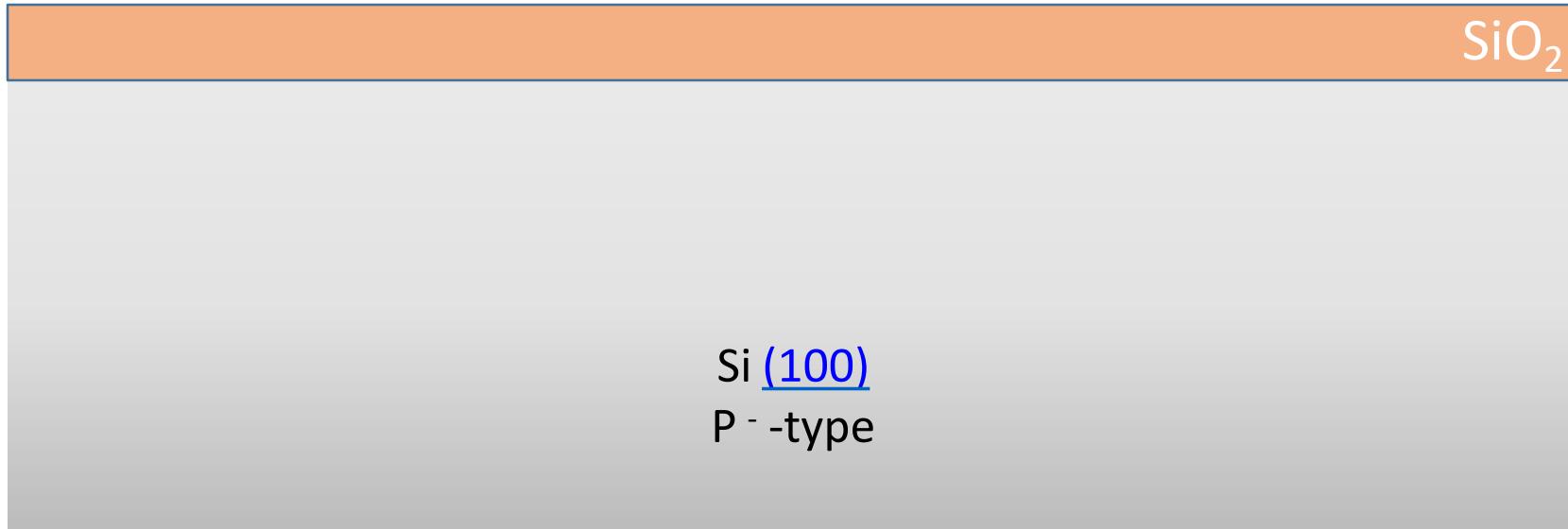
1.Selection of Substrate



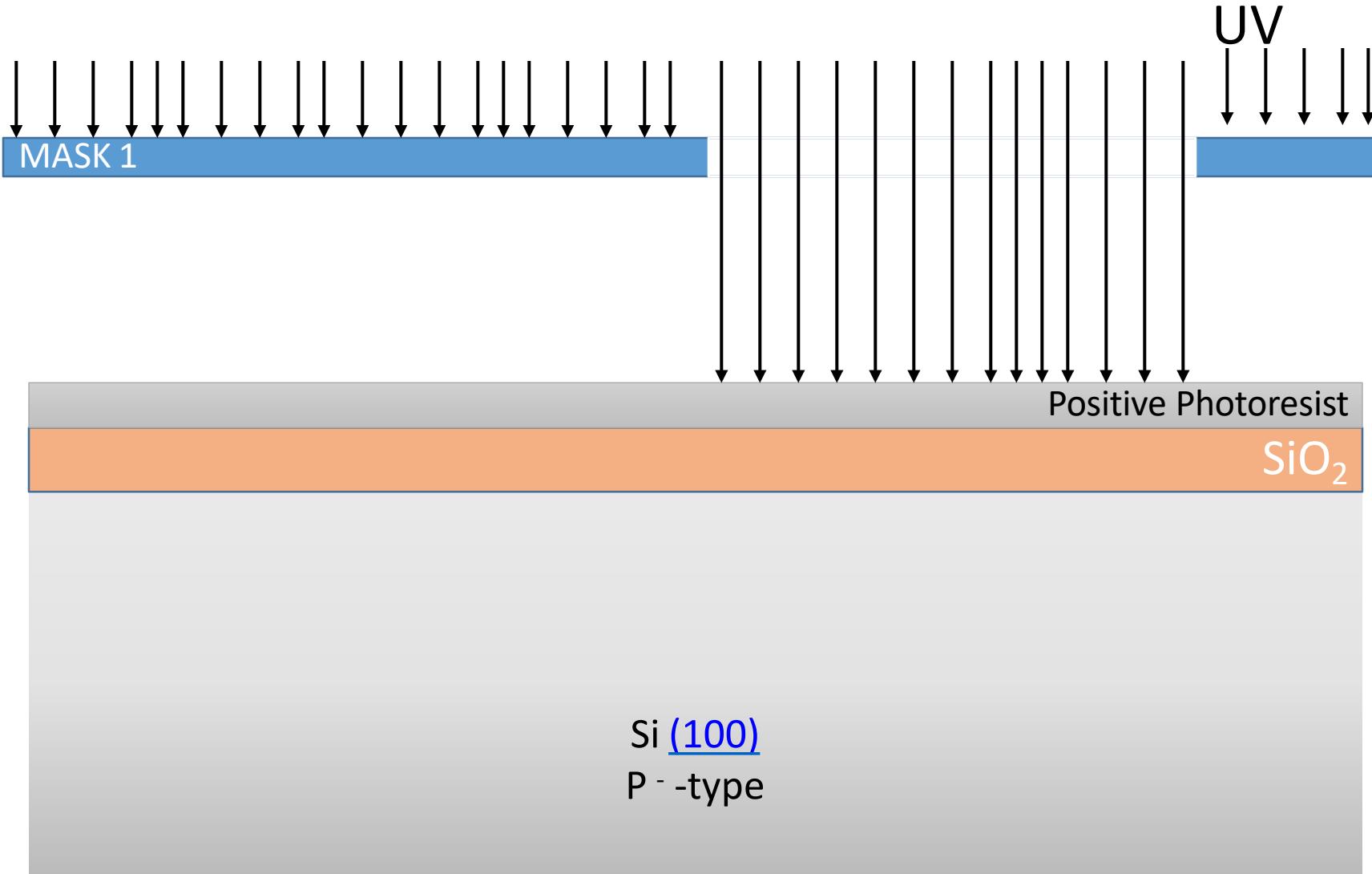
The n-well CMOS process starts with a moderately doped (with impurity concentration typically less than 10^{15} cm^{-3}) P-type silicon substrate.

2. Cleaning of Substrate

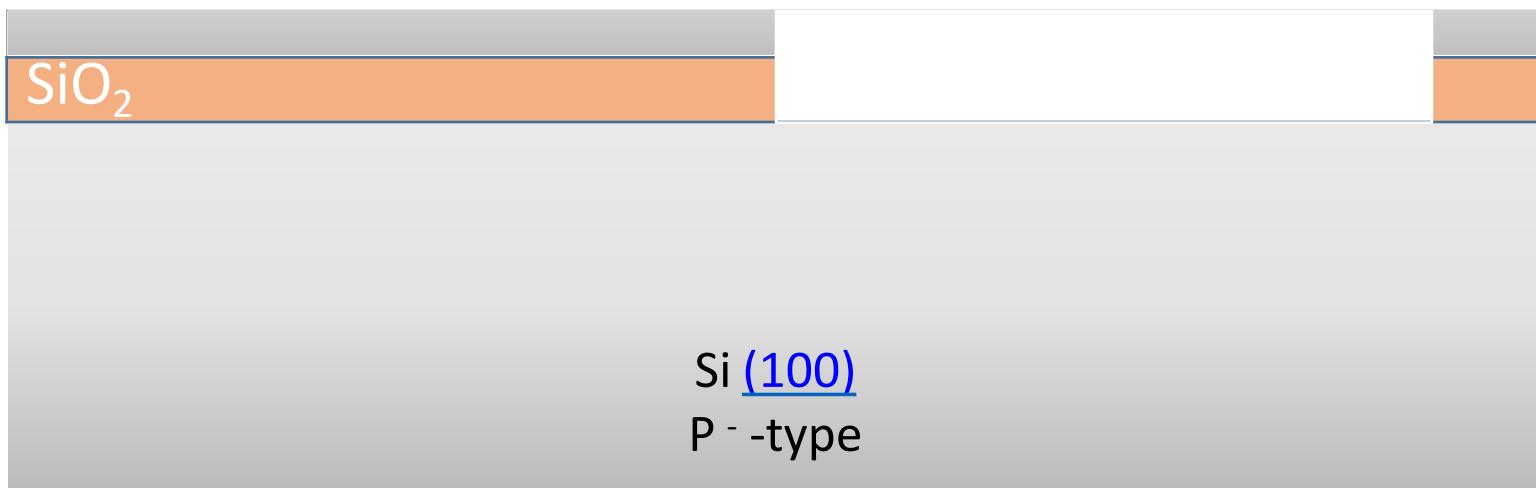
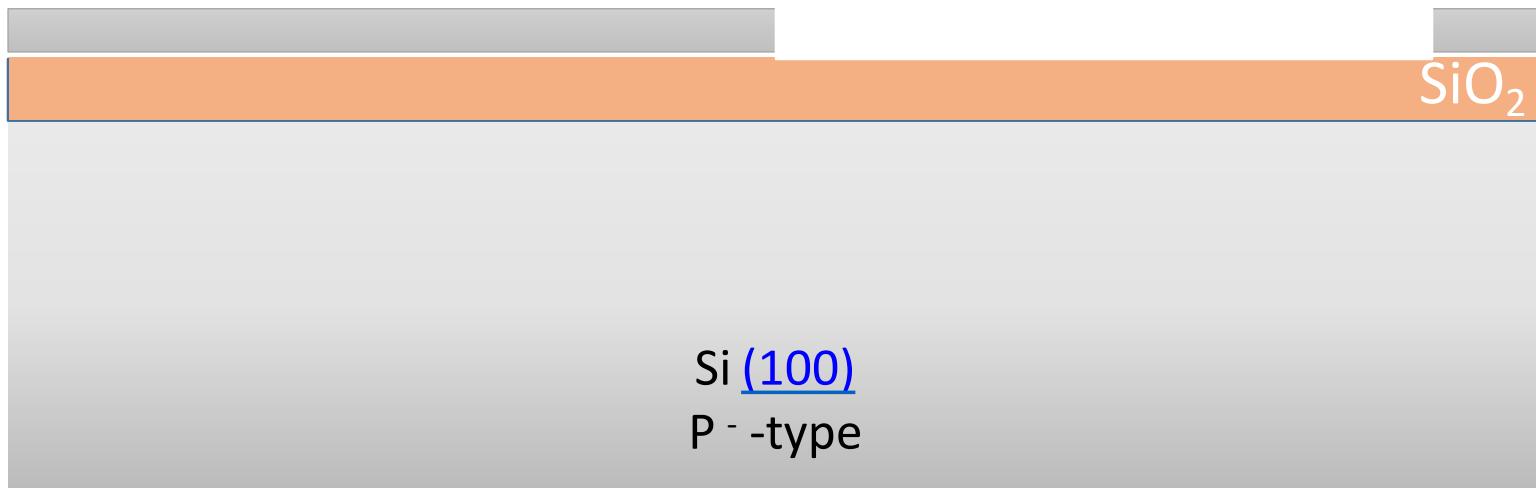
3. Oxidation(1μm)



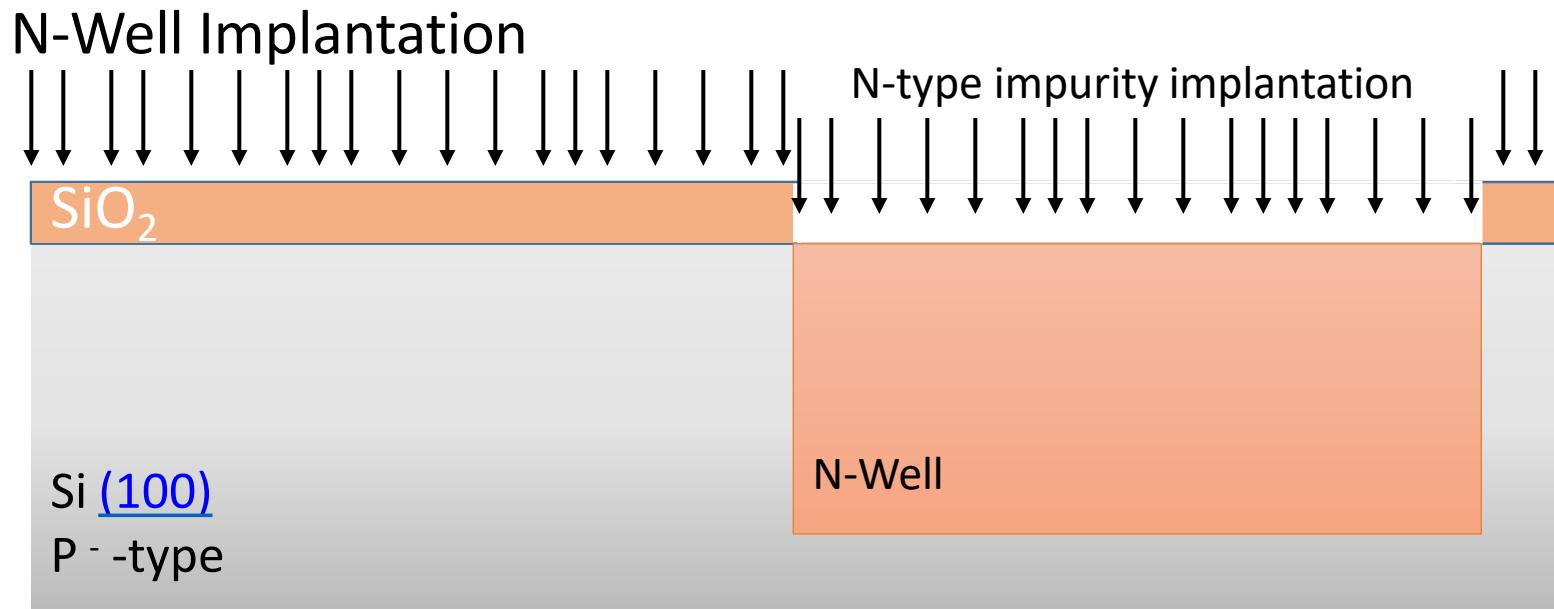
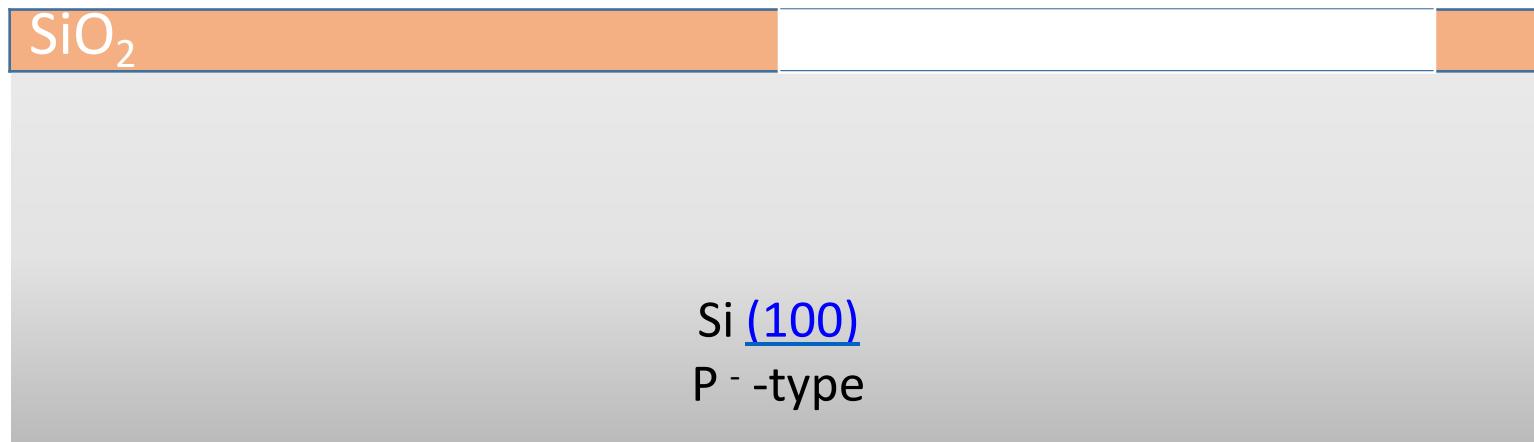
Lithography(MASK1)



HF(Oxide) Cleaning & PR Etching



HF(Oxide) Cleaning & PR Etching



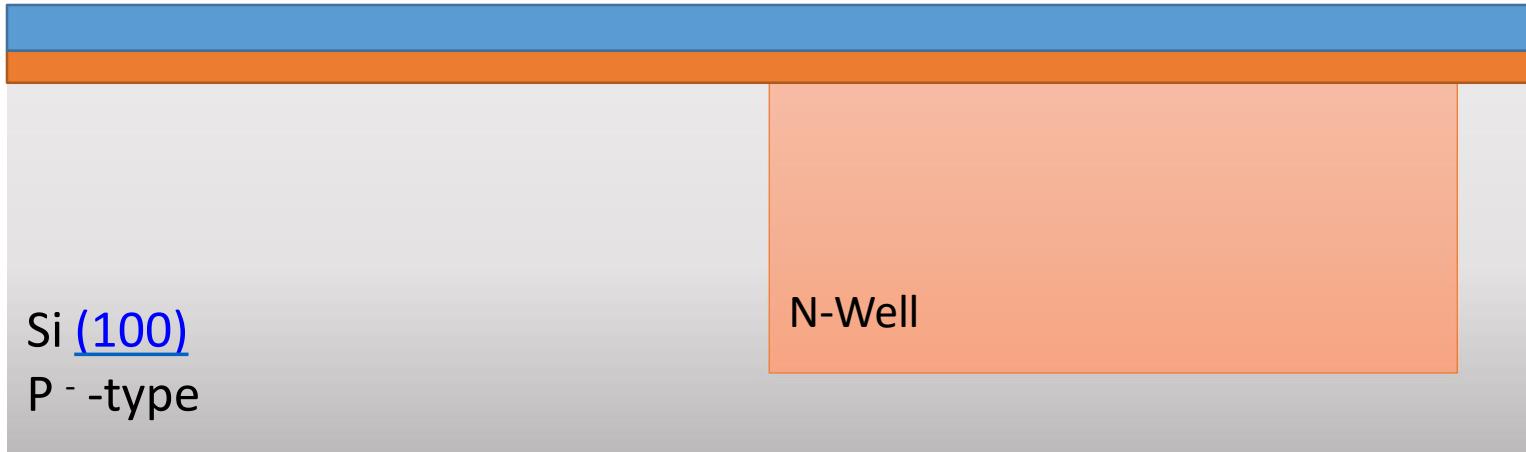
Oxide Cleaning



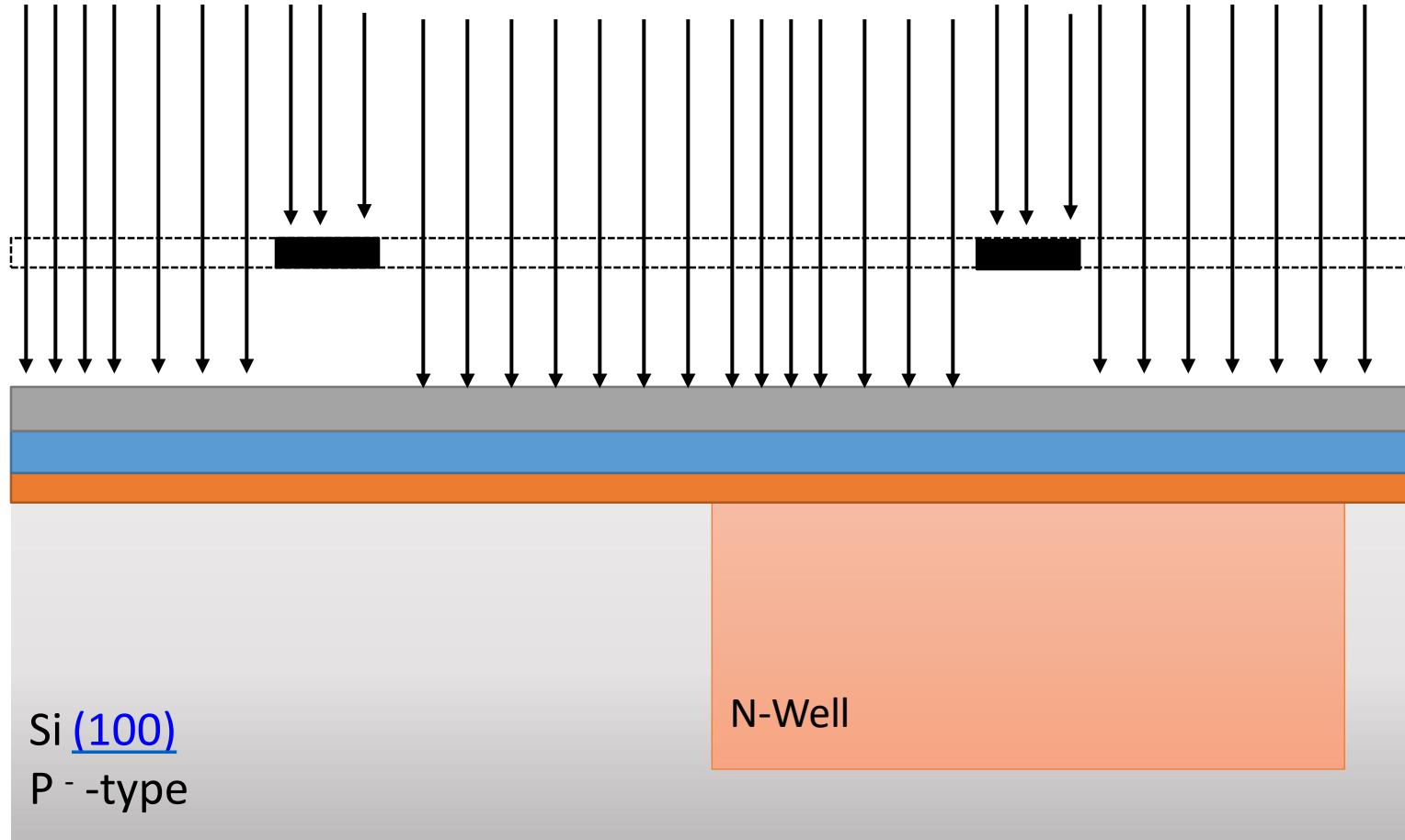
Gate Oxidation



Poly-Silicon Deposition

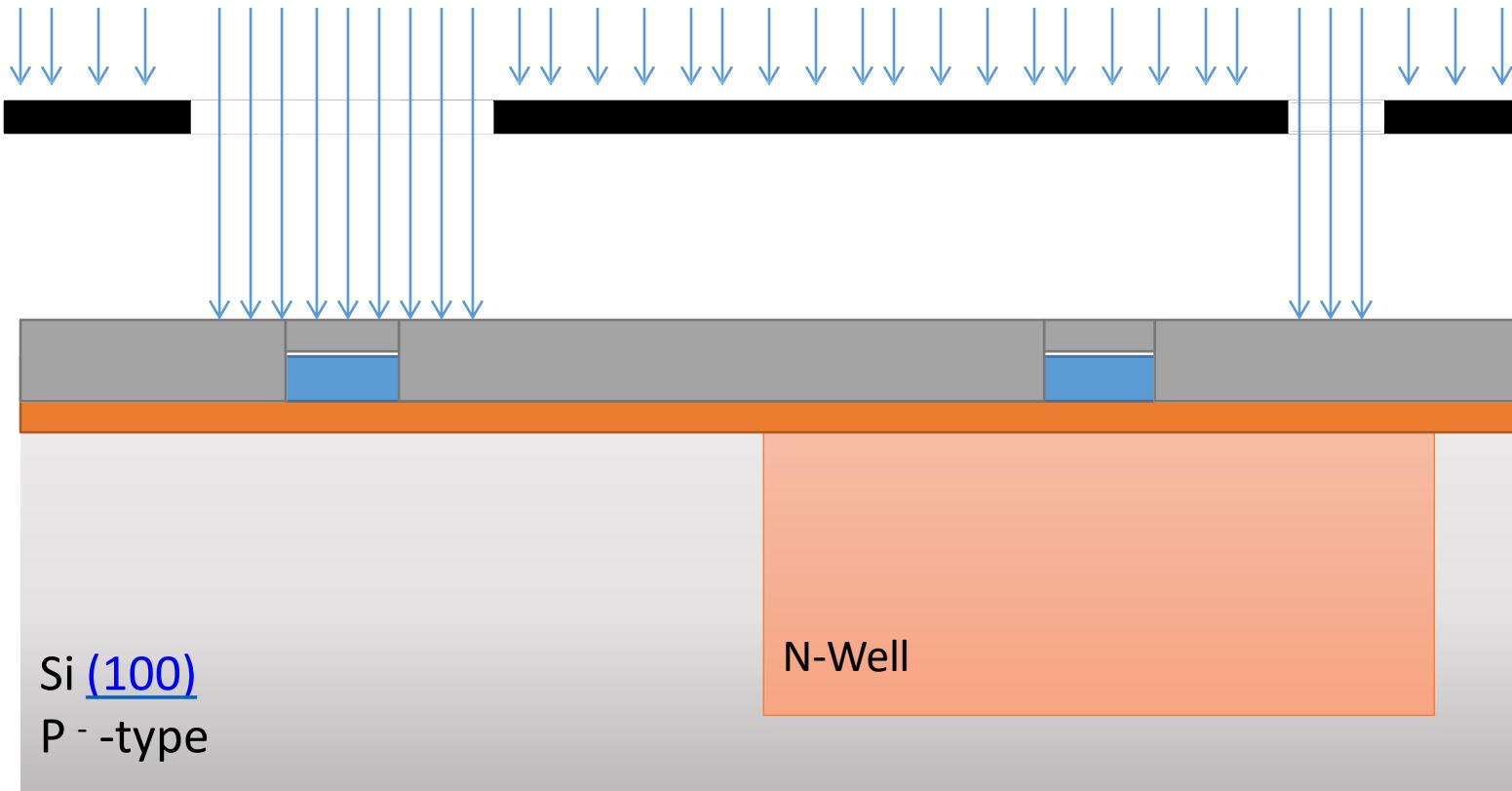


Lithography(MASK2)

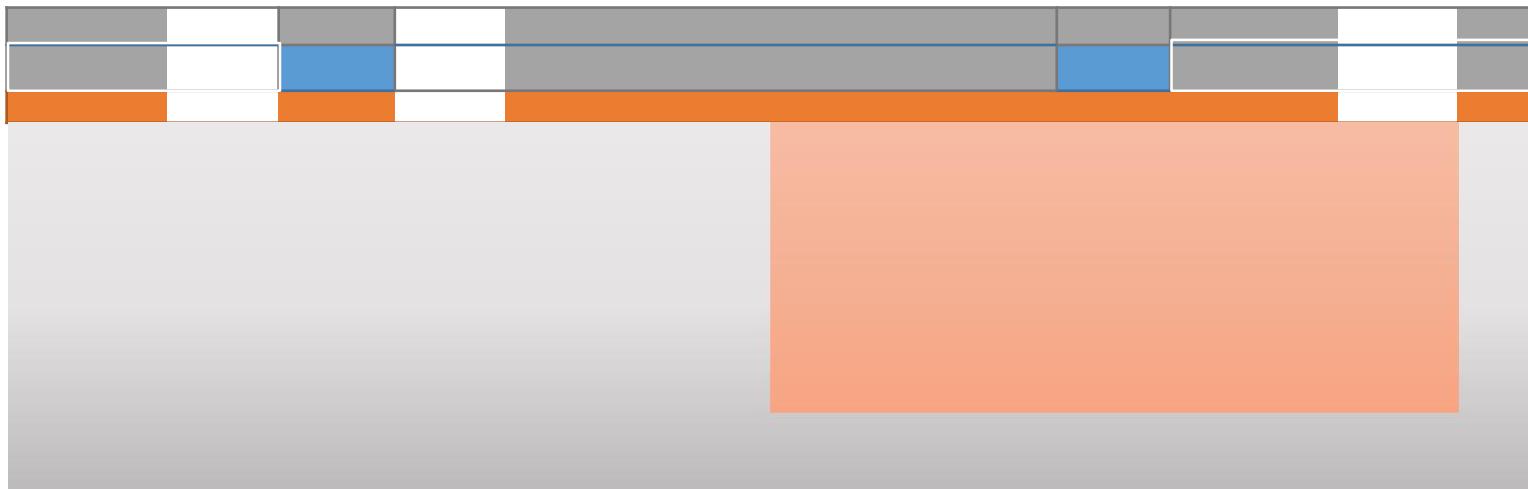




Lithography(MASK3)



HF(Oxide) Cleaning & PR Etching



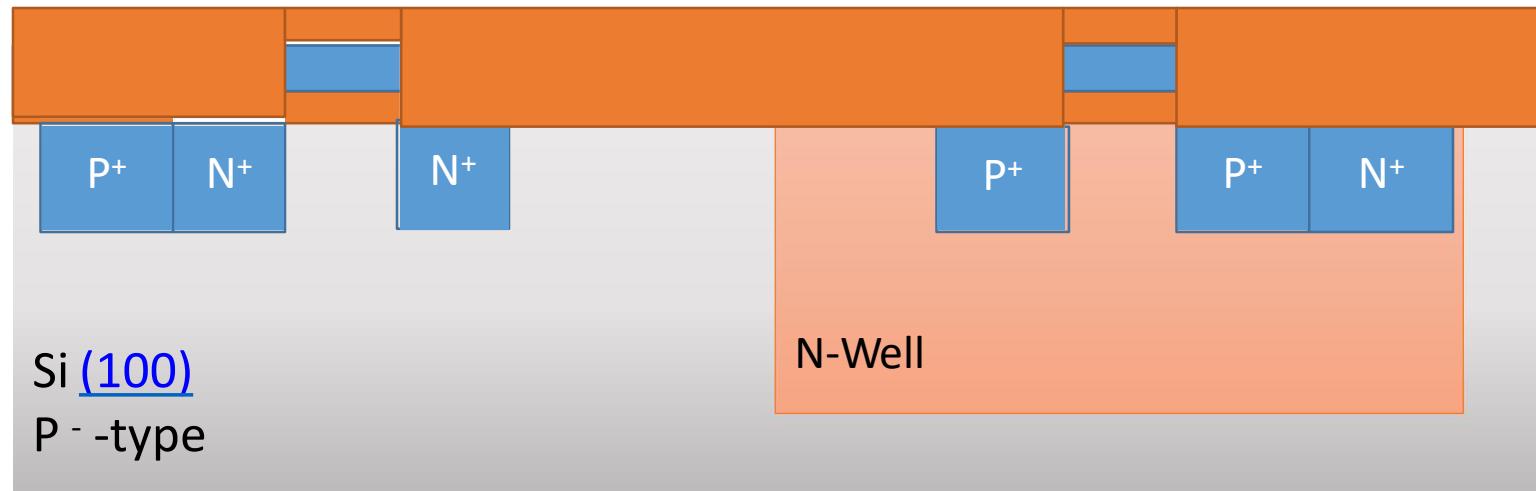
Ion Implantation



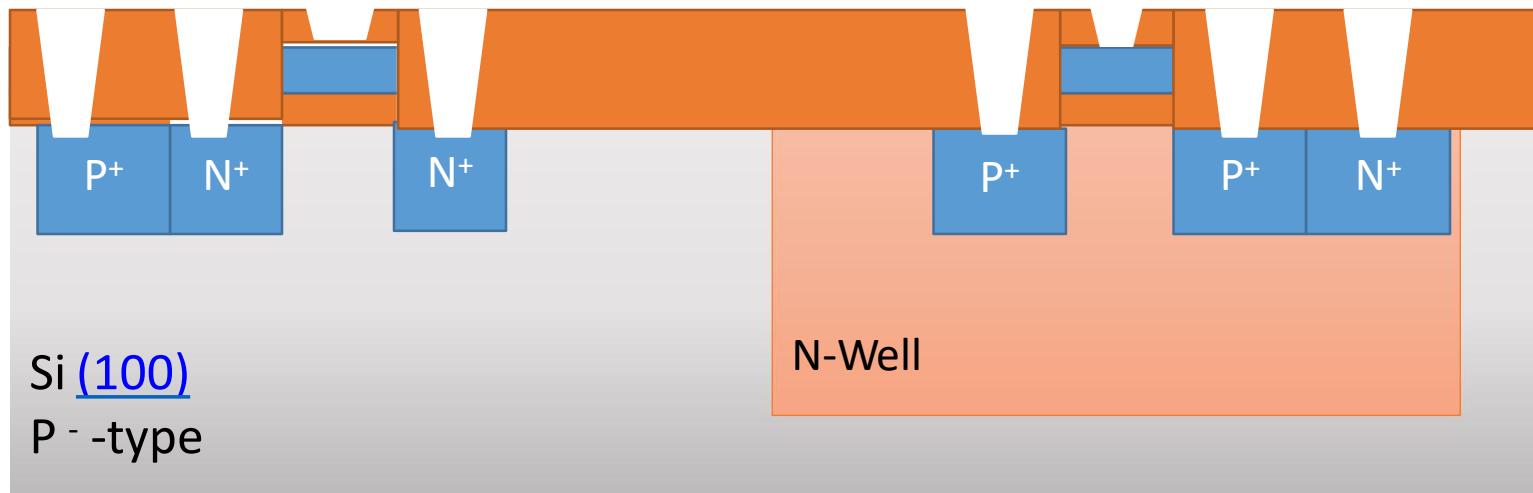
Repeat 12 to 14 for PMOS



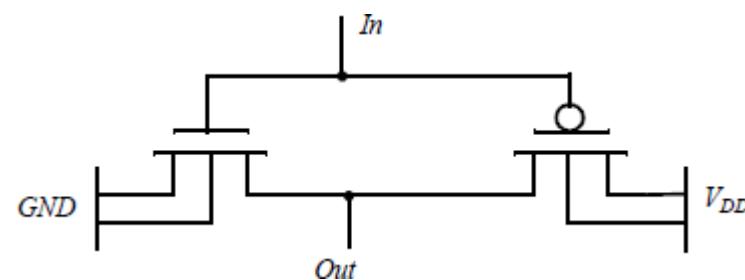
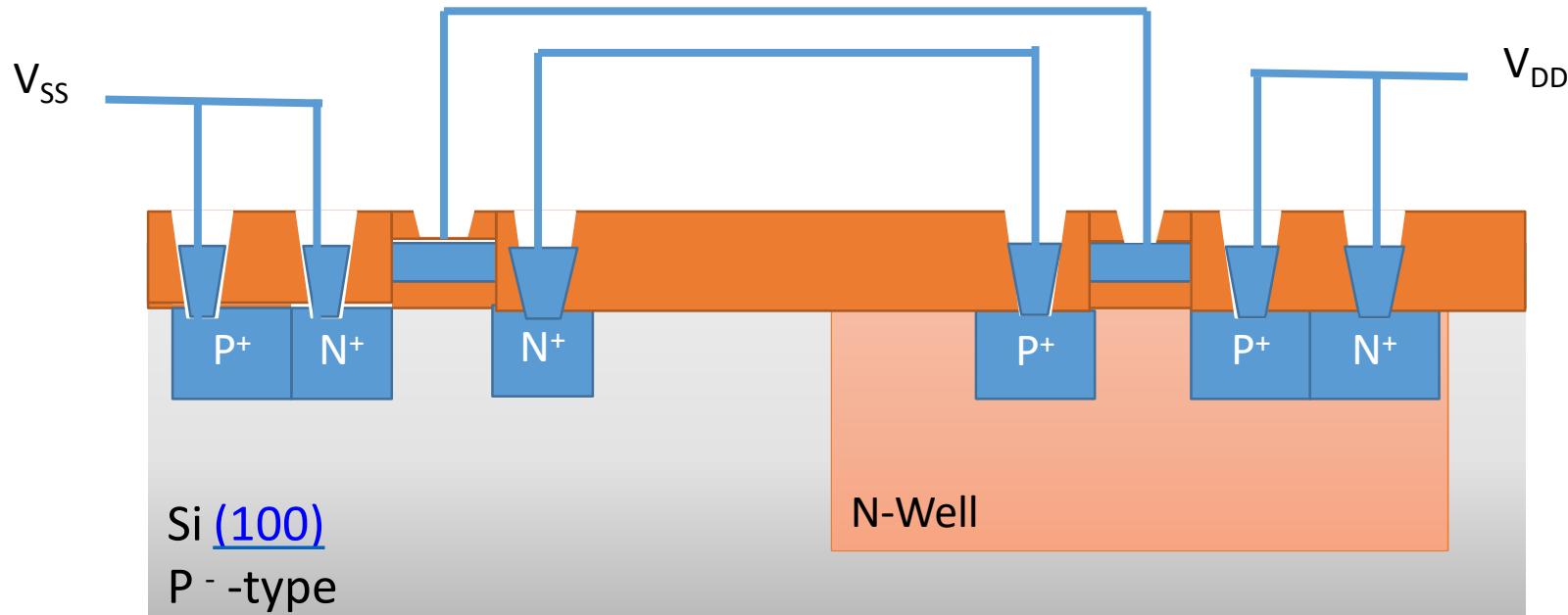
Thick Oxide Deposition



Lithography and Oxide Patterning



Metallization and Patterning



Twin-Tub CMOS Fabrication

1.Selection of Substrate

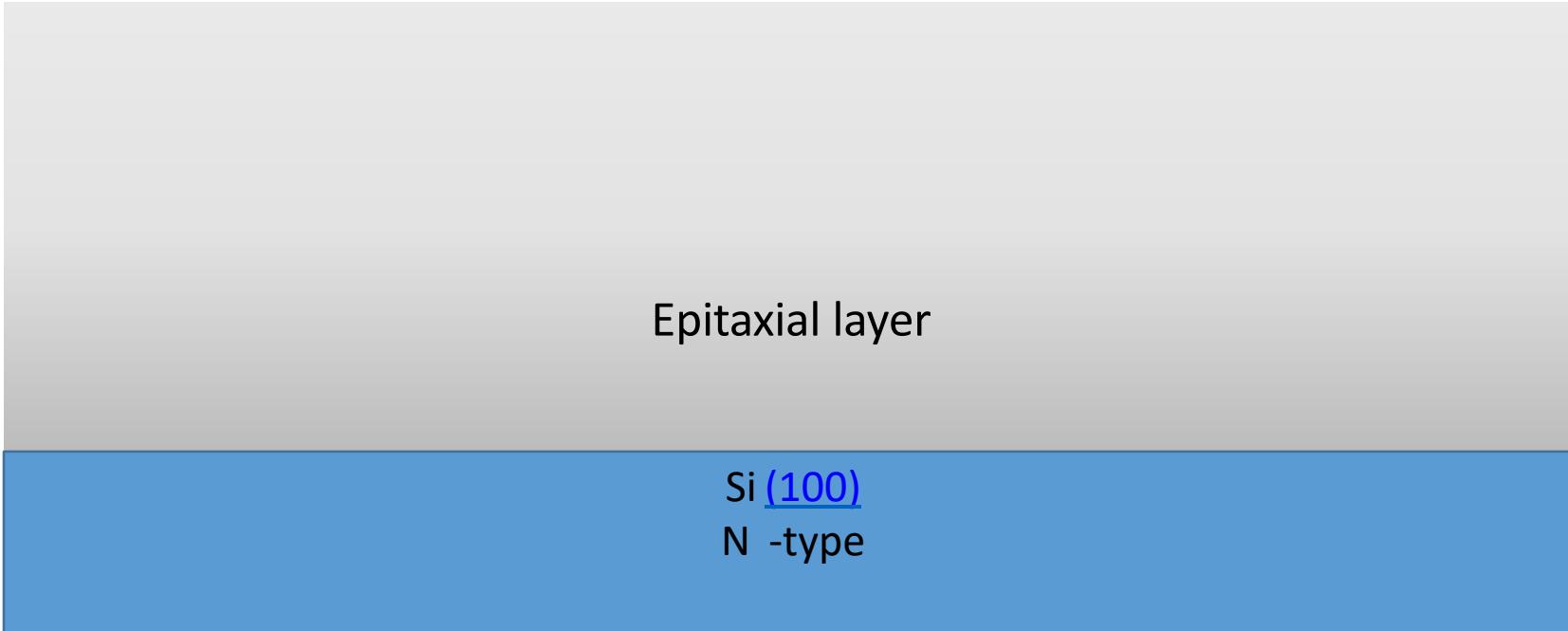
Si [\(100\)](#)

N -type

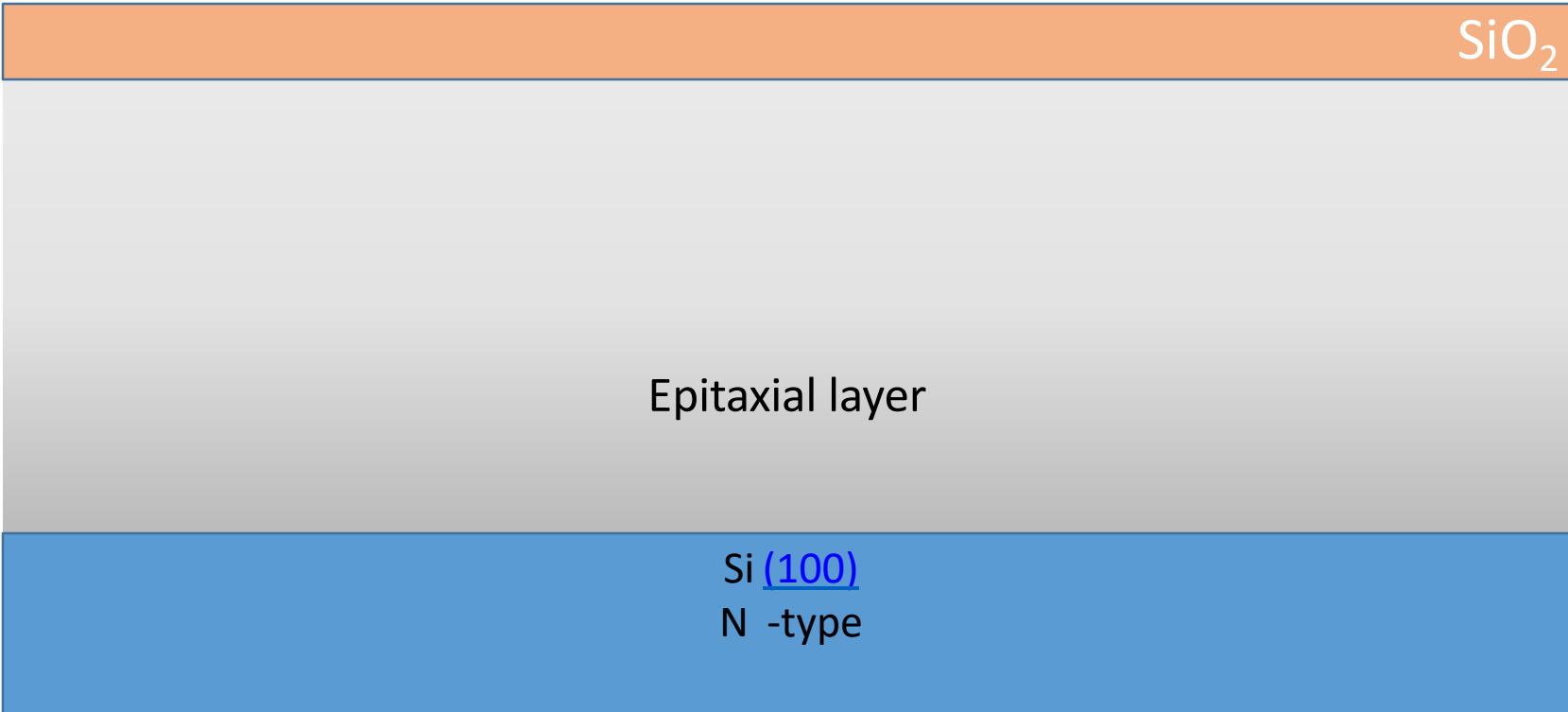
The twin-tub CMOS process starts with a high resistive n-type (100) silicon substrate.

2. Cleaning of Substrate

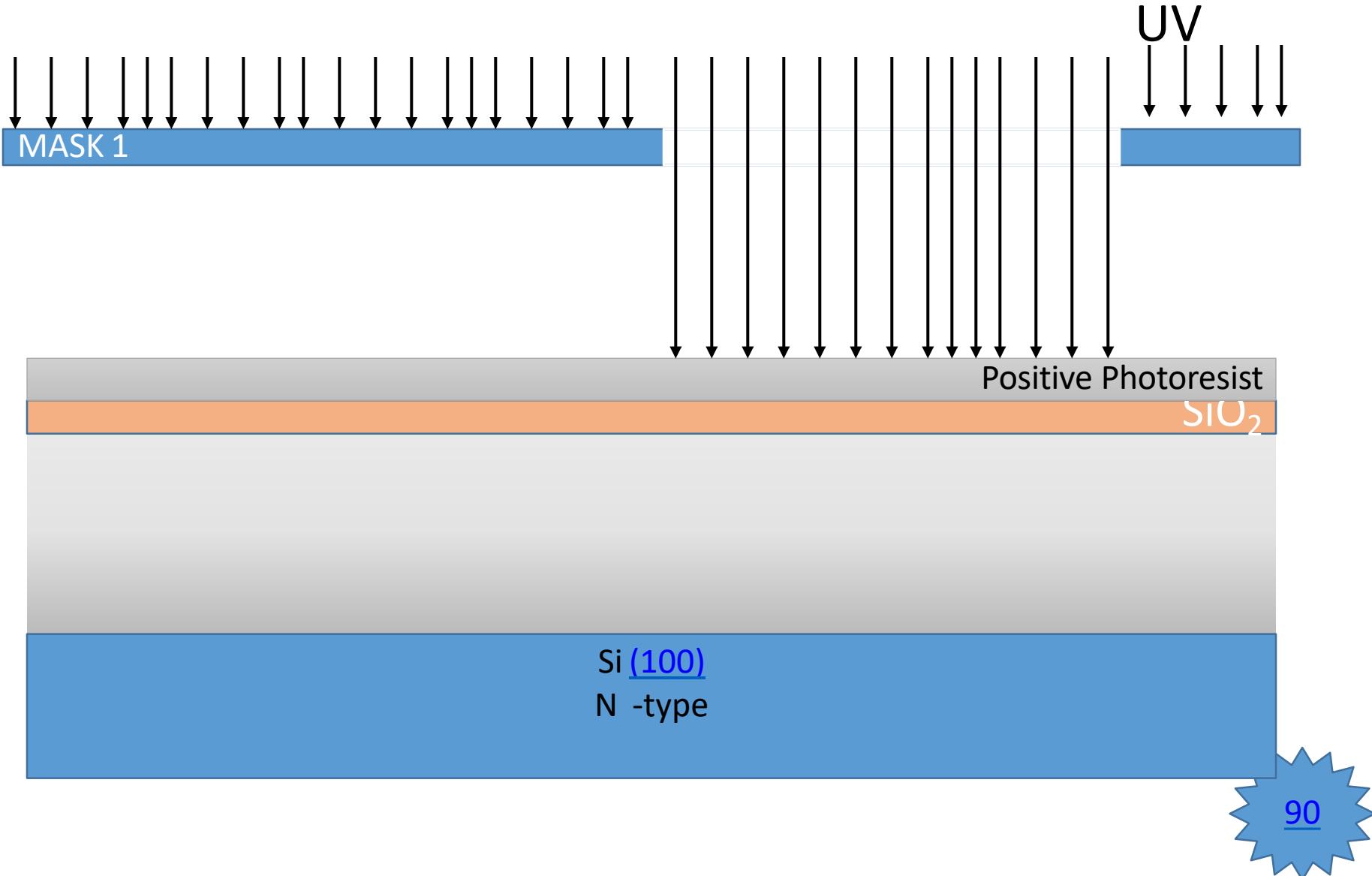
Epitaxial Layer Deposition



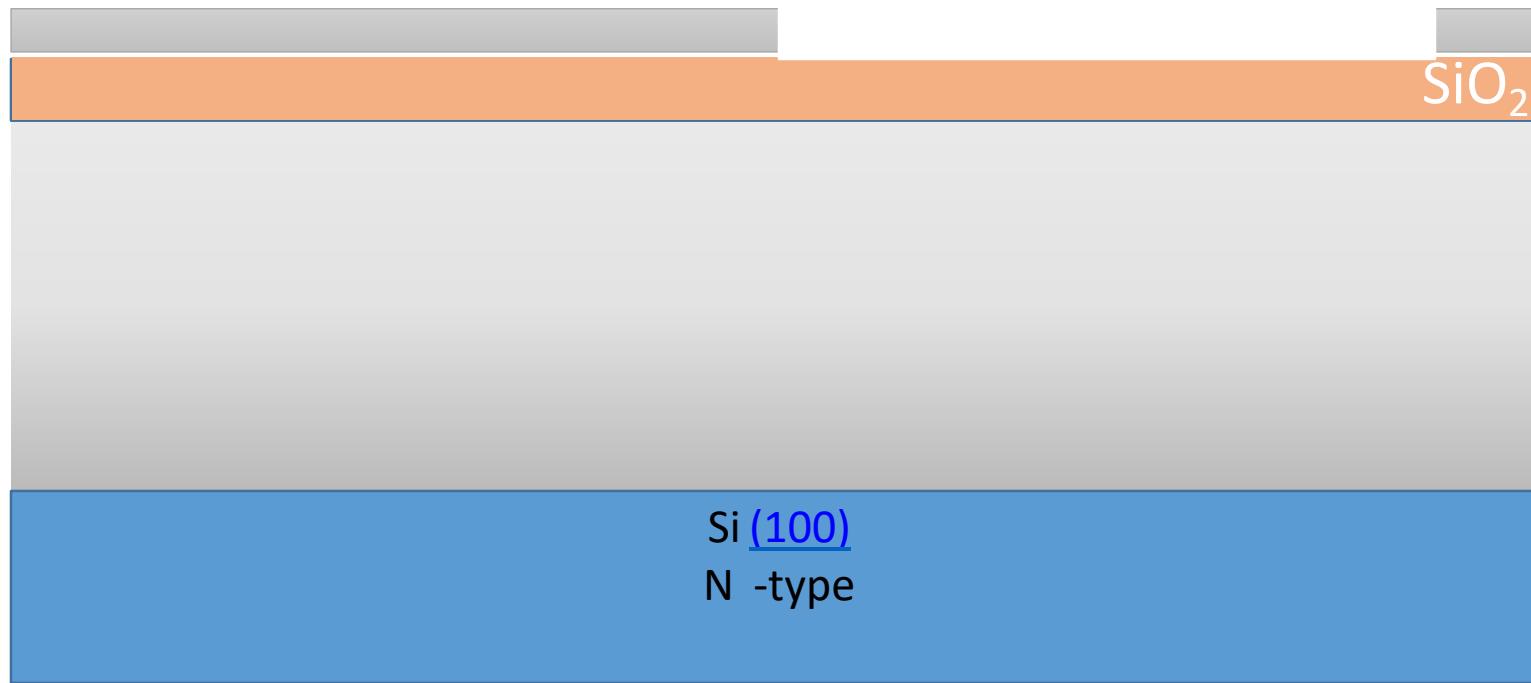
3. Oxidation(1μm)



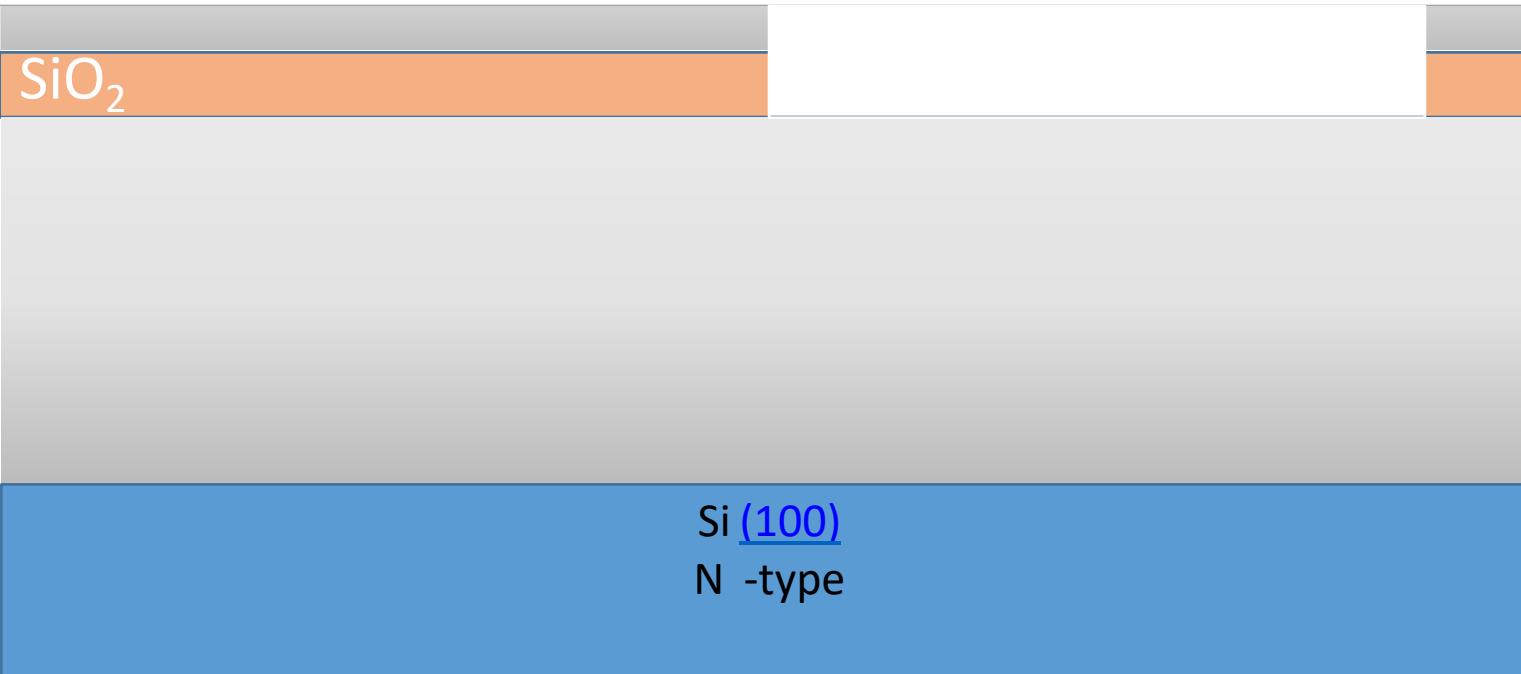
Lithography(MASK1)



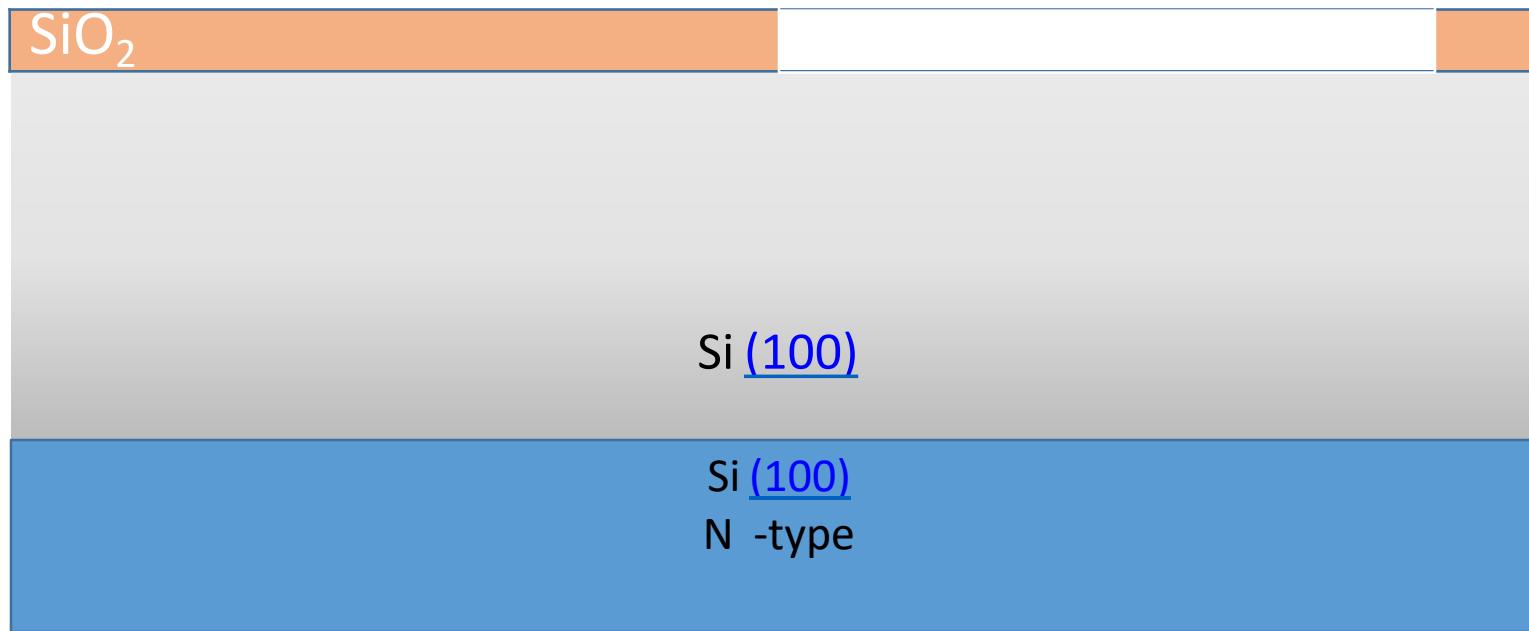
HF(Oxide) Cleaning & PR Etching



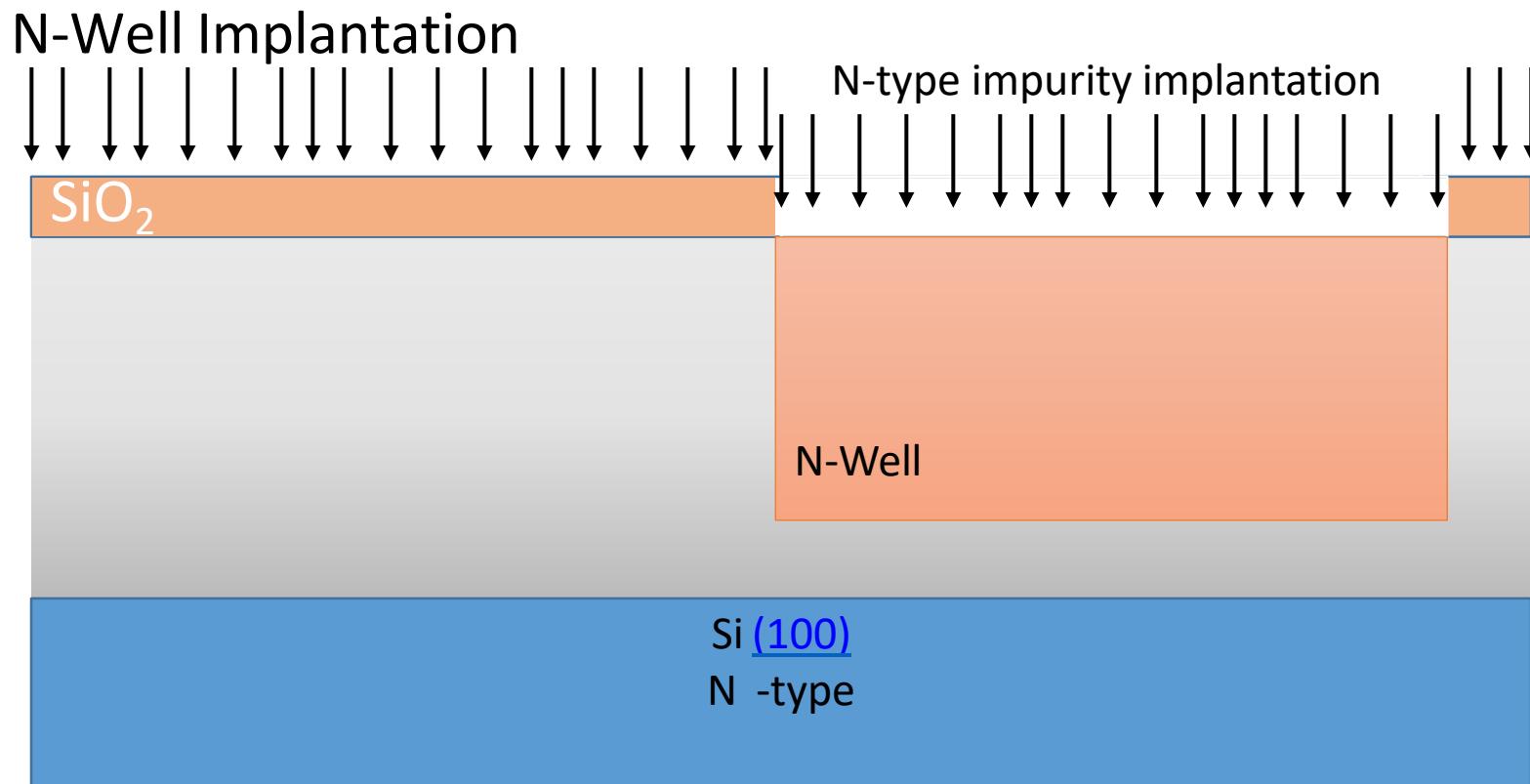
90



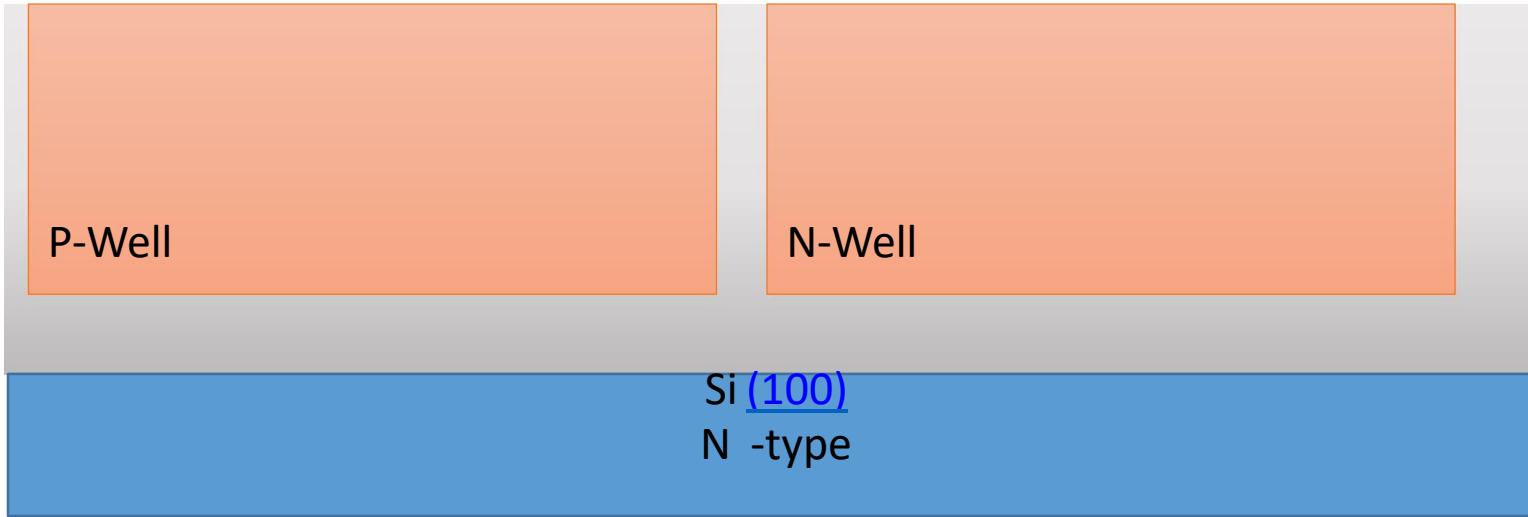
HF(Oxide) Cleaning & PR Etching



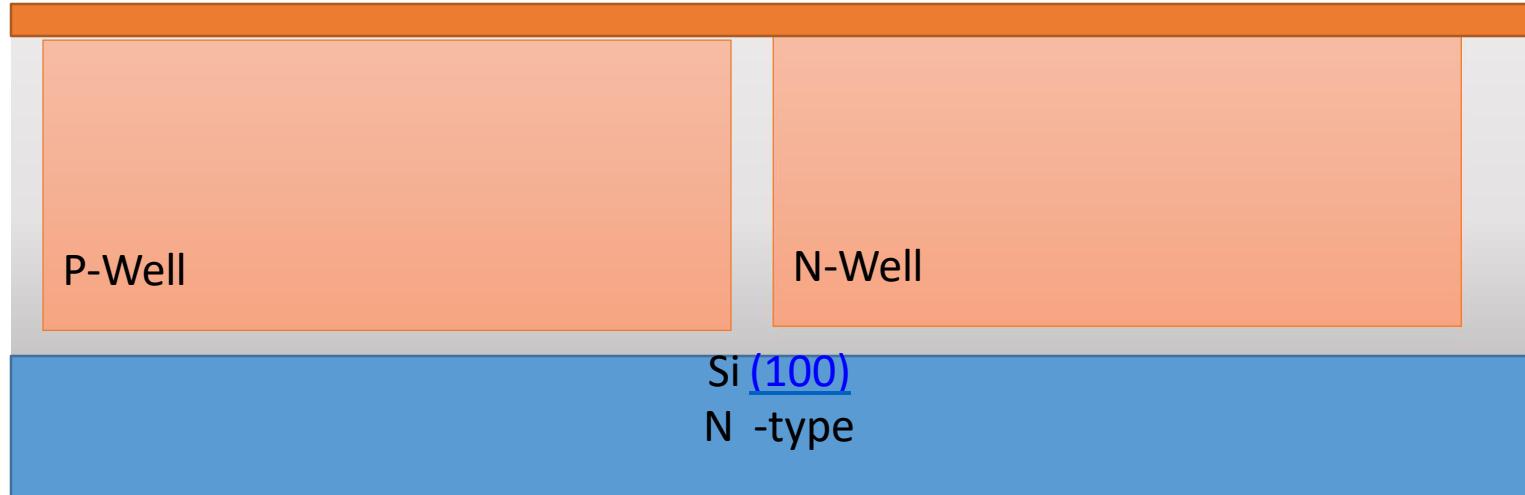
90



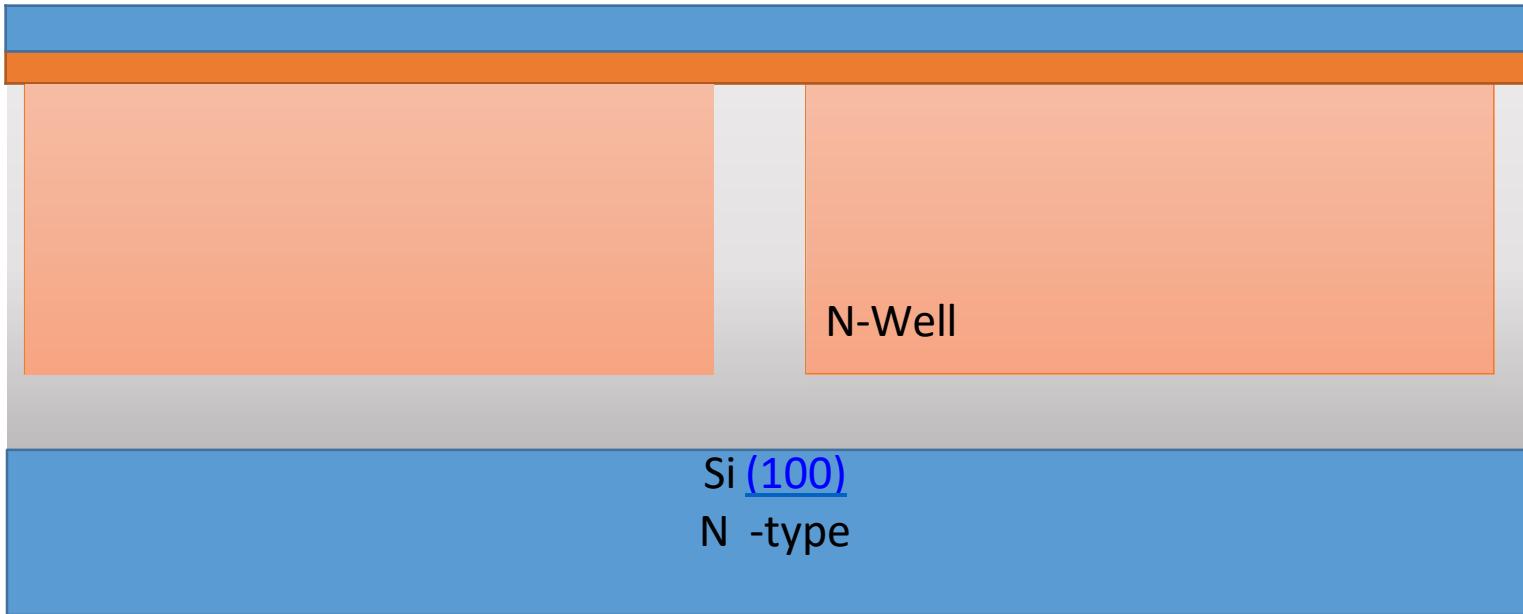
Oxide Cleaning



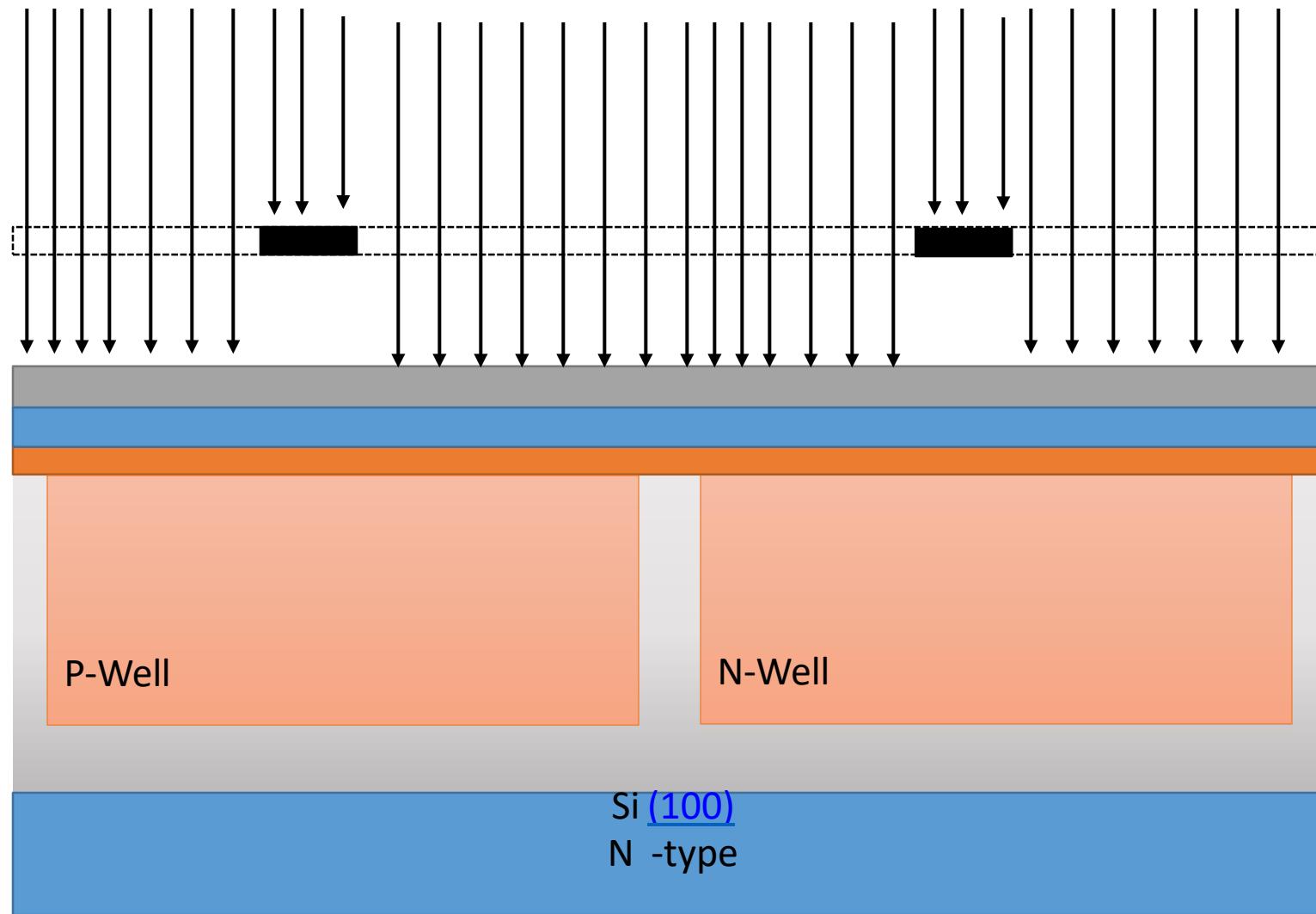
Gate Oxidation

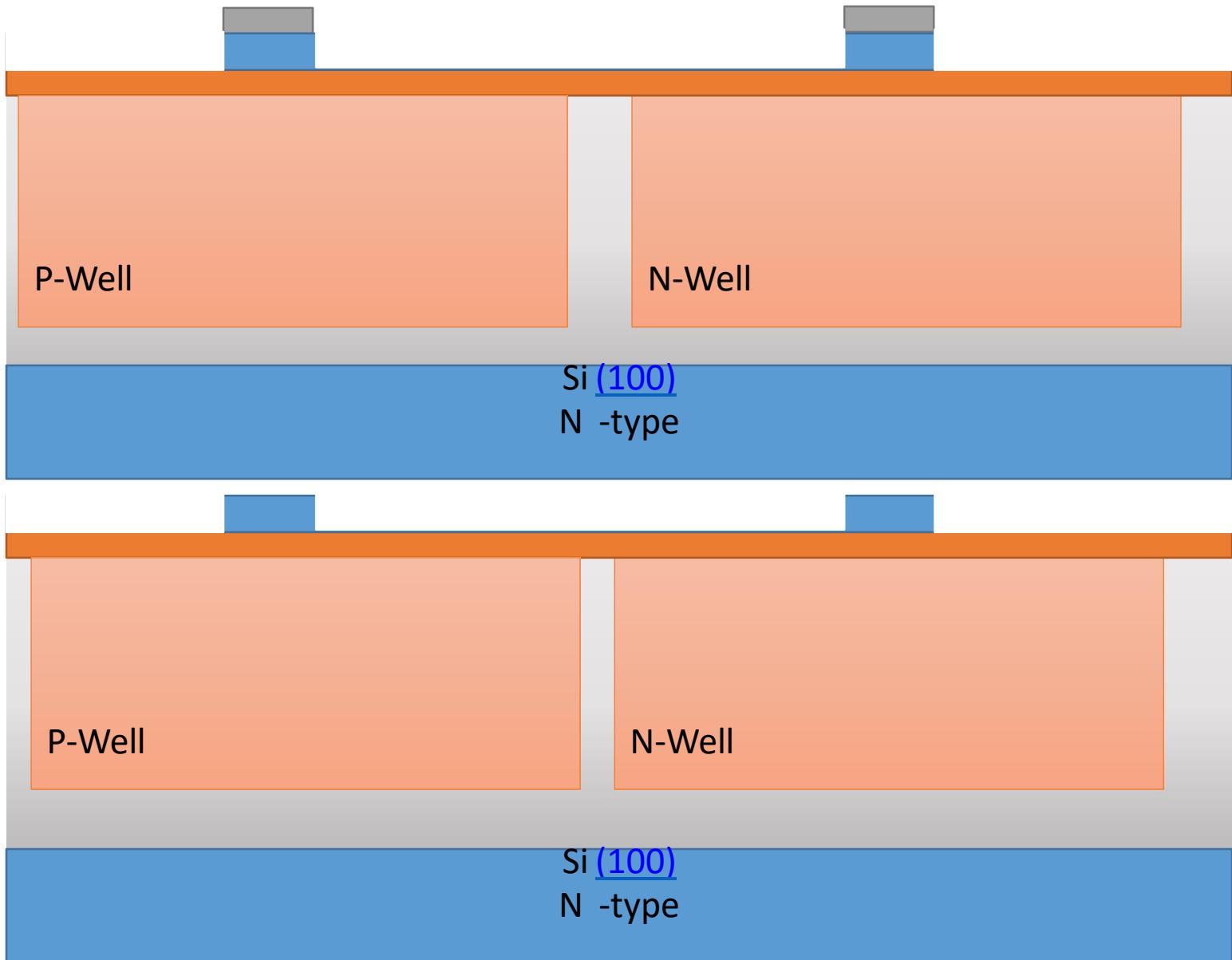


Poly-Silicon Deposition

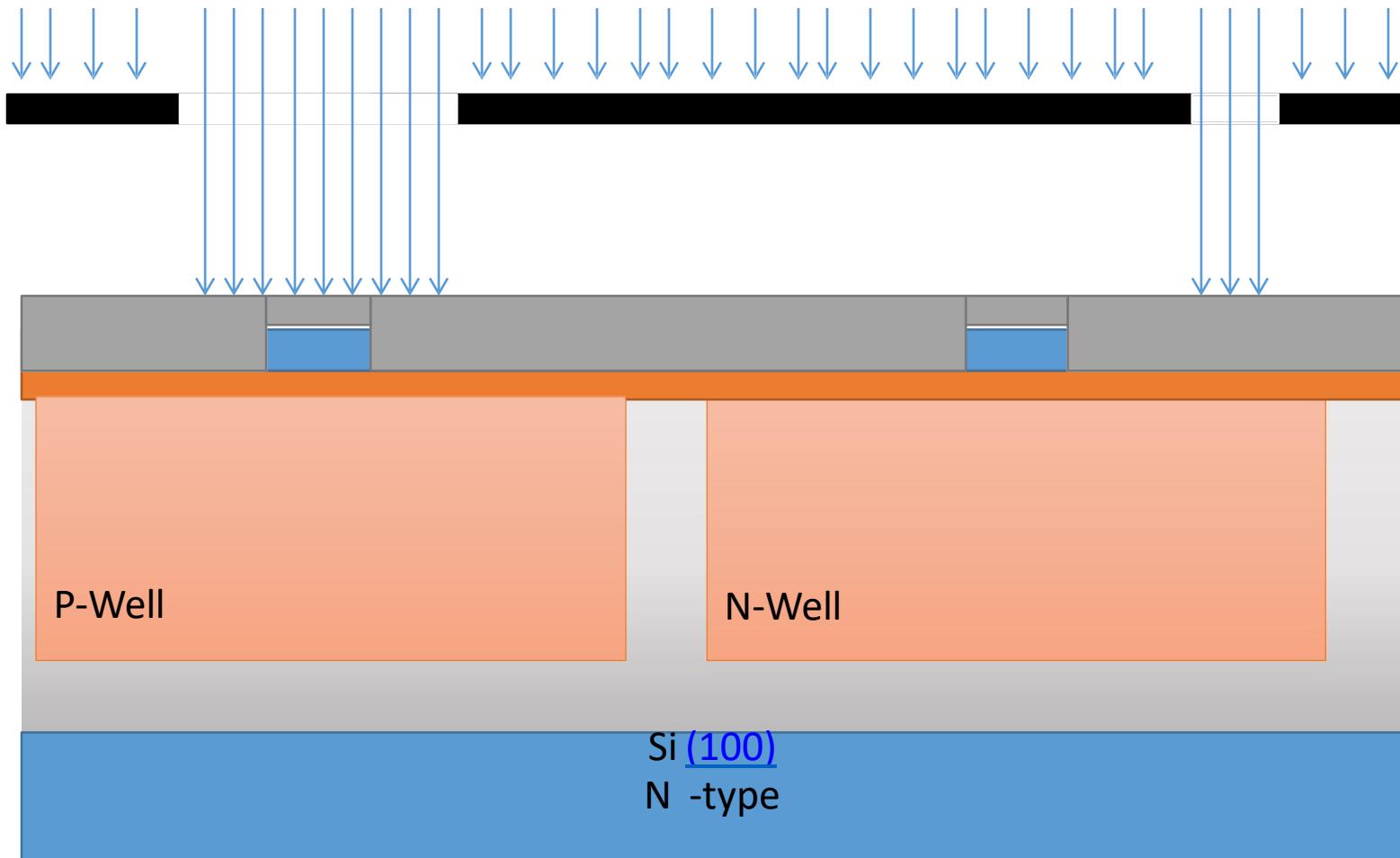


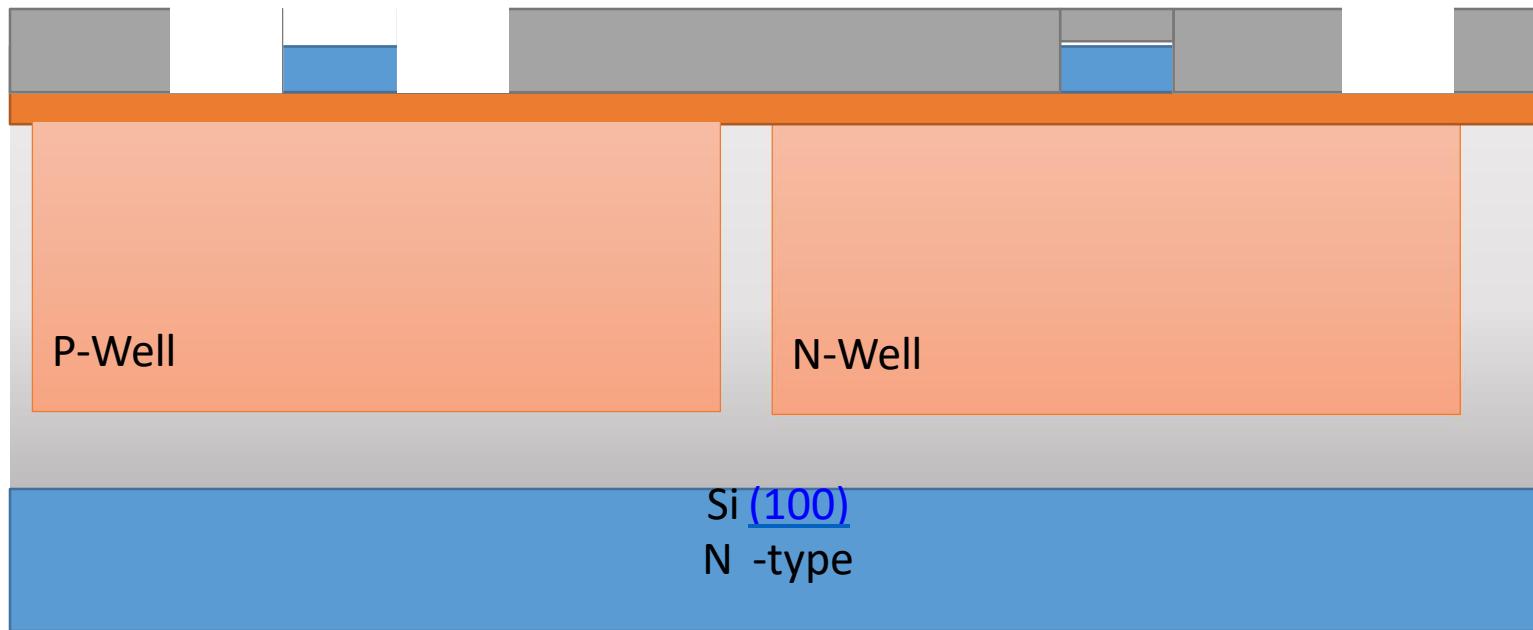
Lithography(MASK2)





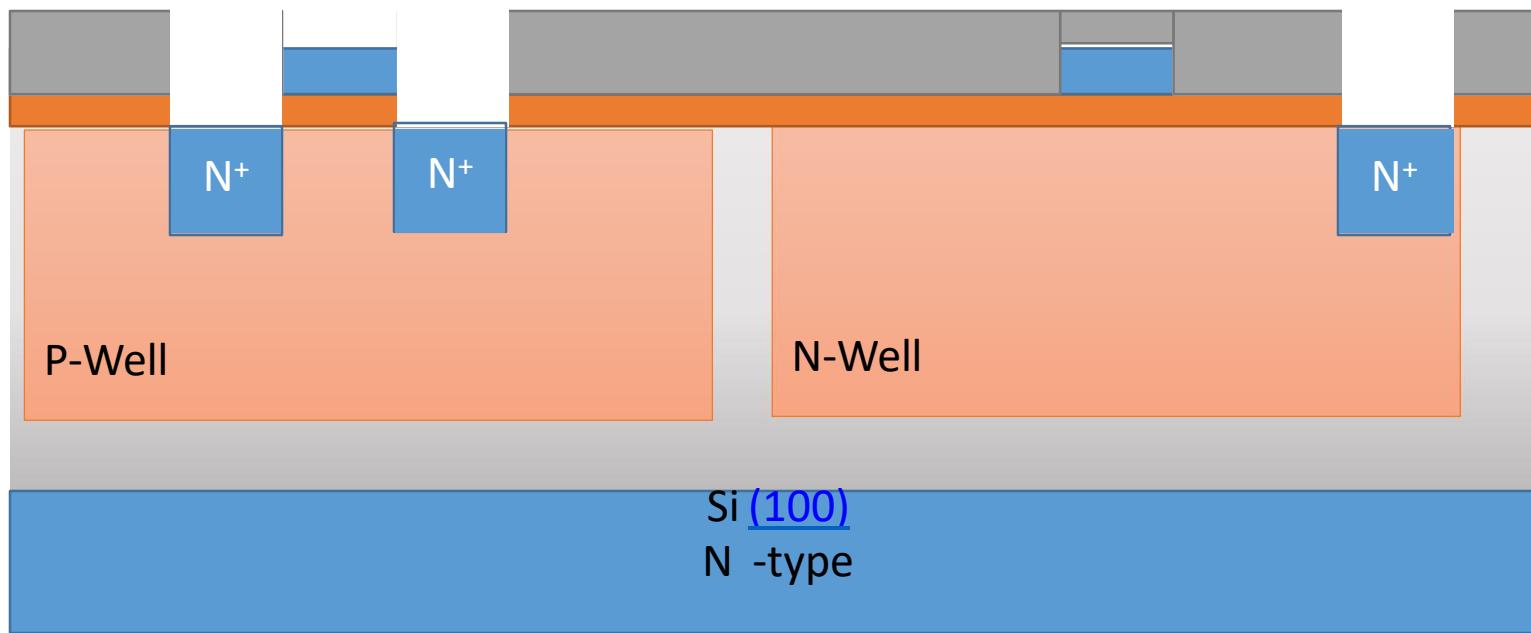
Lithography(MASK3)

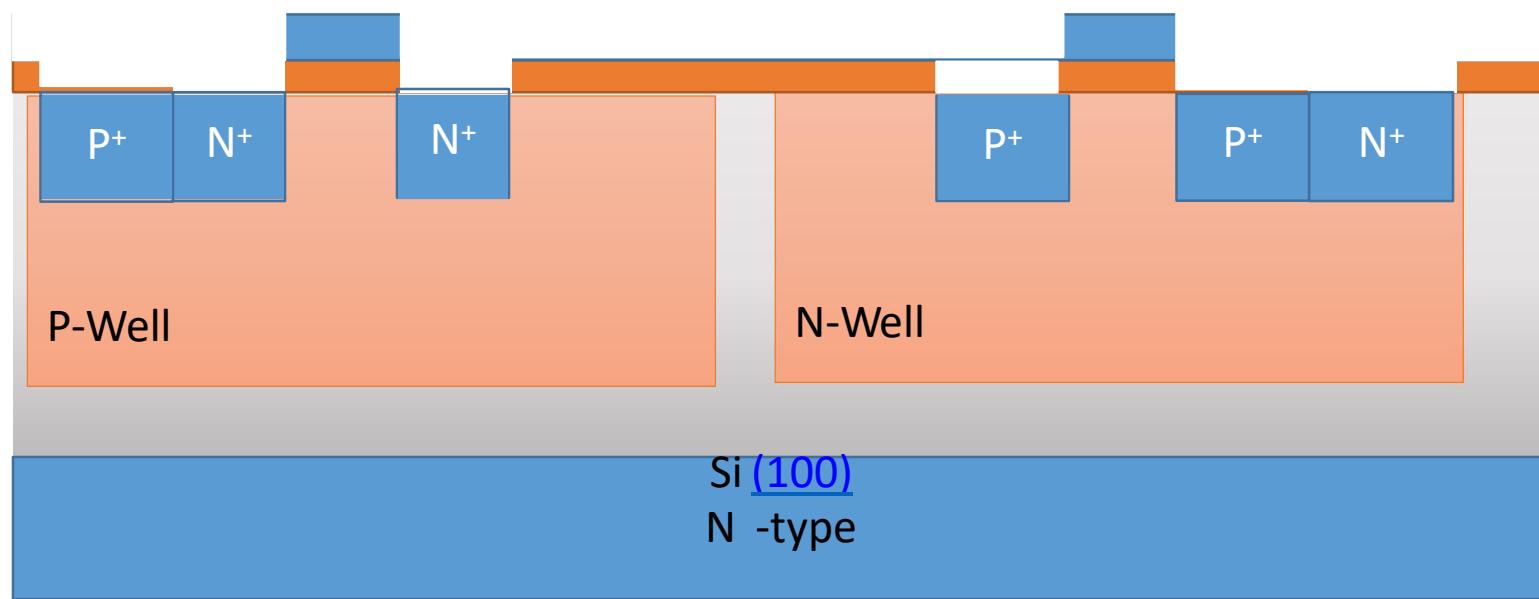




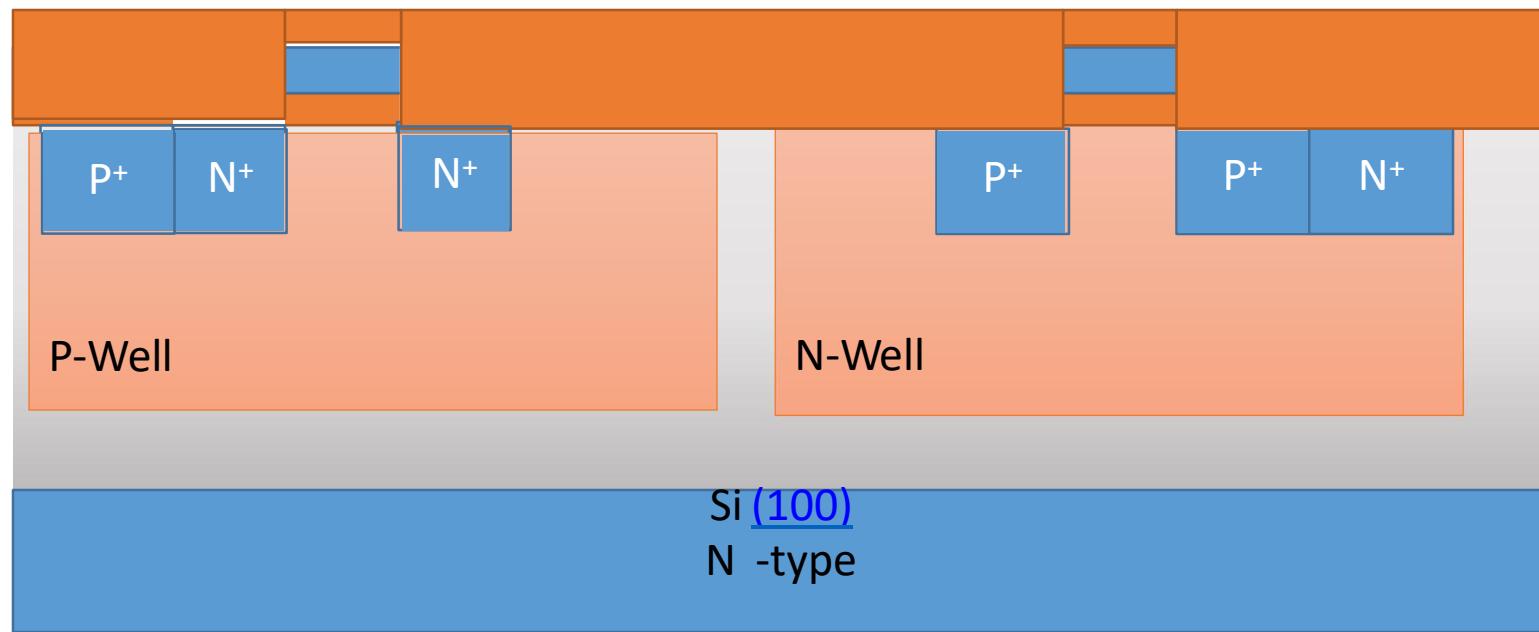
90

Ion Implantation

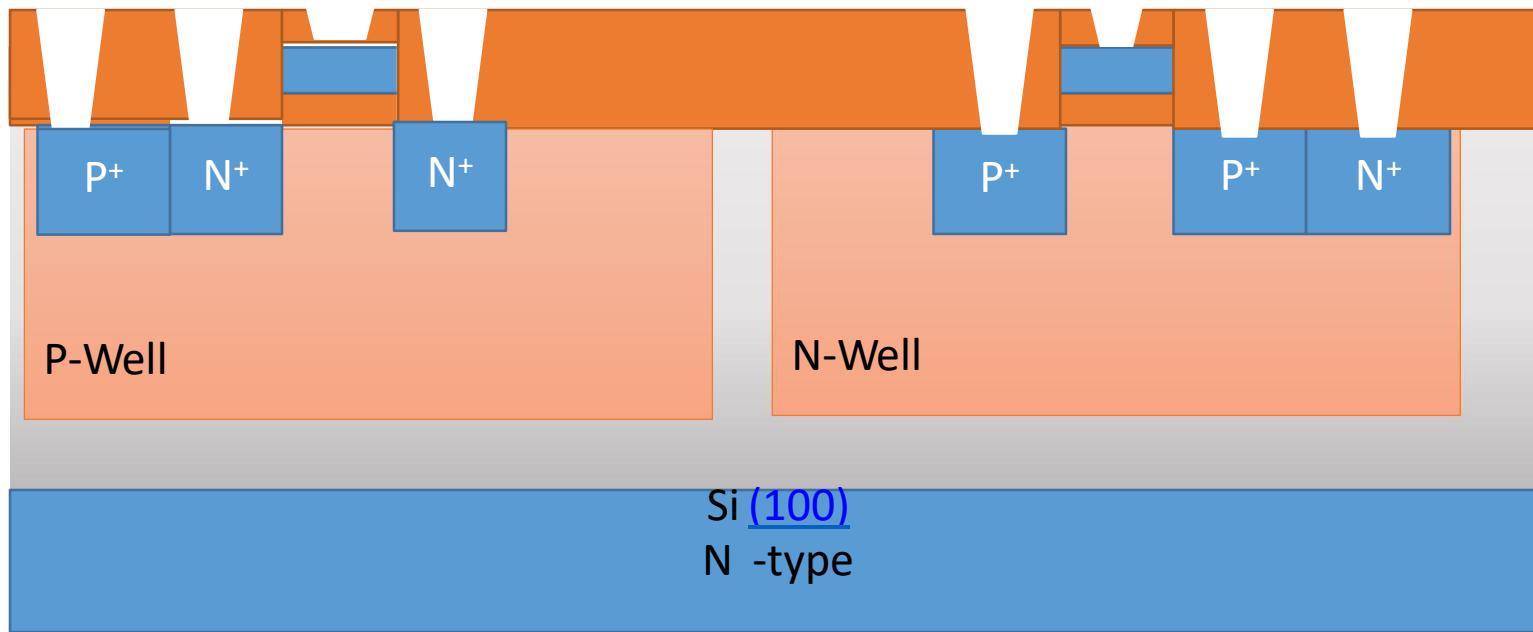




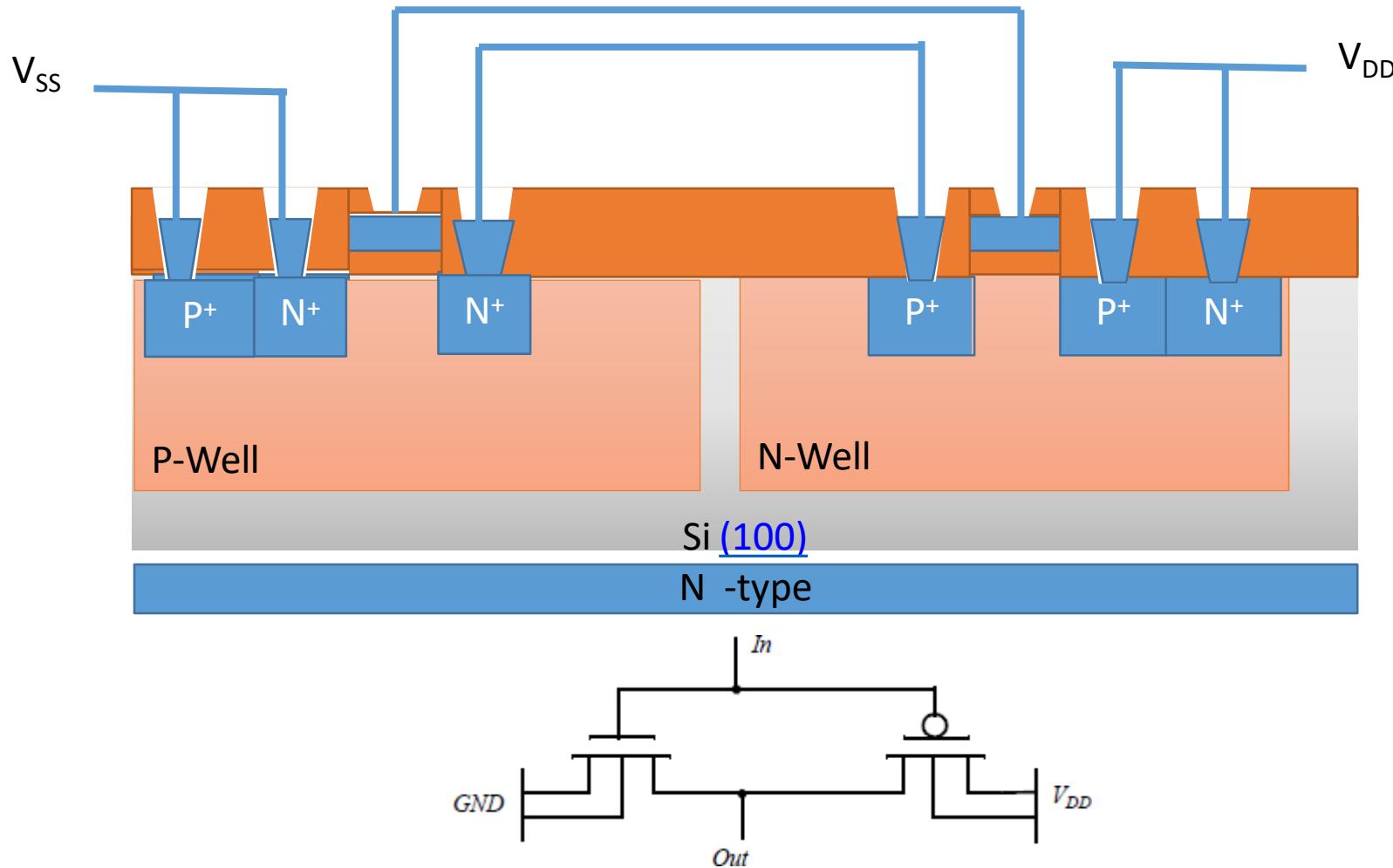
Thick Oxide Deposition



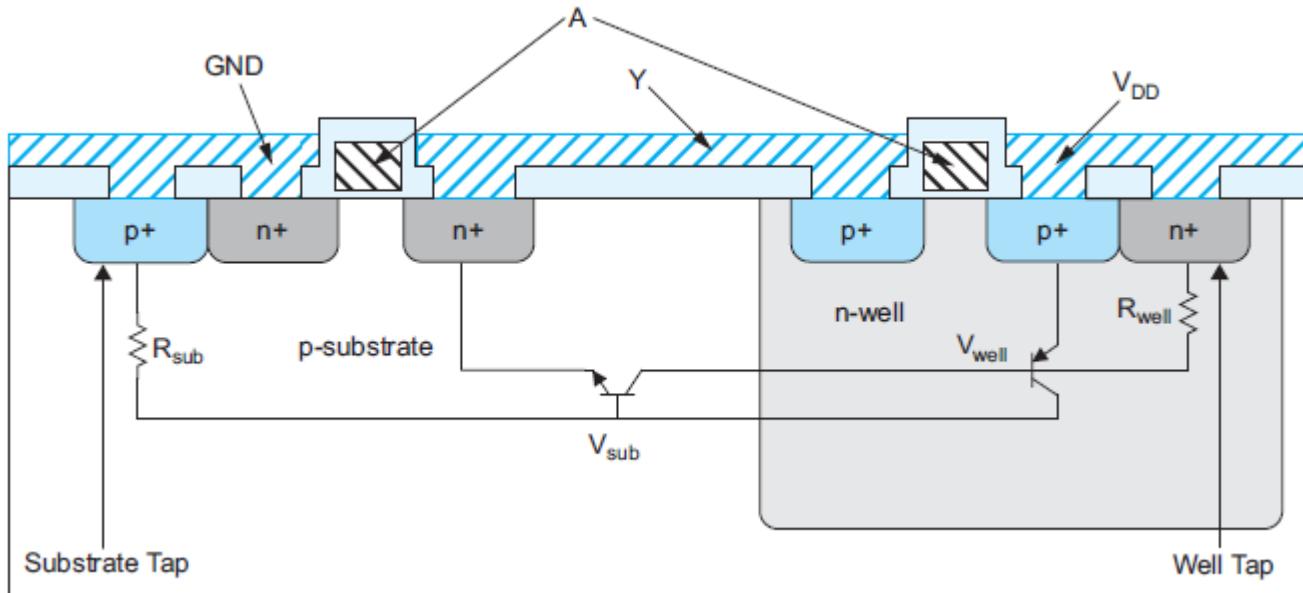
Lithography and Oxide Patterning



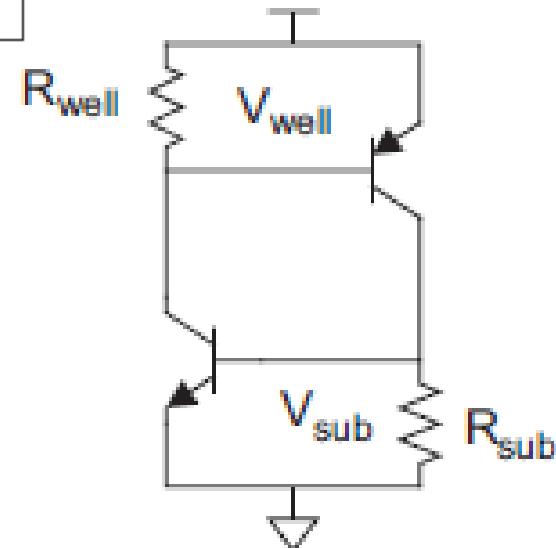
Metallization and Patterning



Latch-Up



Tendency of CMOS chips to develop low-resistance paths between V_{DD} and V_{SS} Called Latch-up

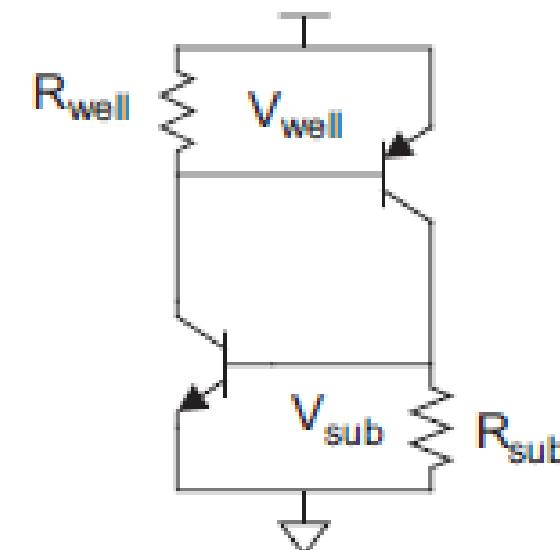
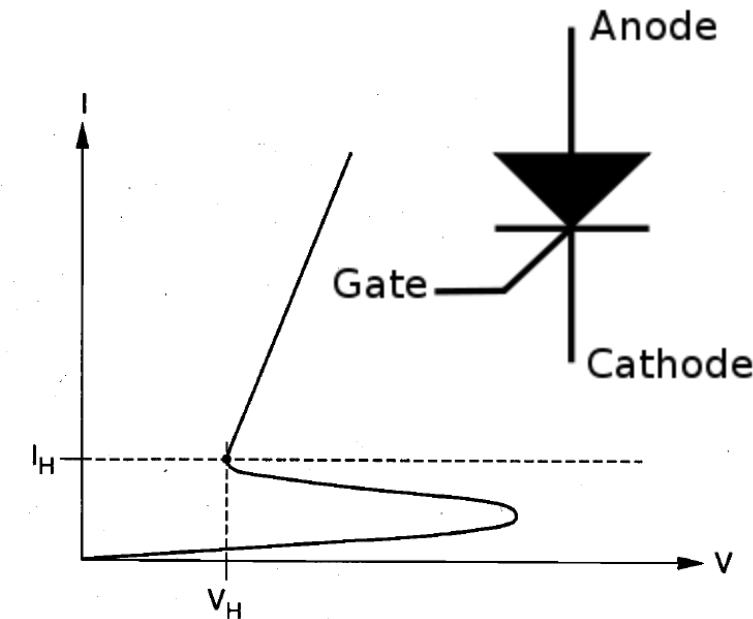


Latch-Up

These BJTs form a silicon-controlled rectifier (SCR) with positive feedback and virtually short circuit the power rail to-ground, thus causing excessive current flows and even permanent device damage.

PNP transistor whose base is formed by the n-well with its base-to-collector current gain (β_1) as high as several hundreds.

NPN transistor with its base formed by the p-type substrate. The base-to-collector current gain β_2 of this lateral transistor may range from a few tenths to tens



Latch-up

R_{well} represents the parasitic resistance in the n-well structure with its value ranging from $1\text{ k}\Omega$ to $20\text{ k}\Omega$.

R_{sub} can be as high as several hundred ohms.

Unless the SCR is triggered by an external disturbance, the collector currents of both transistors consist of the reverse leakage currents of the collector-base junctions and therefore, their current gains are very low.

If the collector current of one of the transistors is temporarily increased by an external disturbance, however, the resulting feedback loop causes this current perturbation to be multiplied by $(\beta_1 \beta_2)$. This event is called the *triggering* of the SCR.

Latch-up

Once triggered, each transistor drives the other transistor with positive feedback, eventually creating and sustaining a low-impedance path between the power and the ground rails, resulting in latch-up. It can be seen that if the condition

$$\beta_1 \cdot \beta_2 \geq 1$$

$$\frac{\alpha_1}{1-\alpha_1} \cdot \frac{\alpha_2}{1-\alpha_2} \geq 1 \Rightarrow \alpha_1 + \alpha_2 \geq 1$$

Technique to overcome Latch-up

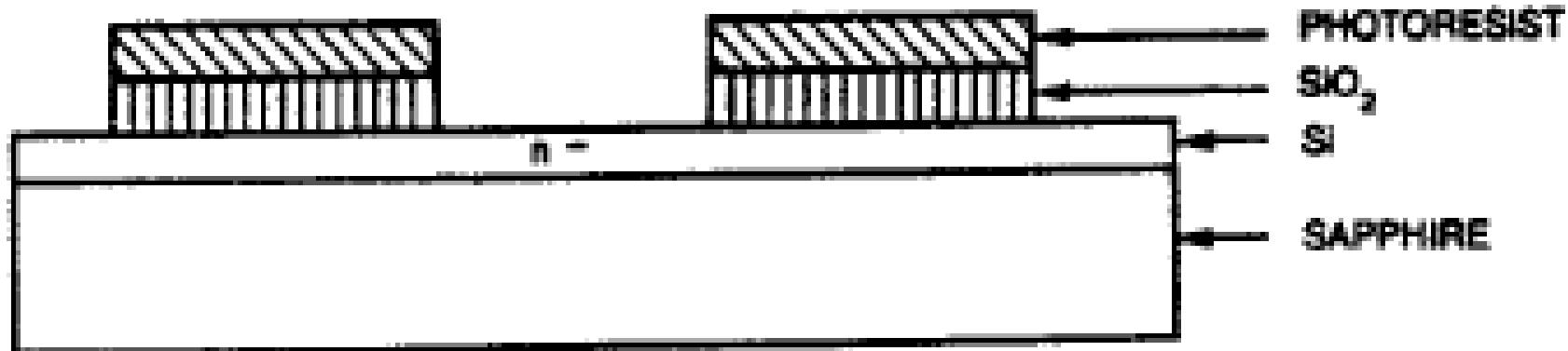
Use p+ guard-band rings connected to ground around nMOS transistors and n+ guard rings connected to VDD around pMOS transistors to reduce R and R_{SUB} and to capture injected minority carriers before they reach the base of the parasitic BJTs.

Place substrate and well contacts as close as possible to the source connections of MOS transistors to reduce the values of R_{well} and R_{sub} .

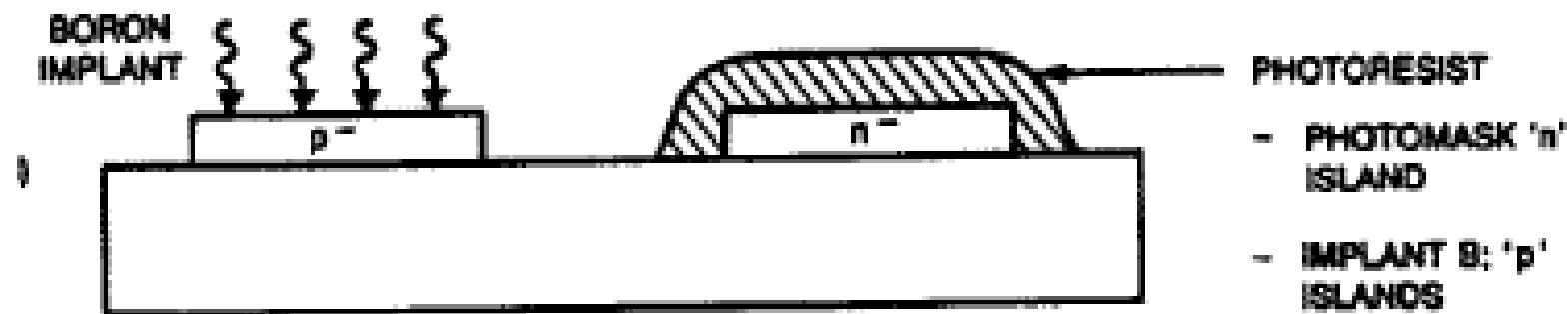
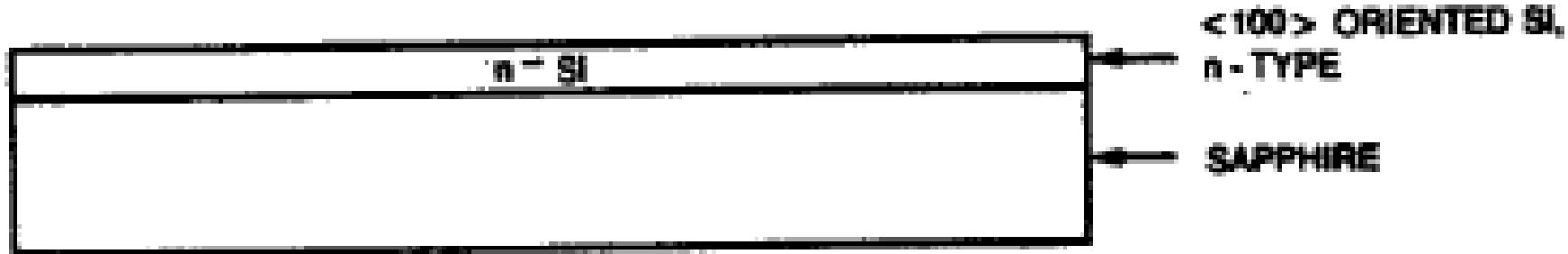
SOI Devices

SOI fabrication

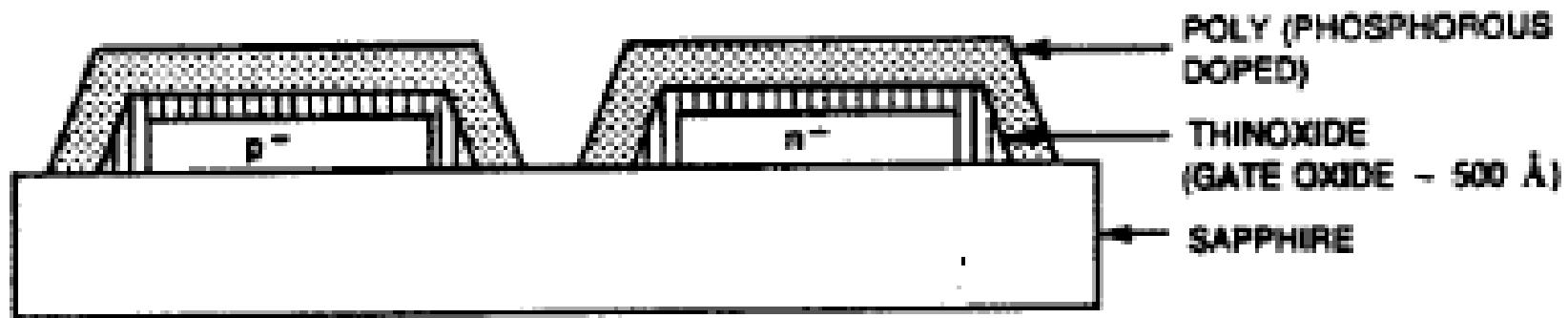
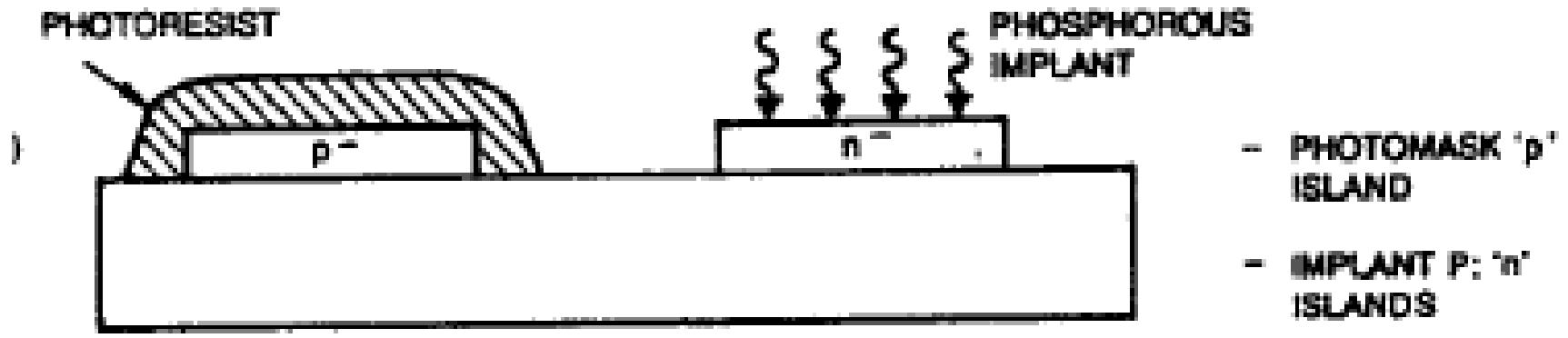
Refer: Principles of CMOS VLSI Design A system perspective 2nd Ed. By Neil H. E. Weste and Kamran Eshraghian Page no. 125-129



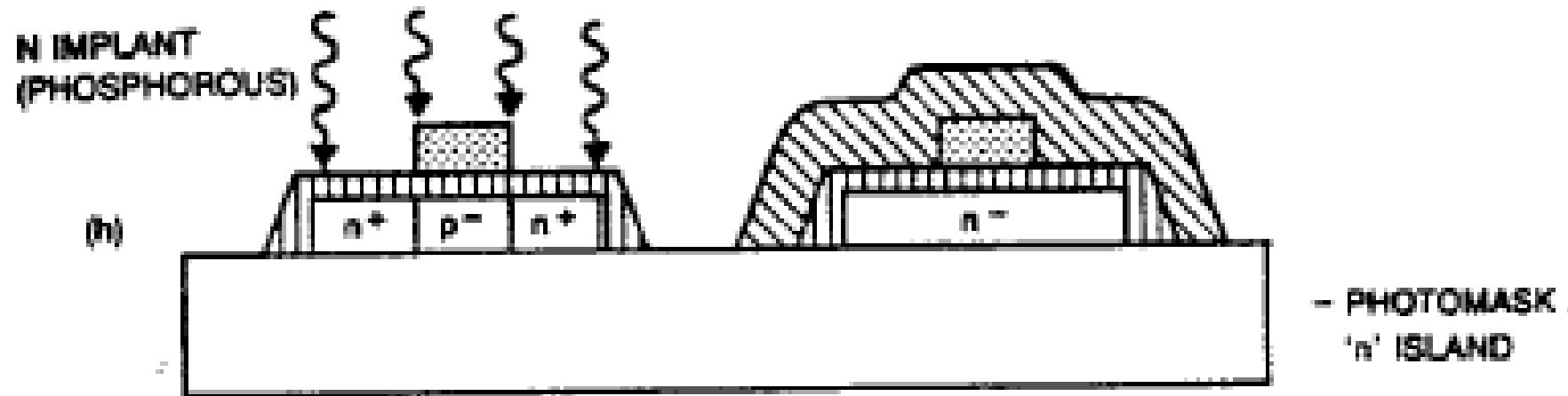
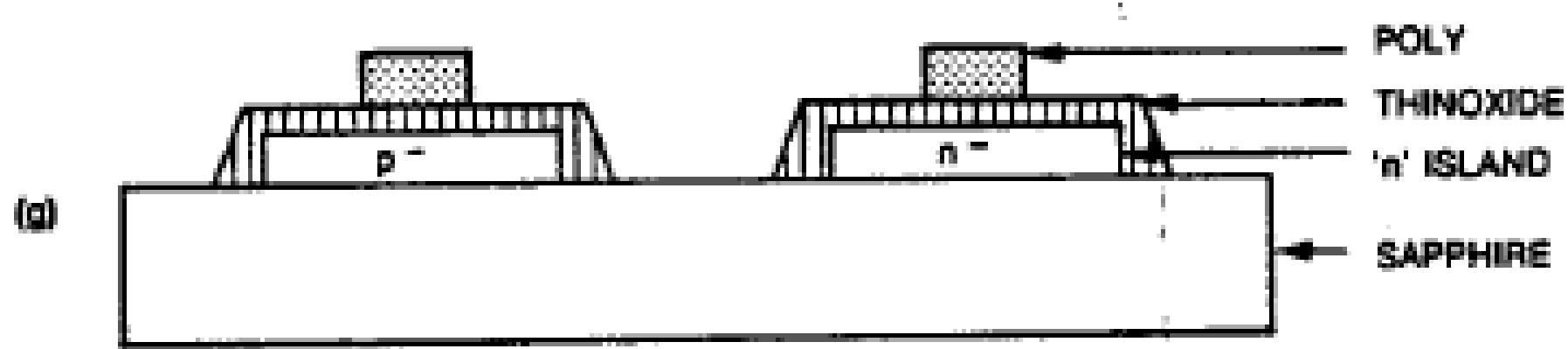
SOI fabrication



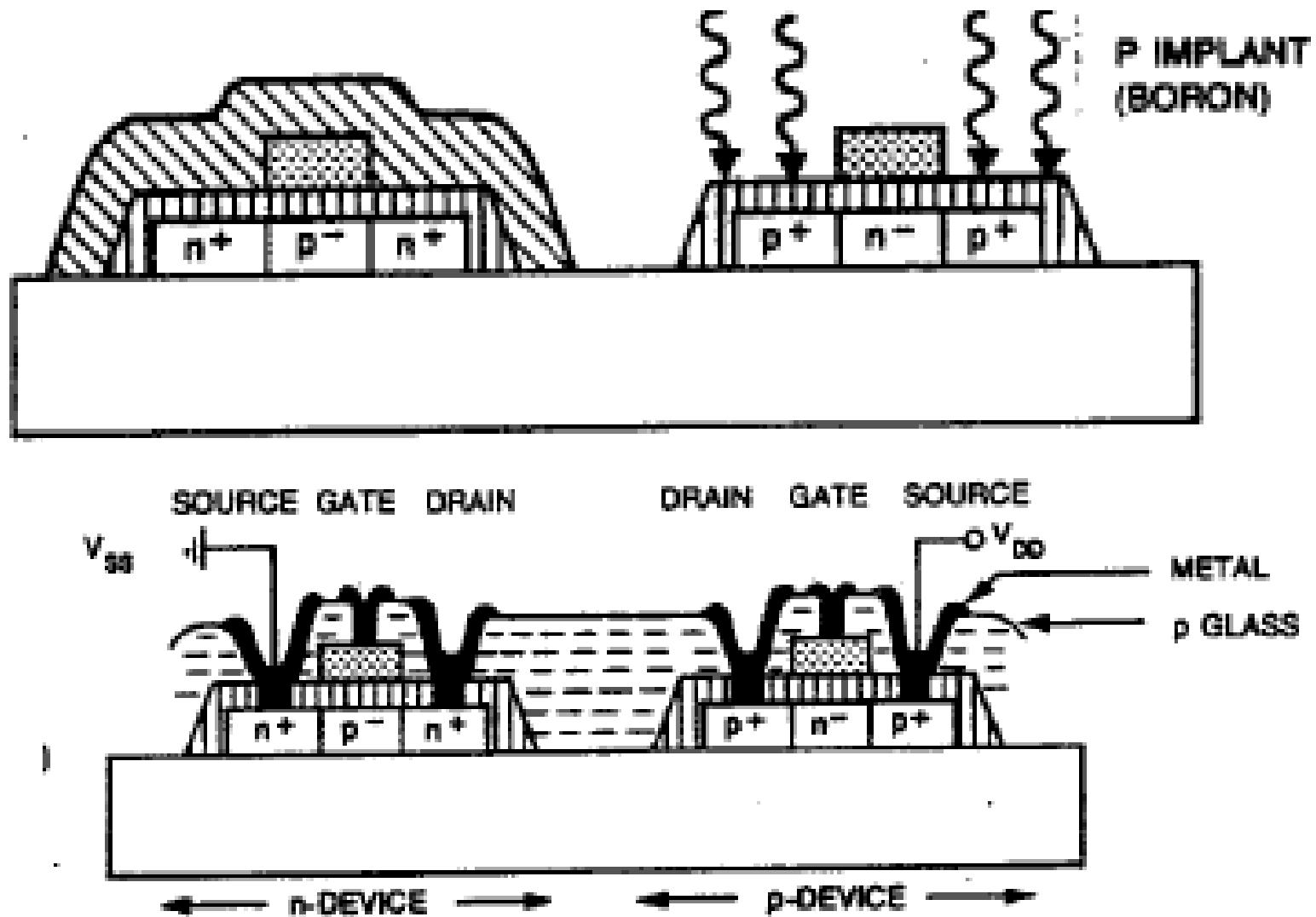
SOI fabrication



SOI fabrication

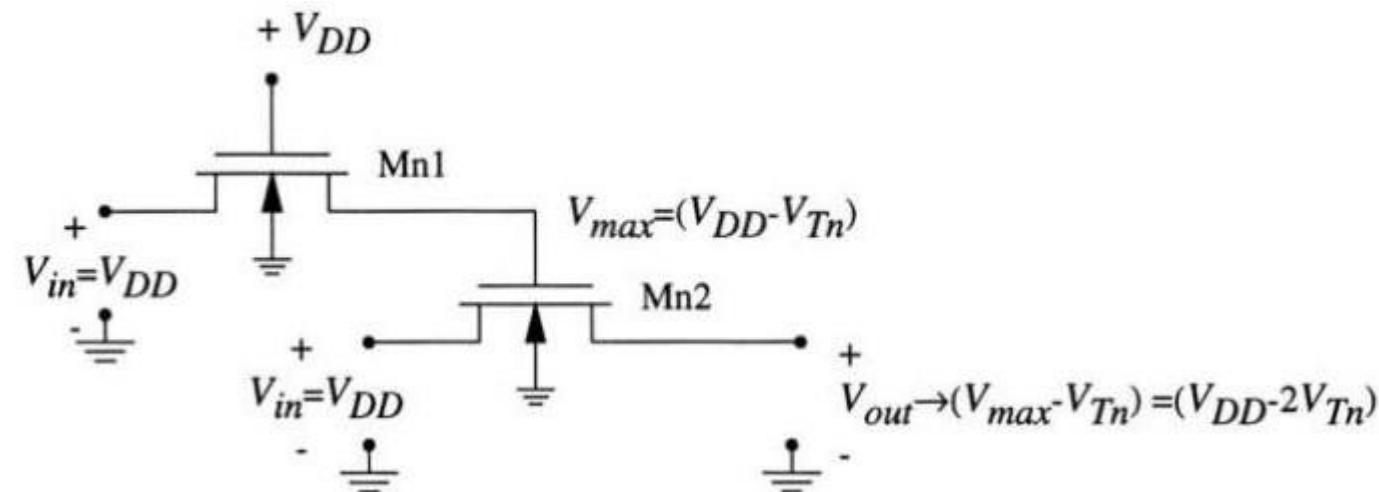
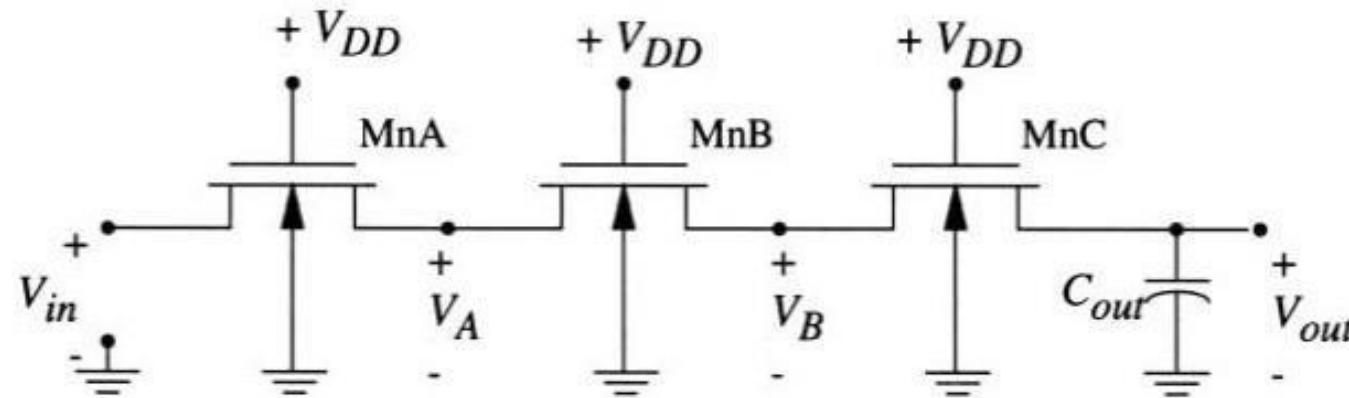


SOI fabrication



Pass Transistor and Transmission Gate

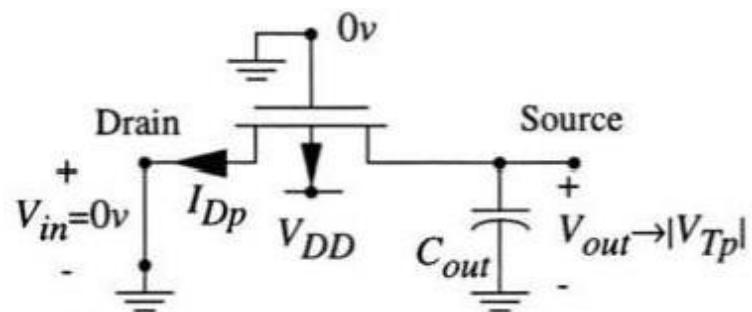
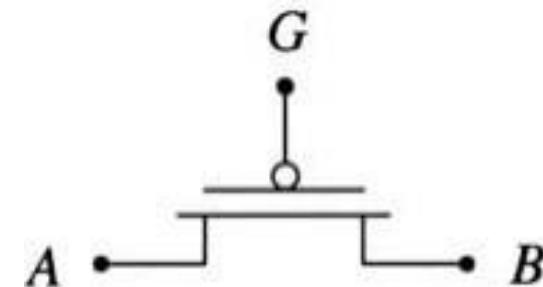
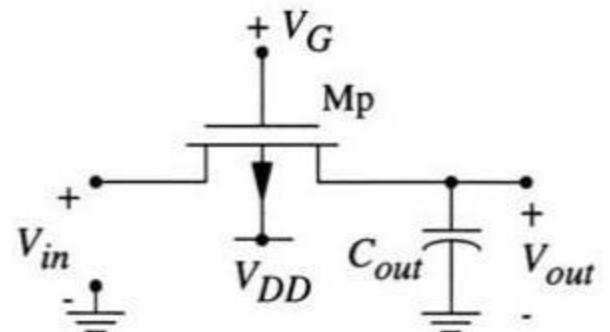
Pass Transistors (NMOS)



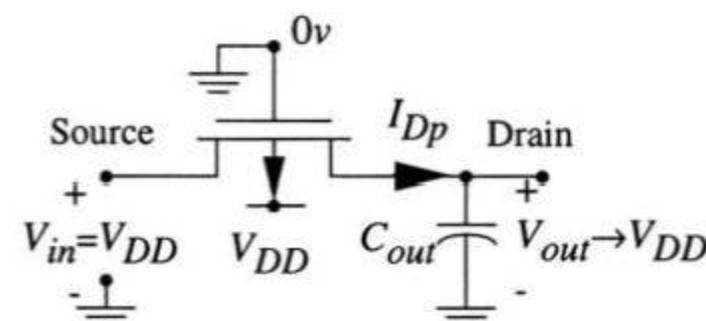
NMOS

- NMOS Pass transistor passes strong logic 0.
- NMOS Pass transistor passes weak logic 1.

Pass Transistors (PMOS)

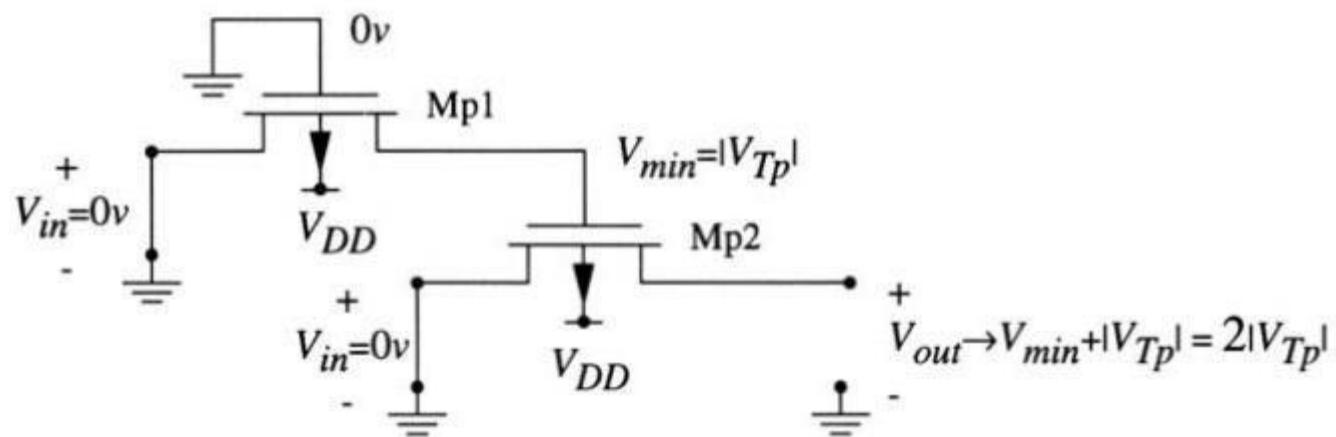
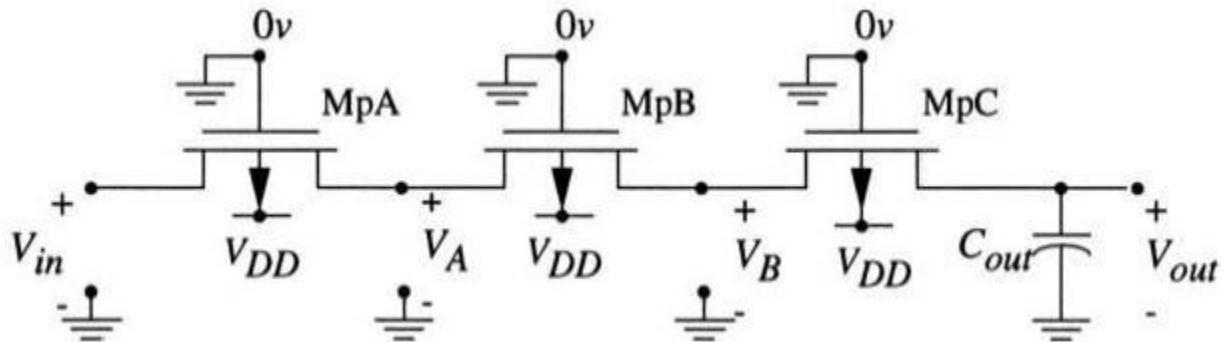


(a) Logic 0 input



(b) Logic 1 input

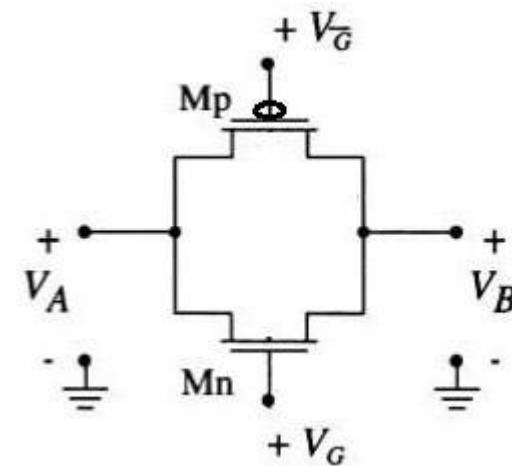
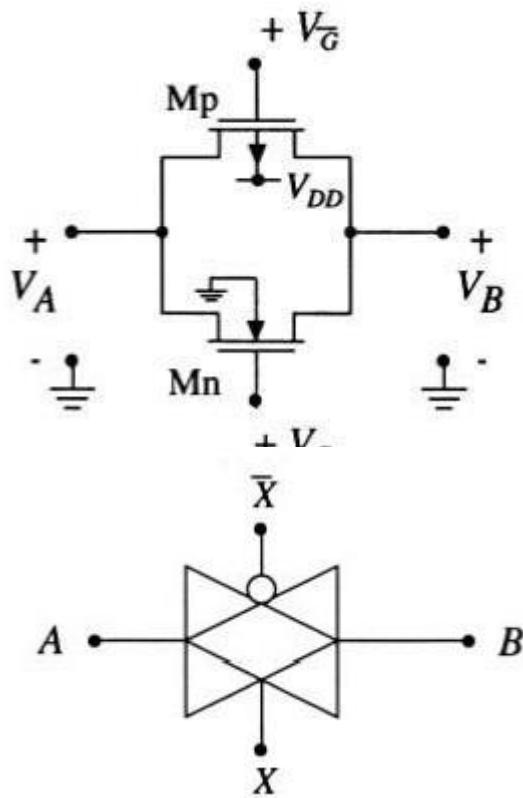
Pass Transistors (PMOS)



PMOS

- PMOS Pass transistor passes weak logic 0.
- PMOS Pass transistor passes strong logic 1.

Transmission Gate

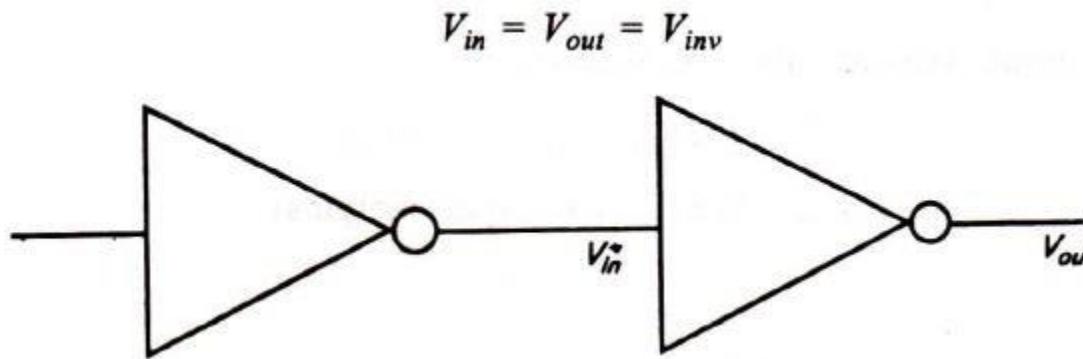


X	A	B
1	0	0
1	1	1
0	0	?
0	1	?

PMOS Pass transistor passes strong logic 1.

NMOS Pass transistor passes strong logic 0.

Determination of pull-up to pull-down ratio ($Z_p.U / Z_p.D.$)
For An Nmos Inverter Driven By Another Nmos Inverter



$$V_{inv} = 0.5V_{DD}$$

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

Note → $K = \frac{\mu \epsilon_{ins} \epsilon_0}{t_{ox}}$

In the depletion mode

$$I_{ds} = K \frac{W_{p.u.}}{L_{p.u.}} \frac{(-V_{td})^2}{2} \text{ since } V_{gs} = 0$$

and in the enhancement mode

$$I_{ds} = K \frac{W_{p.d.}}{L_{p.d.}} \frac{(V_{inv} - V_t)^2}{2} \text{ since } V_{gs} = V_{inv}$$

Equating (since currents are the same) we have

$$\frac{W_{p.d.}}{L_{p.d.}} (V_{inv} - V_t)^2 = \frac{W_{p.u.}}{L_{p.u.}} (-V_{td})^2$$

$$Z_{p.d.} = \frac{L_{p.d.}}{W_{p.d.}}; Z_{p.u.} = \frac{L_{p.u.}}{W_{p.u.}}$$

$$\frac{1}{Z_{p.d.}} (V_{inv} - V_t)^2 = \frac{1}{Z_{p.u.}} (-V_{td})^2$$

$$V_{inv} = V_t - \frac{V_{td}}{\sqrt{Z_{p.u.}/Z_{p.d.}}}$$

substitute typical values as follows

$$V_t = 0.2V_{DD}; V_{td} = - 0.6V_{DD}$$

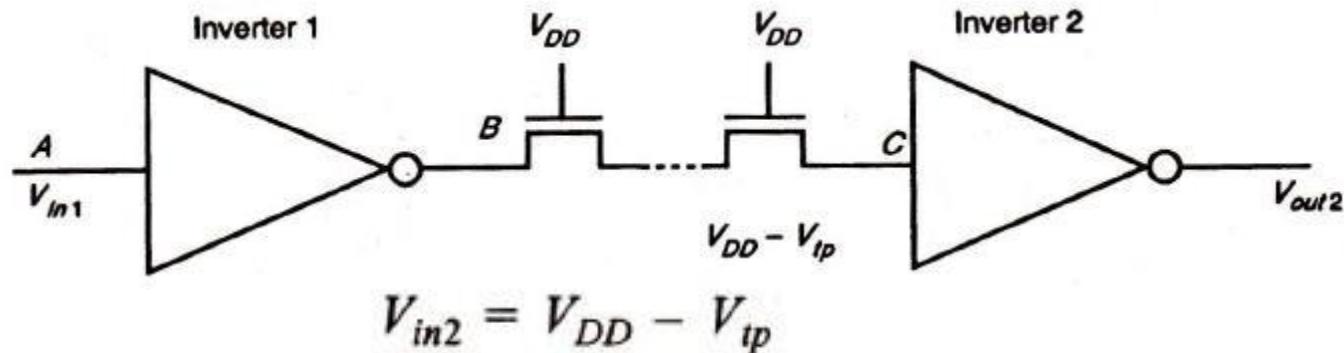
$$V_{inv} = 0.5V_{DD} \text{ (for equal margins)}$$

$$0.5 = 0.2 + \frac{0.6}{\sqrt{Z_{p.u.}/Z_{p.d.}}}$$

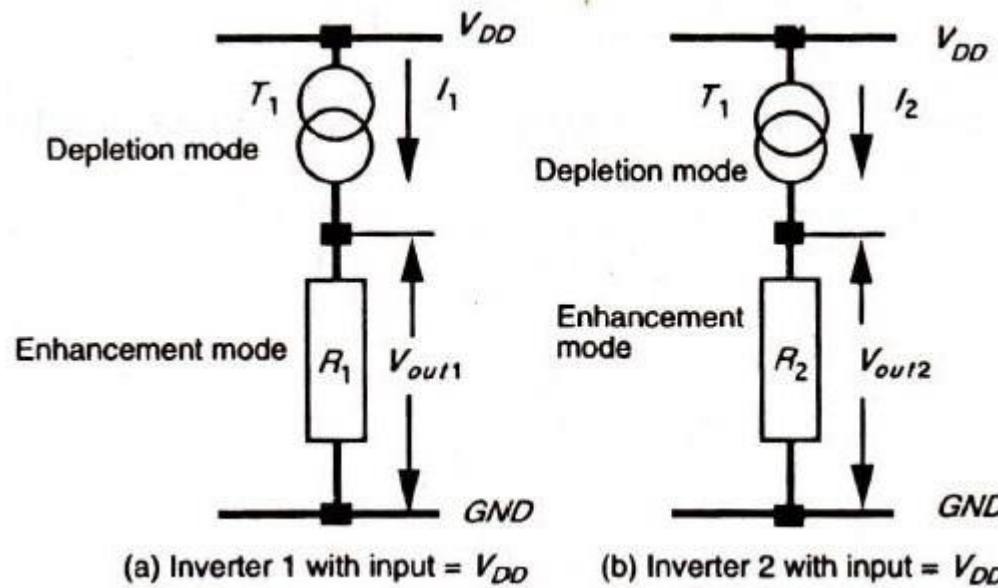
$$\sqrt{Z_{p.u.}/Z_{p.d.}} = 2$$

$$Z_{p.u.}/Z_{p.d.} = 4/1$$

Pull-up to pull-down ratio for an Nmos inverter driven through one or more pass transistors



V_{tp} = threshold voltage for a pass transistor



For the p.d. transistor

$$I_{ds} = K \frac{W_{p.d.1}}{L_{p.d.1}} \left((V_{DD} - V_t) V_{ds1} - \frac{V_{ds1}^2}{2} \right)$$

$$R_1 = \frac{V_{ds1}}{I_{ds}} = \frac{1}{K} \frac{L_{p.d.1}}{W_{p.d.1}} \left(\frac{1}{V_{DD} - V_t - \frac{V_{ds1}}{2}} \right)$$

$$R_1 = \frac{1}{K} Z_{p.d.1} \left(\frac{1}{V_{DD} - V_t} \right)$$

Now, for depletion mode p.u. in saturation with $V_{GS} = 0$

$$I_1 = I_{ds} = K \frac{W_{p.u.1}}{L_{p.u.1}} \frac{(-V_{td})^2}{2}$$

$$\text{Note} \rightarrow K = \frac{\mu \epsilon_{ins} \epsilon_0}{t_{ox}}$$

$$I_1 R_1 = V_{out\ 1}$$

$$V_{out1} = I_1 R_1 = \frac{Z_{p.d.1}}{Z_{p.u.1}} \left(\frac{1}{V_{DD} - V_t} \right) \frac{(V_{td})^2}{2}$$

Consider inverter 2 (Figure 2.10(b)) when input = $VDD - V_{tp}$.

$$R_2 = \frac{1}{K} Z_{p.d.2} \frac{1}{((V_{DD} - V_{tp}) - V_t)}$$

$$I_2 = K \frac{1}{Z_{p.u.2}} \frac{(-V_{td})^2}{2}$$

Taking typical values

$$V_t = 0.2V_{DD}$$

$$V_{tp} = 0.3V_{DD}^*$$

$$\frac{Z_{p.u.2}}{Z_{p.d.2}} = \frac{Z_{p.u.1}}{Z_{p.d.1}} \frac{0.8}{0.5}$$

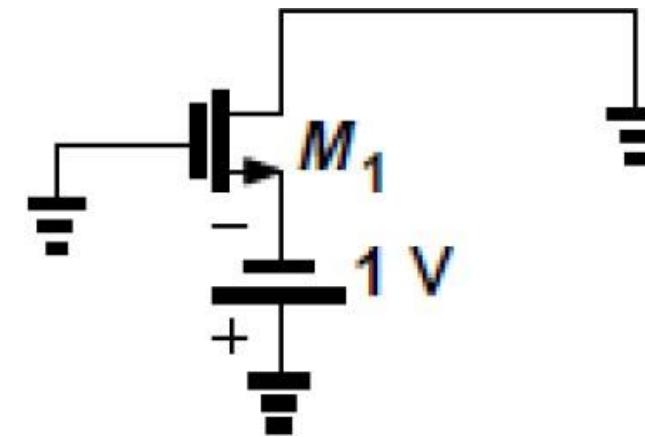
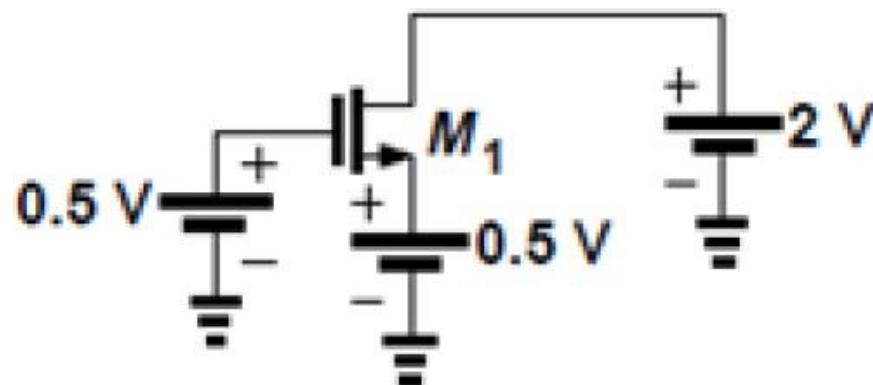
$$\frac{Z_{p.u.2}}{Z_{p.d.2}} \doteq 2 \quad \frac{Z_{p.u.1}}{Z_{p.d.1}} = \frac{8}{1}$$

Problem:

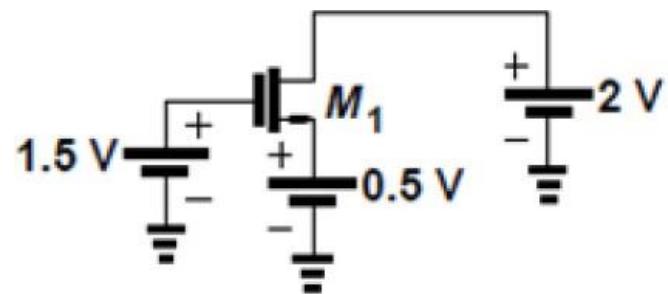
Compare BJT and MOSFET with respect to following parameters:

- (i) Transconductance value
- (ii) Input impedance
- (iii) Noise
- (iv) Current handling capability

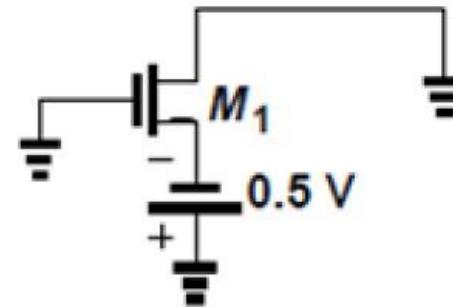
Find the region of operation for NMOS based circuits given in FIG. 3C.
Assume $V_T = 0.4V$



Find the region of operation for the circuit given in **FIG. Q 2C** (i) and (ii).
Threshold voltage $V_{TH} = 0.4$ V.



(i)



(ii)

FIG. Q 2C

BiCMOS Inverter

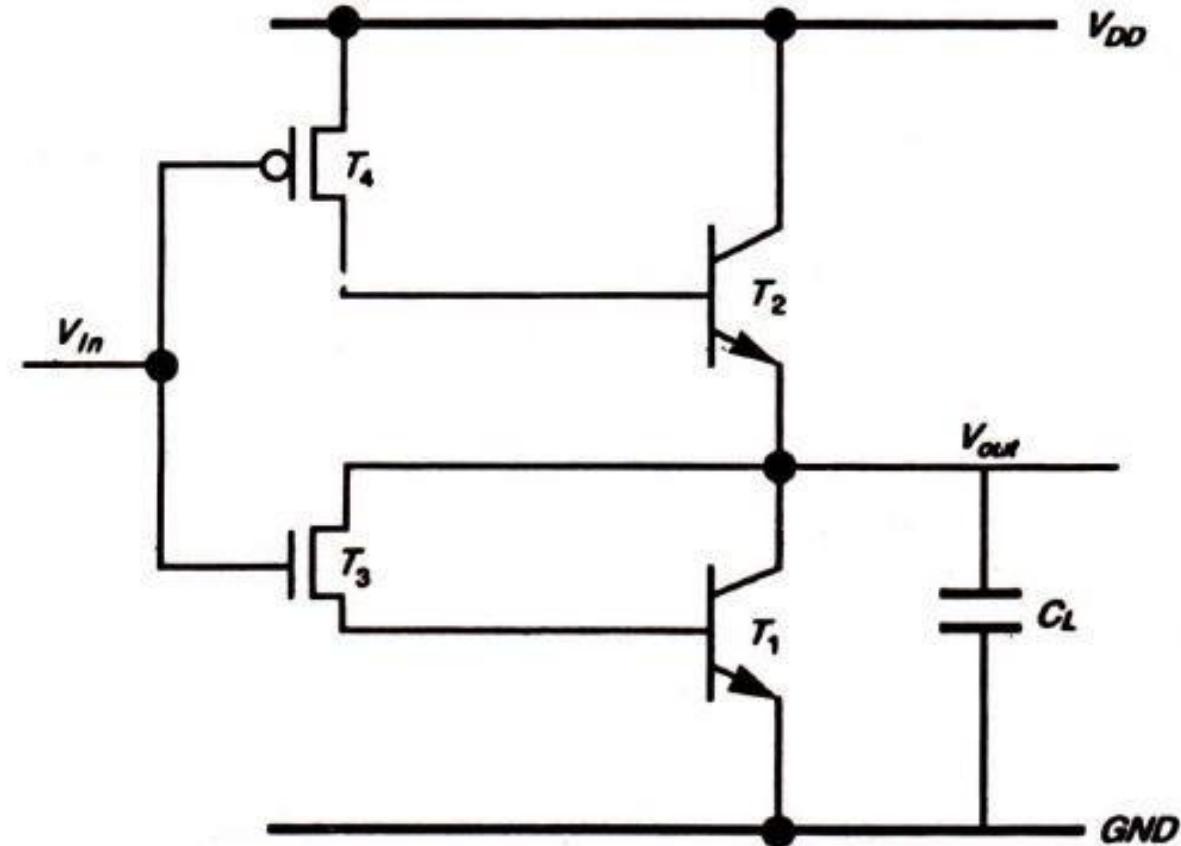
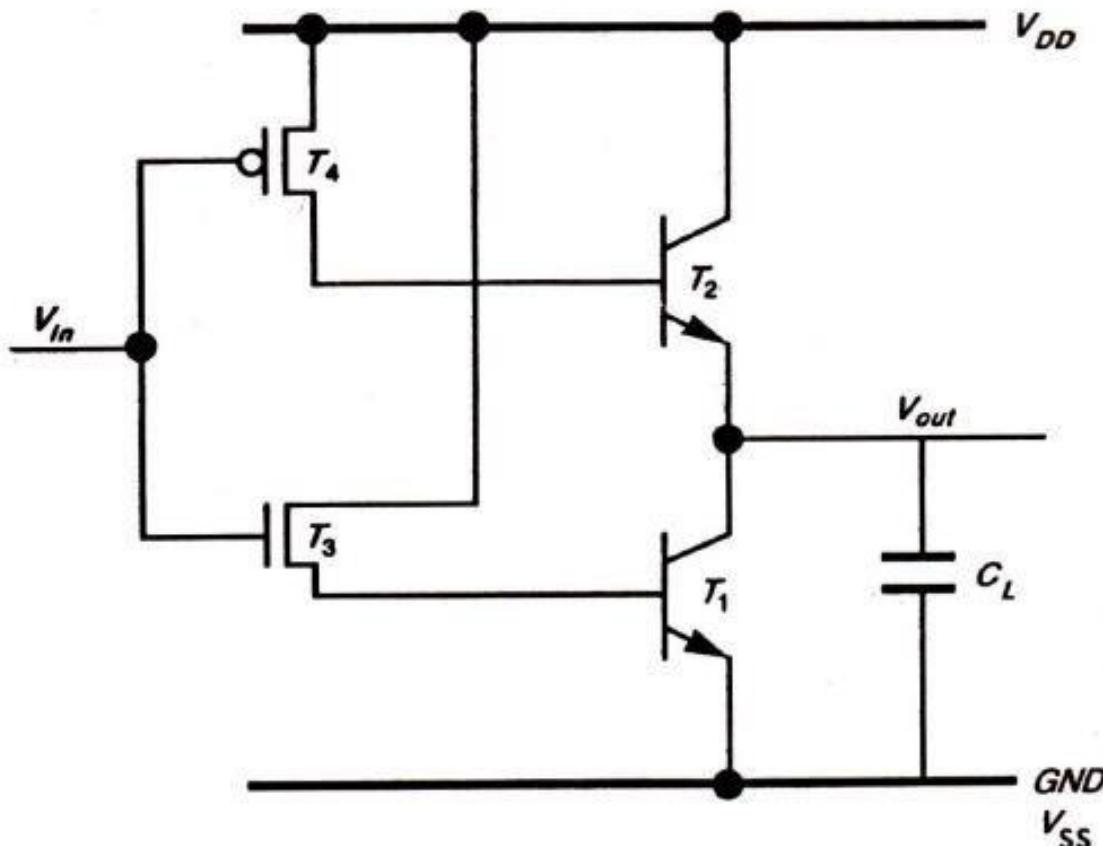
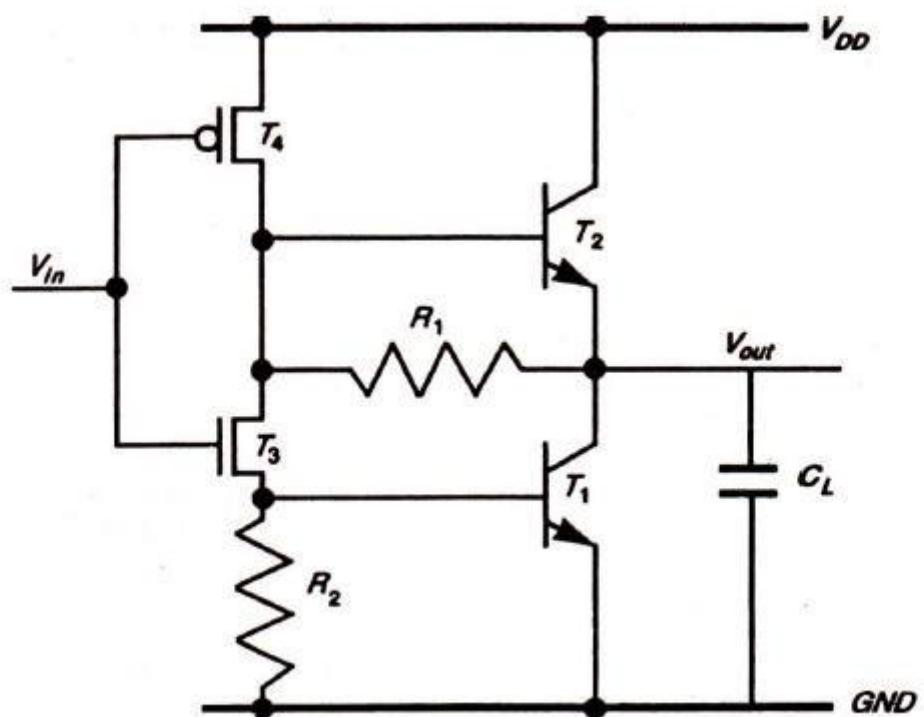
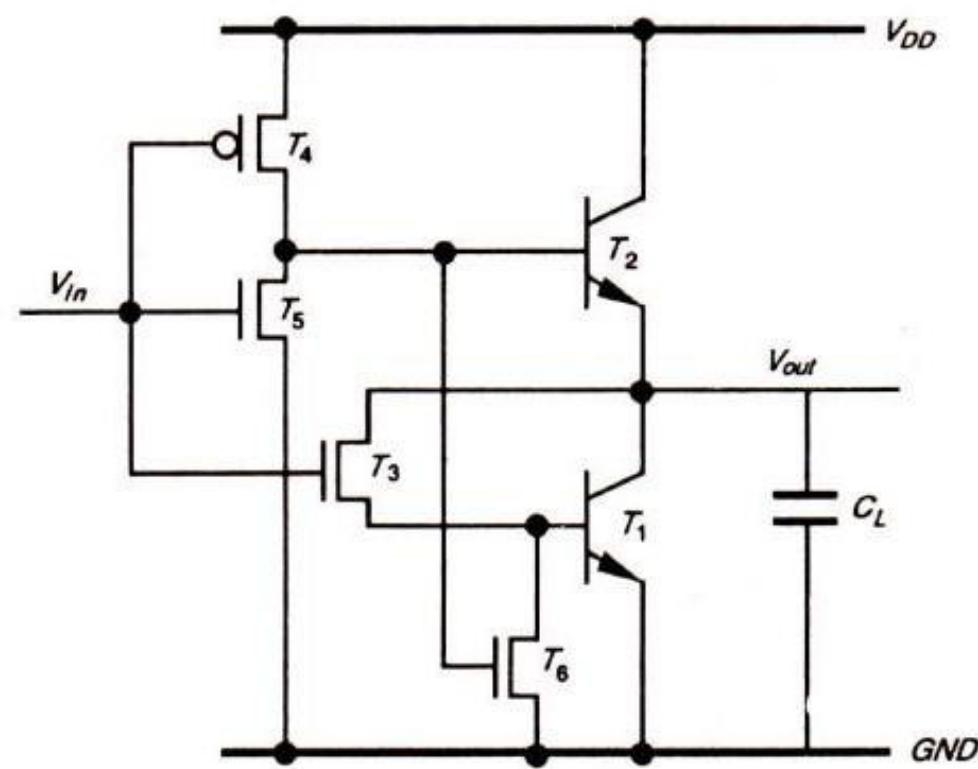


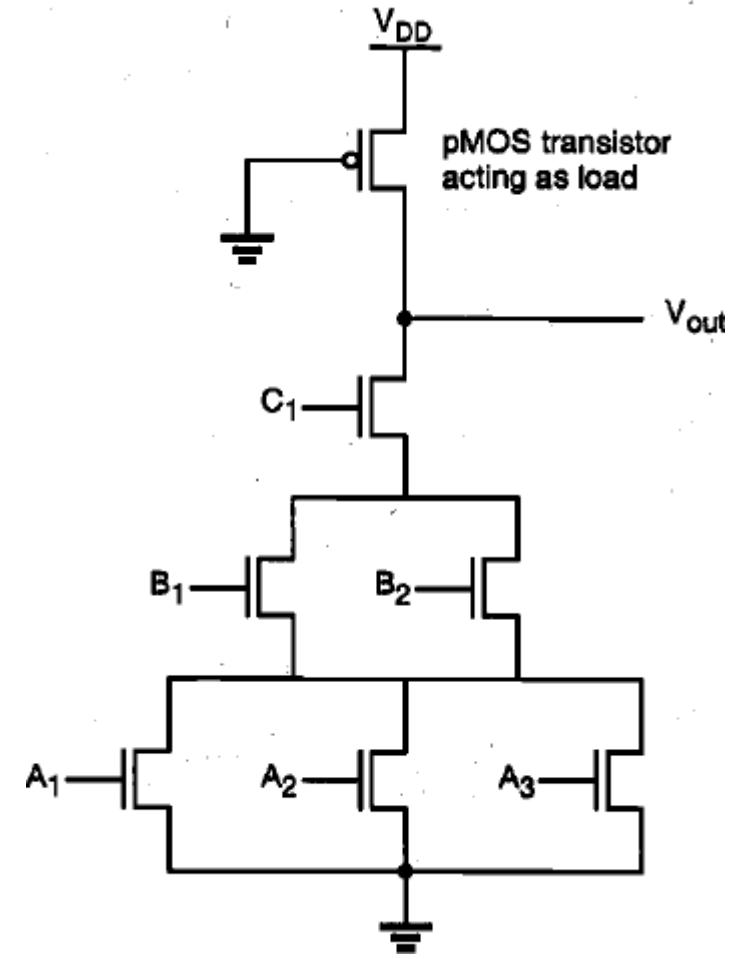
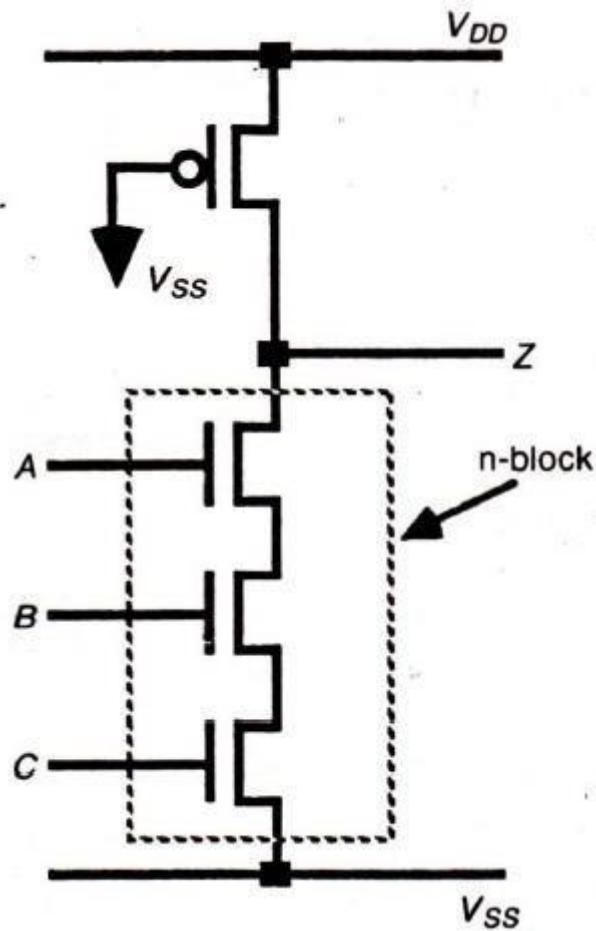
FIGURE 2.17 A simple BiCMOS inverter.



RE 2.19 An improved BiCMOS inverter with better output logic levels.



Pseudo NMOS Logic



The following parameters are given for an nMOS process:

$$t_{ox} = 500 \text{ \AA}$$

$$\text{substrate doping } N_A = 1 \cdot 10^{16} \text{ cm}^{-3}$$

$$\text{polysilicon gate doping } N_D = 1 \cdot 10^{20} \text{ cm}^{-3}$$

$$\text{oxide-interface fixed-charge density } N_{ox} = 2 \cdot 10^{10} \text{ cm}^{-3}$$

- (a) Calculate V_T for an unimplanted transistor.

(a) For unimplanted transistor,

$$\phi_F(\text{substrate}) = \frac{kT}{q} \ln \frac{n_i}{N_A} = 0.026 \ln \frac{1.45 \times 10^{10}}{1.0 \times 10^{16}} = -0.35[V]$$

$$\phi_F(\text{gate}) = \frac{kT}{q} \ln \frac{N_{D,poly}}{n_i} = 0.026 \ln \frac{1 \times 10^{20}}{1.45 \times 10^{10}} = 0.59[V]$$

$$\Phi_{GC} = \phi_F(\text{substrate}) - \phi_F(\text{gate}) = -0.35 - 0.59 = -0.94[V]$$

$$\begin{aligned} Q_{B0} &= \sqrt{2qN_{A,sub}\epsilon_{si}|2\phi_F|} \\ &= \sqrt{2 \times 1.6 \times 10^{-19} \times 10^{16} \times 11.7 \times 8.85 \times 10^{-14} \times 2 \times 0.35} \\ &= -4.82 \times 10^{-8}[C/cm^2] \end{aligned}$$

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3.9 \times 8.85 \times 10^{-14}}{500 \times 10^{-8}} = 6.9 \times 10^{-8}[F/cm^2]$$

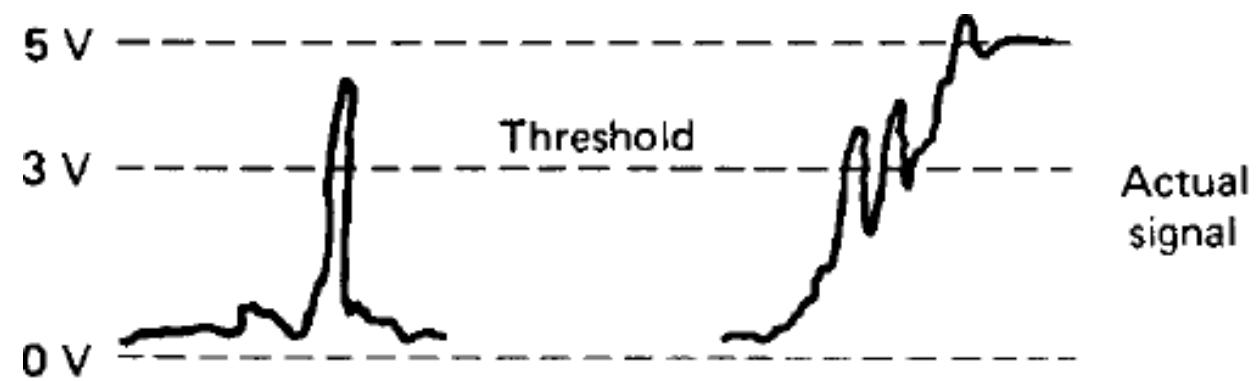
$$\begin{aligned} V_{T0} &= \Phi_{GC} - 2\phi_F - \frac{Q_{B0}}{C_{ox}} - \frac{Q_{ox}}{C_{ox}} \\ &= -0.94 - 2 \times (-0.35) - \frac{(-4.82 \times 10^{-8})}{6.9 \times 10^{-8}} - \frac{2 \times 10^{10} \times 1.6 \times 10^{-19}}{6.9 \times 10^{-8}} \\ &= 0.41[V] \end{aligned}$$

Clocked Logic

- Capacitance plays an important role in all digital circuits.
- The transitions between 1 and 0 represent the physical process of charging and discharging of the capacitance of the terminals and of interconnecting wires called the *parasitic capacitance*
- The voltages representing the logic values take finite response time to attain their final values contributing to the *delay* in the circuit

$$I = C \frac{dv}{dt}$$

- Capacitance means electrical inertia.
- Without this, circuits will become extremely sensitive to "glitches" or "hazards."
- There are two kinds of hazards. If the signal changes its value for a brief time, it is called a *static* hazard.
- A *dynamic* hazard occurs if the signal during its transition to its value bounces back and forth at least once.



(a)



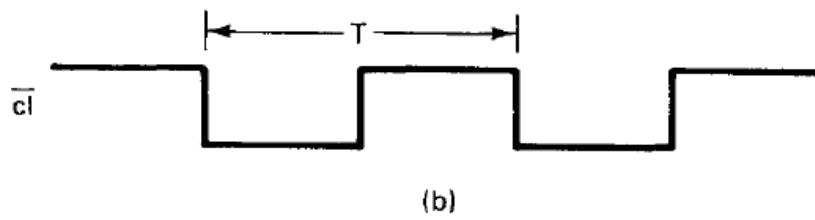
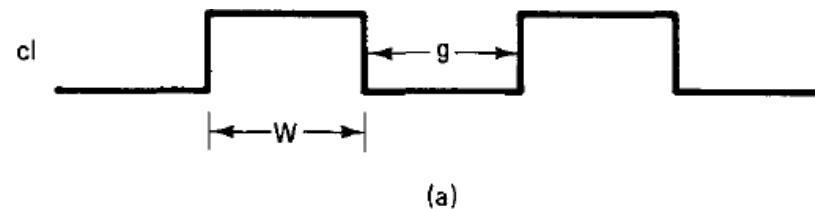
(b)

Logic
representation

- If the charging and discharging processes occur at discrete moments in synchrony with a *clock* pulse, the circuits are said to be *clocked* or *synchronous*.
- *Asynchronous* circuits are free-running circuits without clock pulse in which the timing relations between signals are determined by the inherent circuit delays.

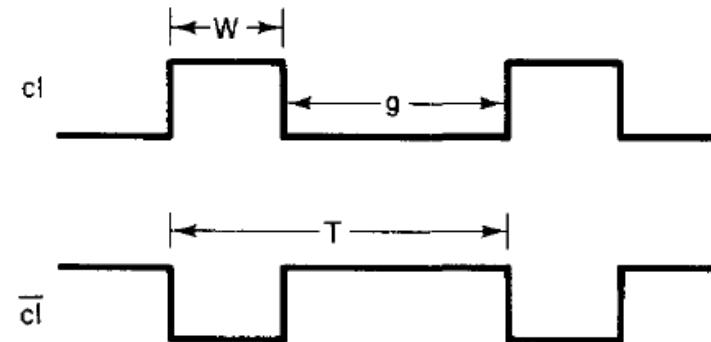
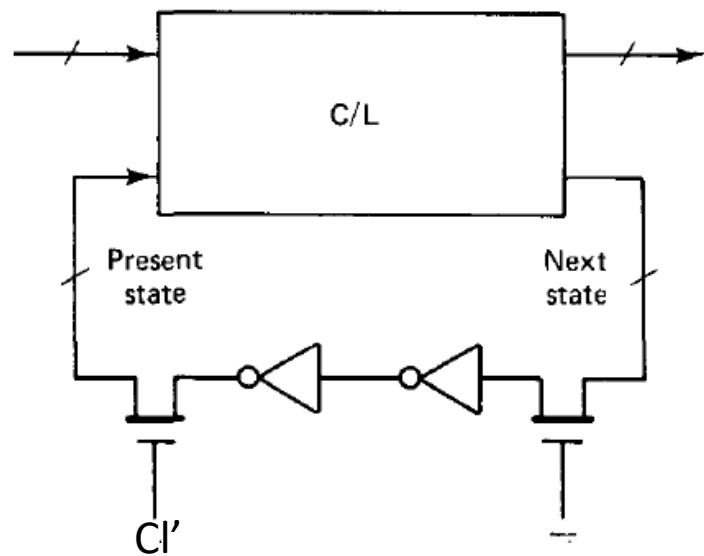
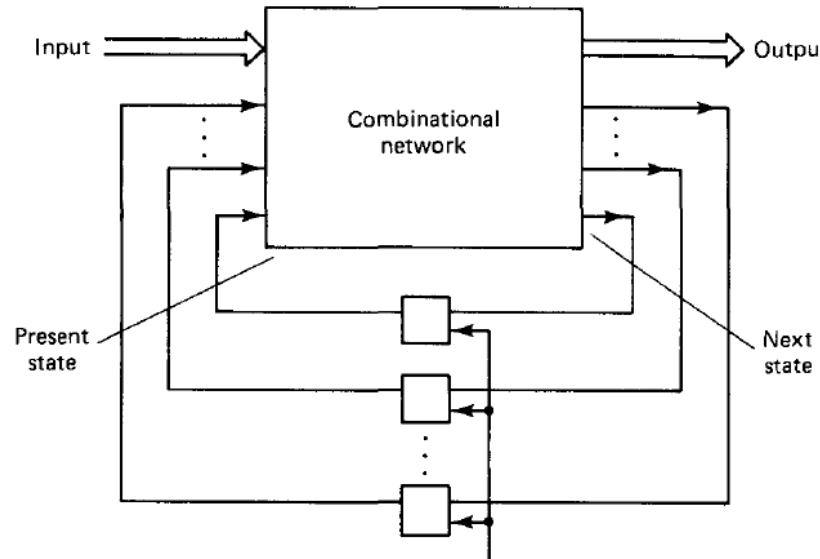
Two most commonly used clocking schemes are

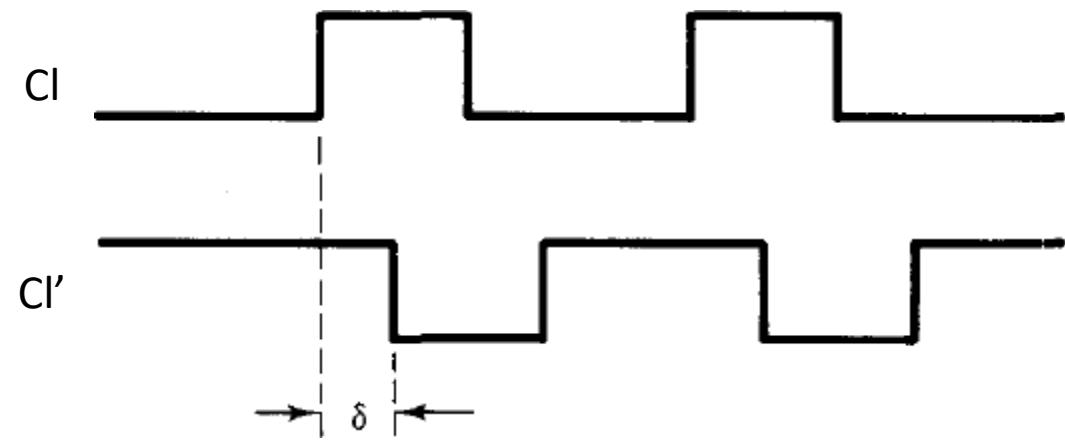
- *One-phase Clock* and
- *Two-phase nonoverlapping Clock*



(a) Single-phase clock; (b) its complement.

Problem with Single Phase Clock



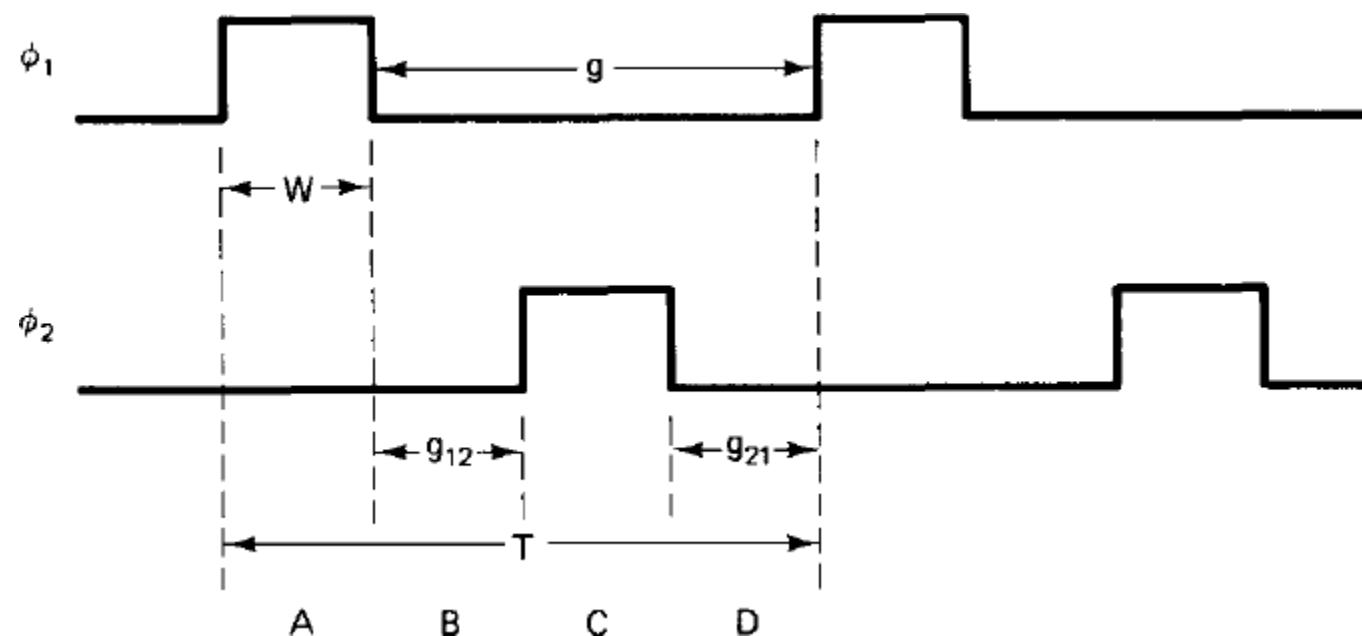
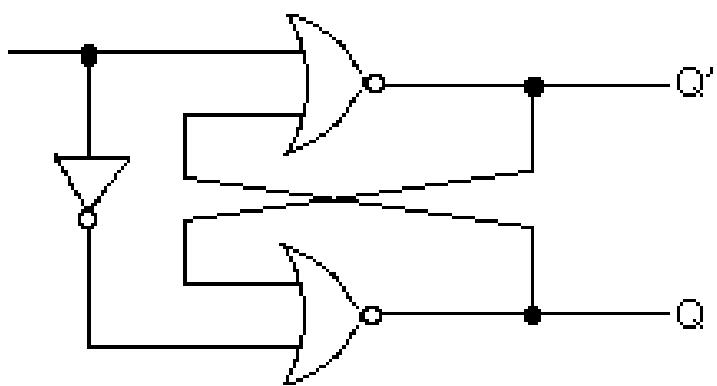


Problem with Single Phase Clock

- The clock width W must be greater than the delay time for the combinational logic, but cannot be more than the minimum delay through the network because this might cause a *multistepping* or *race* condition
- Clock period T must be greater than the longest delay in the combinational network so that the computation of the *next state* propagates through the feedback path to be latched by the *next* clock pulse. But T cannot be indefinitely large if dynamic storage elements are used to store the next-state information, because information may be lost due to leakage.

Problem with Single Phase Clock

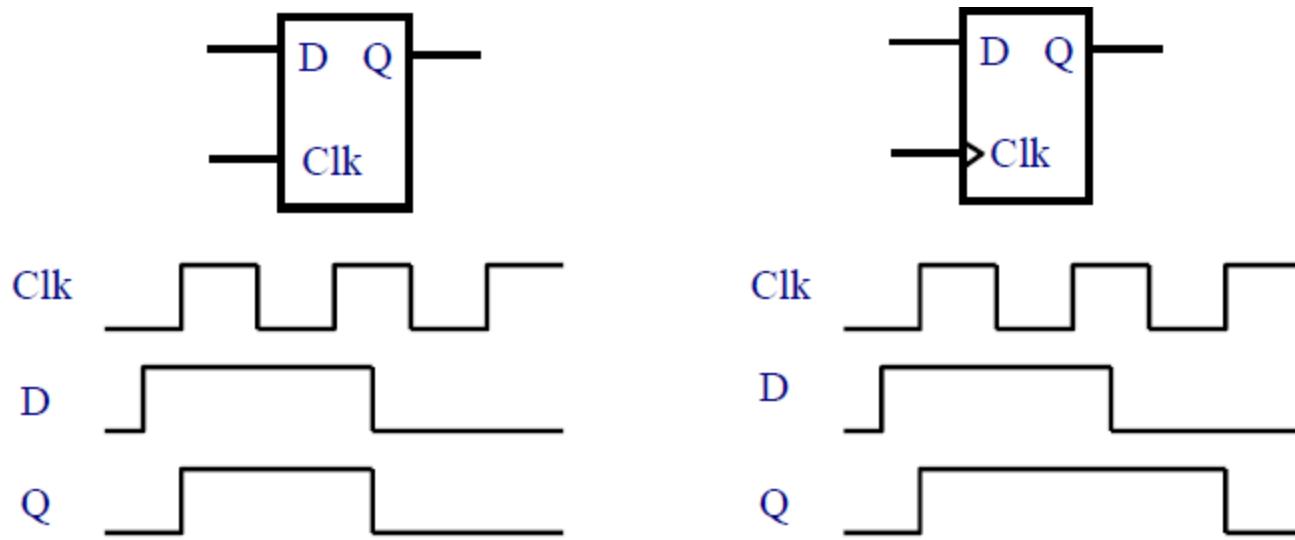
- The clock period must be less than the necessary refresh time for the dynamic elements.
- The interclock gap g must also satisfy some condition in order to make the circuit insensitive to spurious hazards at the input. This can be done by making g long enough so that any transient activity due to the glitches may fade out before the arrival of the next clock



Two-phase nonoverlapping clock

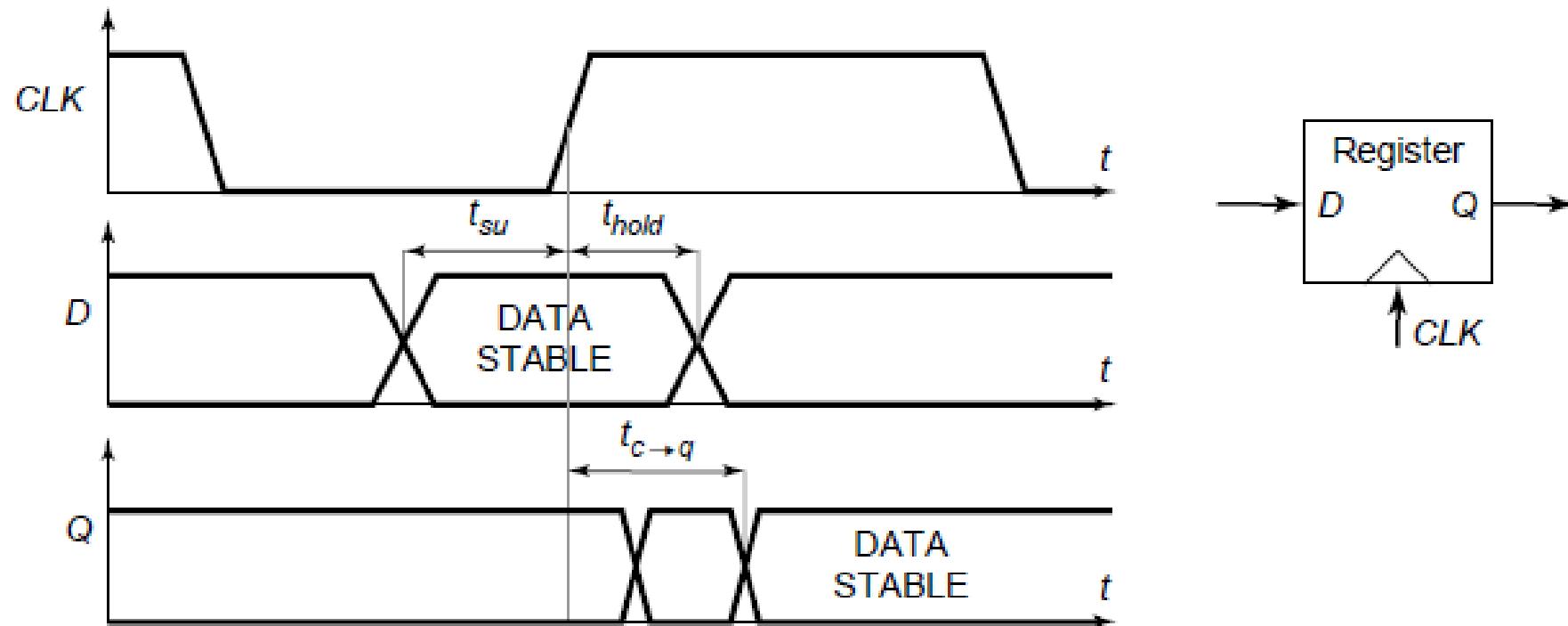
Dynamic Memory element & Timing

Sequential Elements



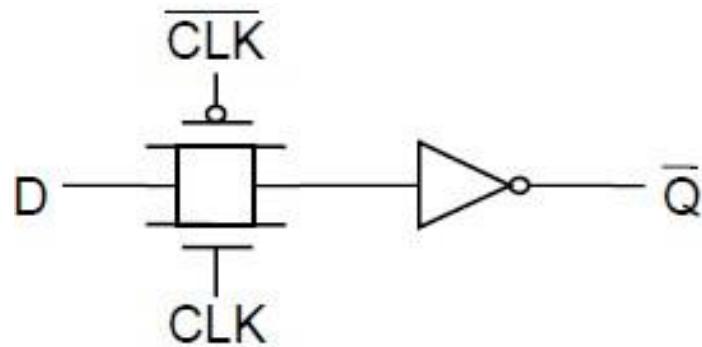
- Latch – level sensitive**
 - Clk=0: “opaque”
 - Clk-1: “transparent”
- Flip-flop – edge triggered**
 - Stores new data when Clk rises

Dynamic Memory element & Timing

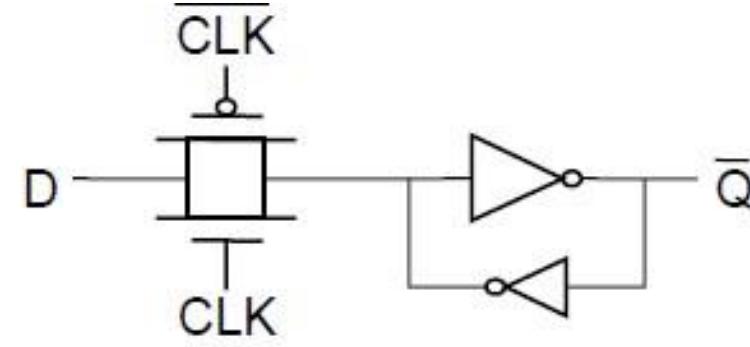


Dynamic Memory element & Timing

Storage Mechanisms

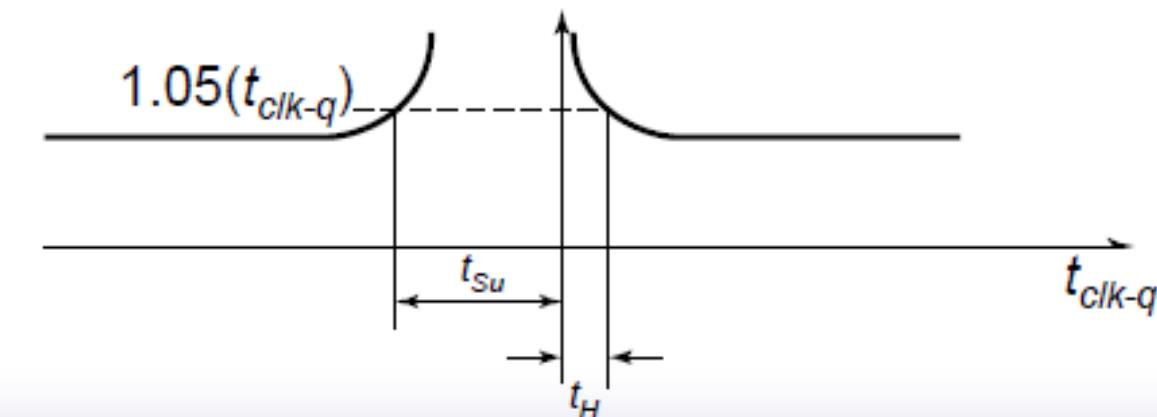
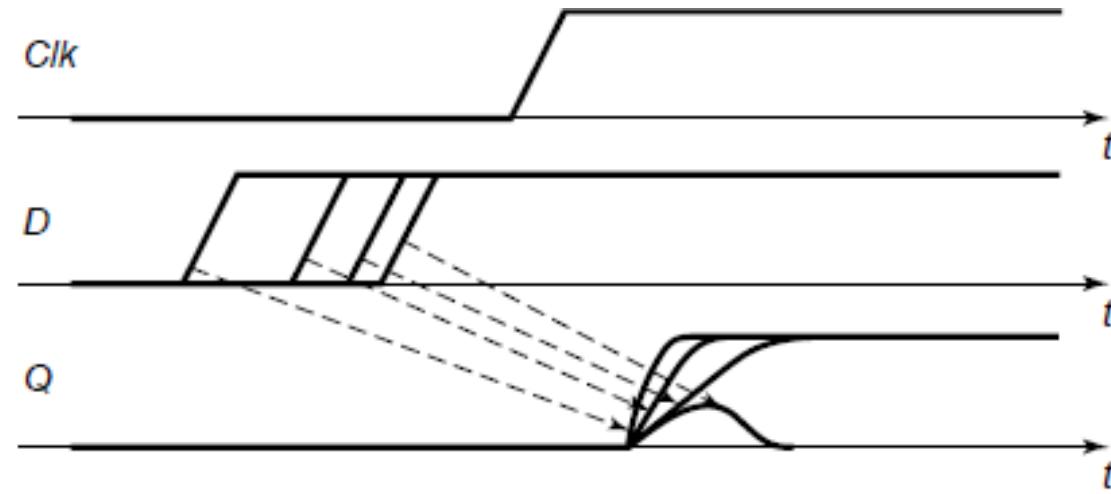


Dynamic Latch



Static Latch

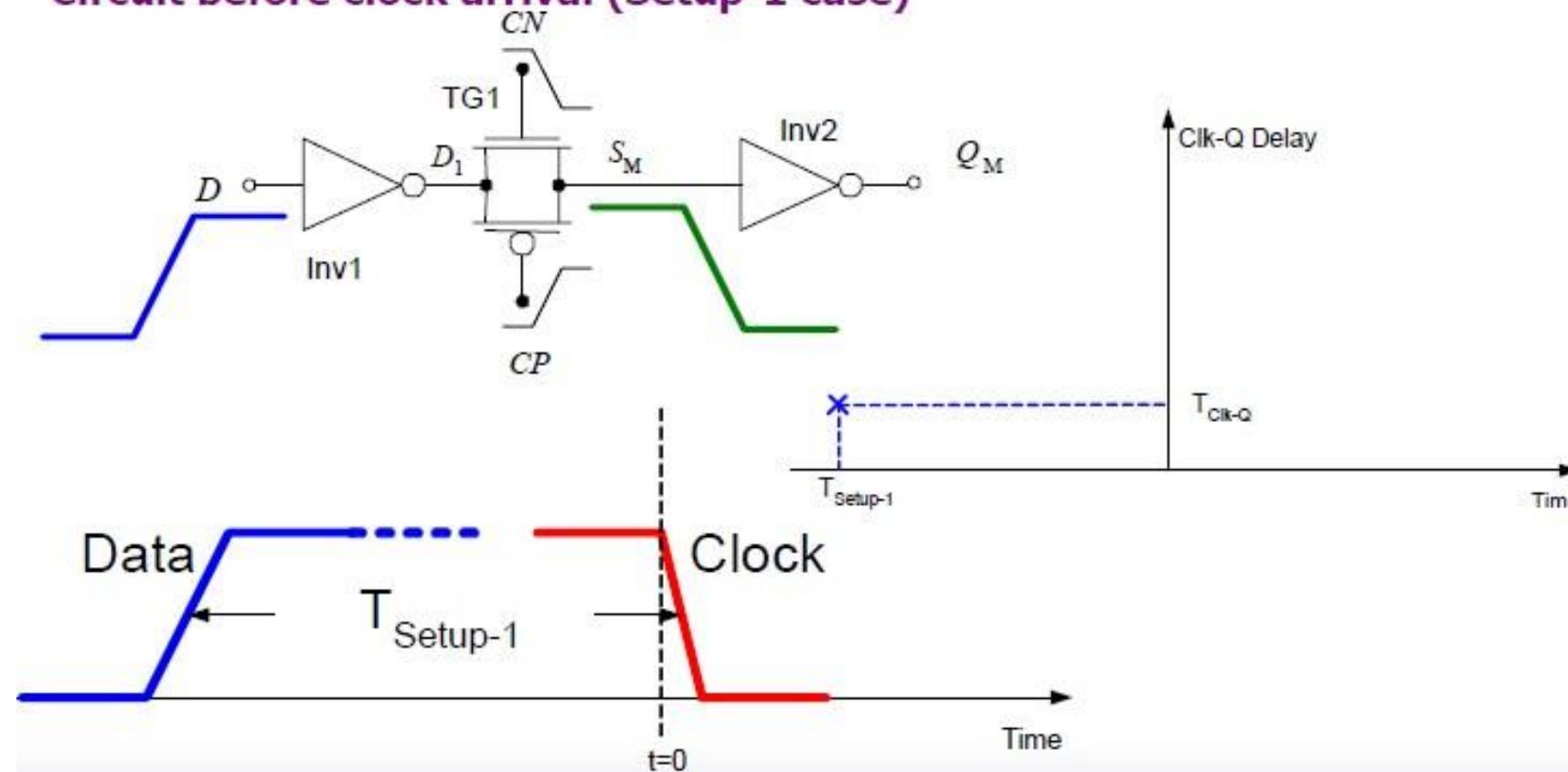
Timing



Timing

Setup Time Illustrations

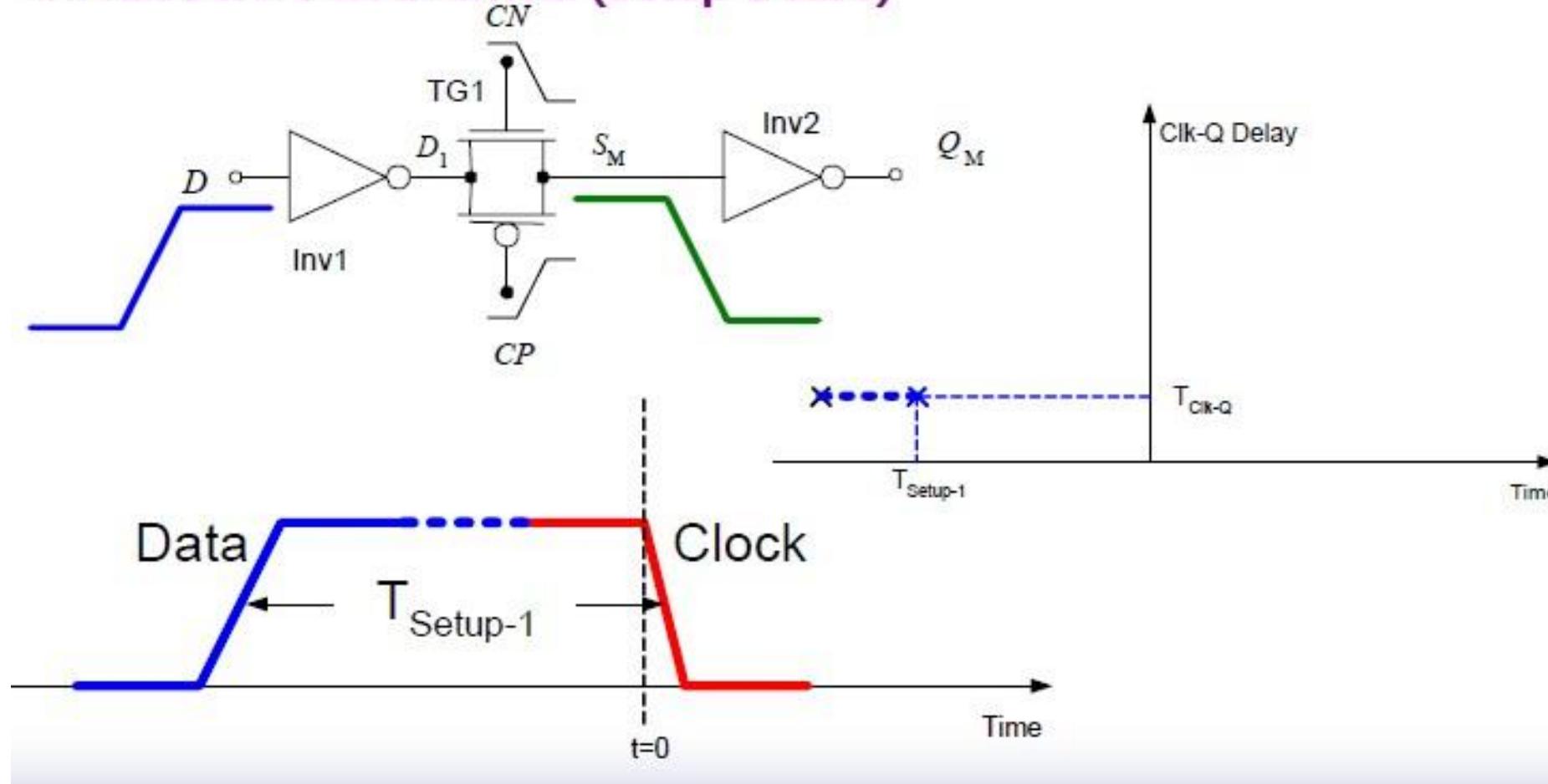
Circuit before clock arrival (Setup-1 case)



Timing

Setup Time Illustrations

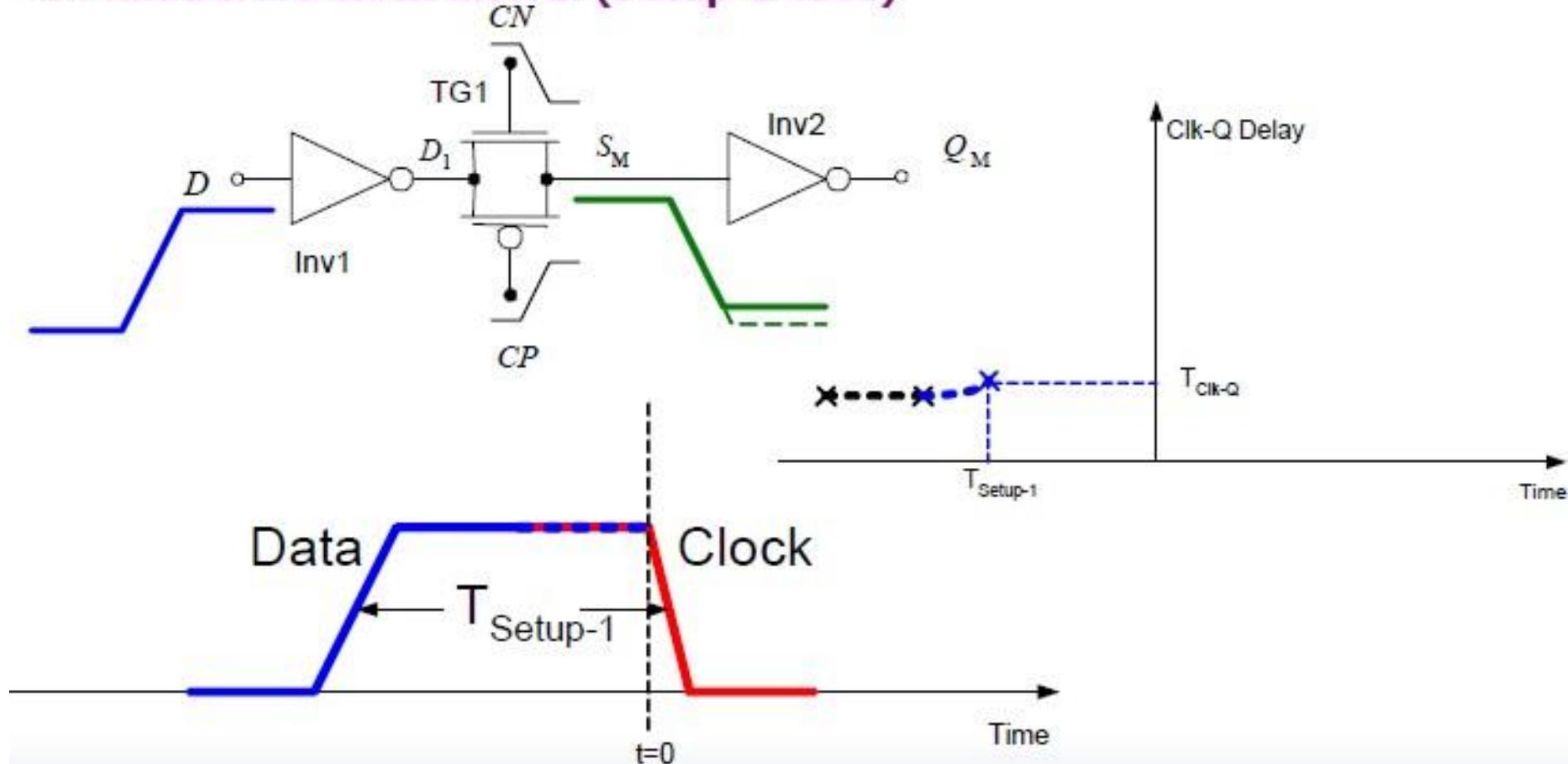
Circuit before clock arrival (Setup-1 case)



Timing

Setup Time Illustrations

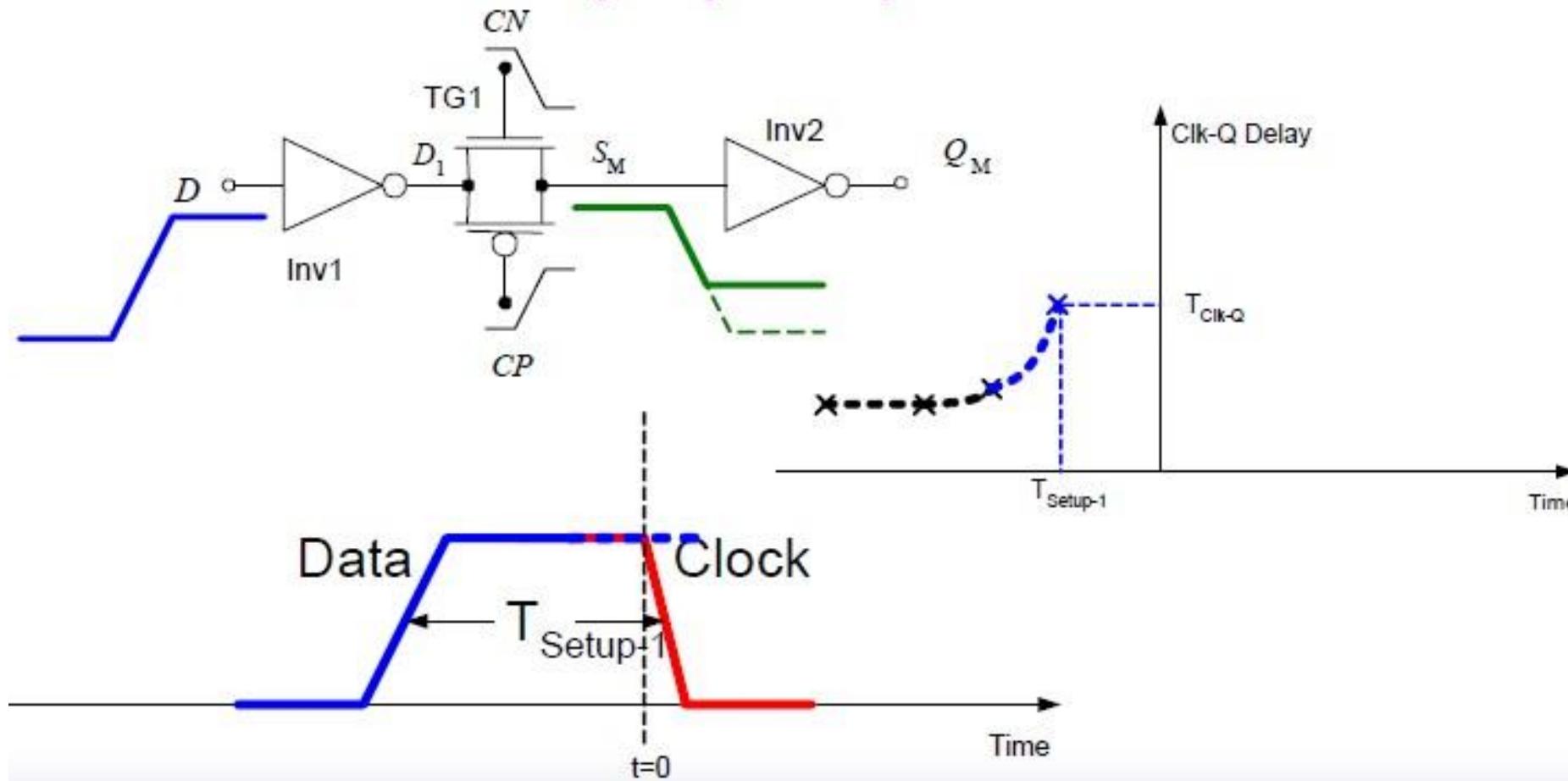
Circuit before clock arrival (Setup-1 case)



Timing

Setup Time Illustrations

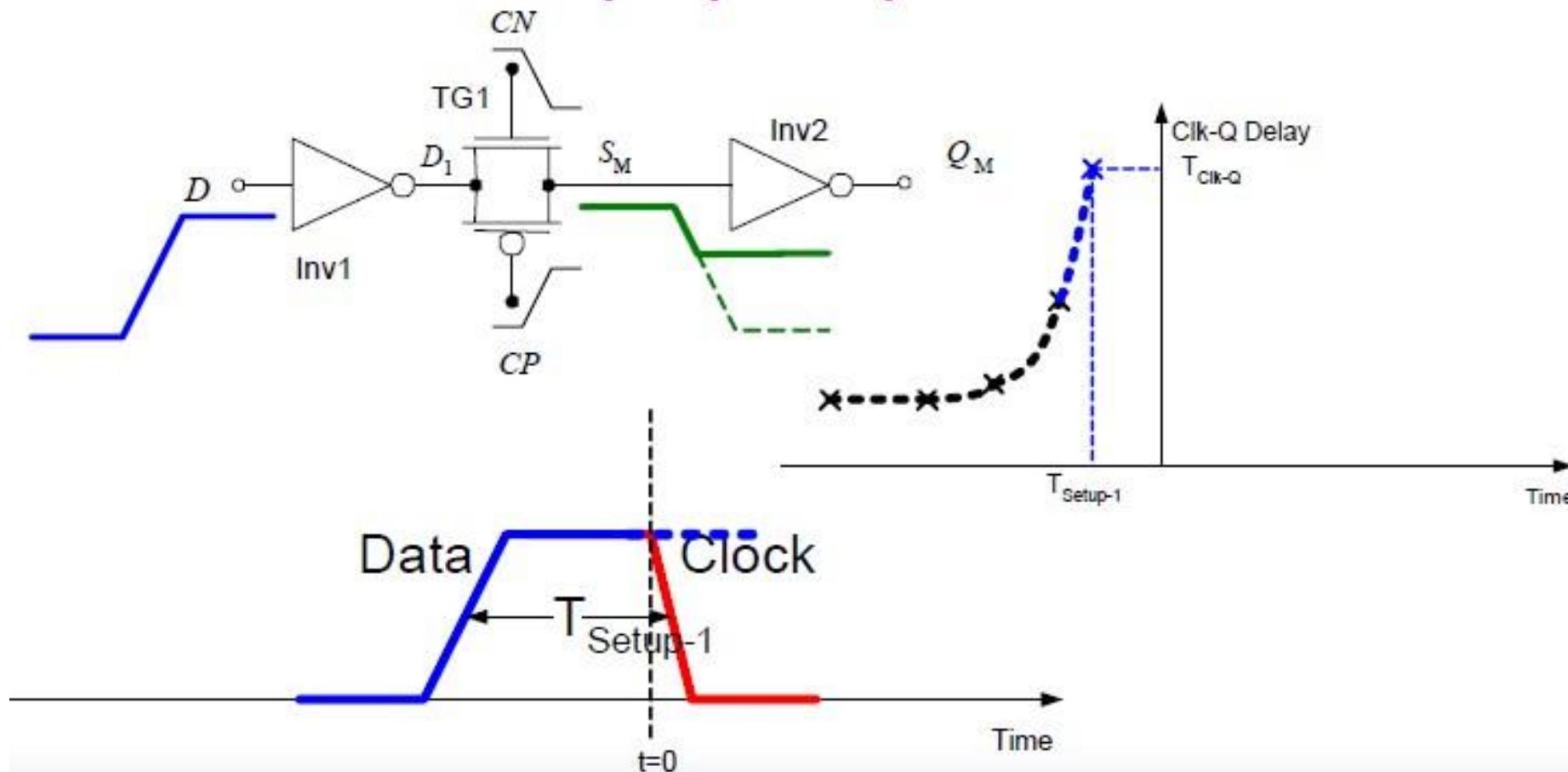
Circuit before clock arrival (Setup-1 case)



Timing

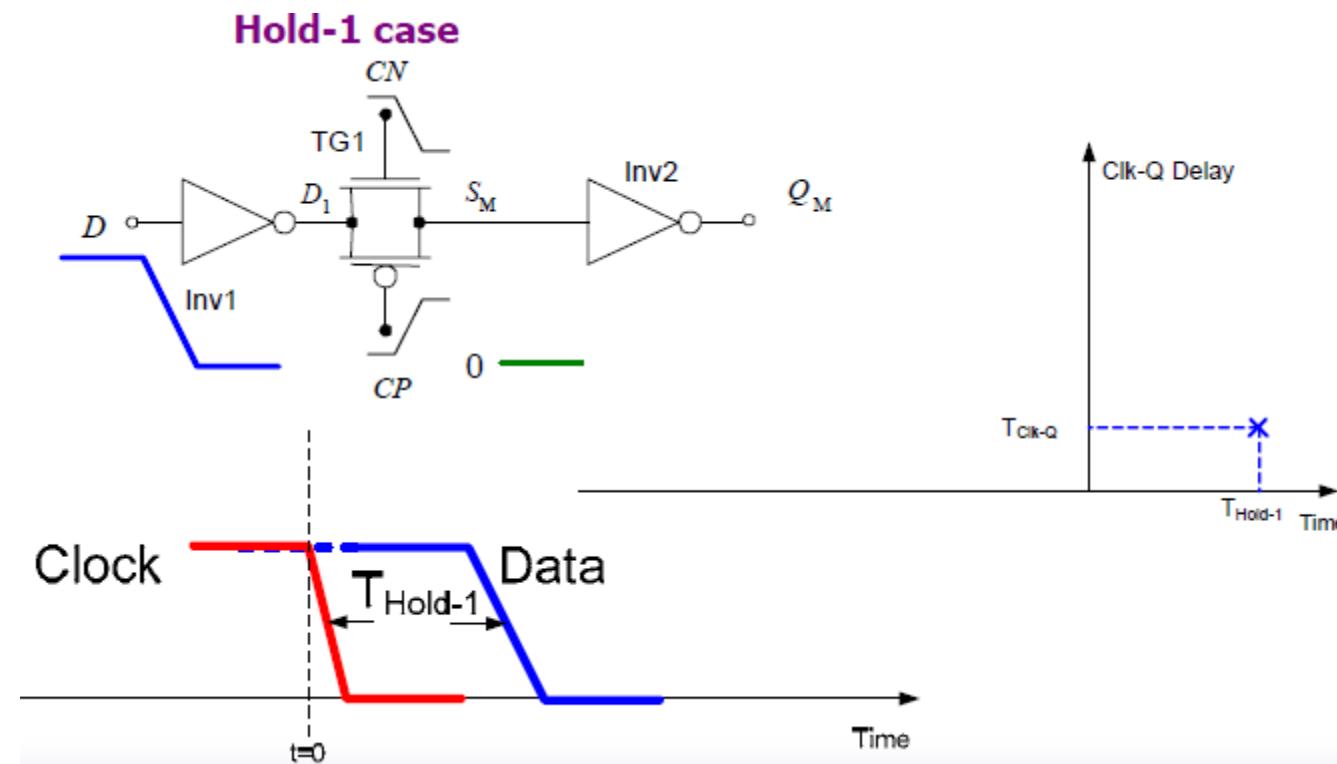
Setup Time Illustrations

Circuit before clock arrival (Setup-1 case)



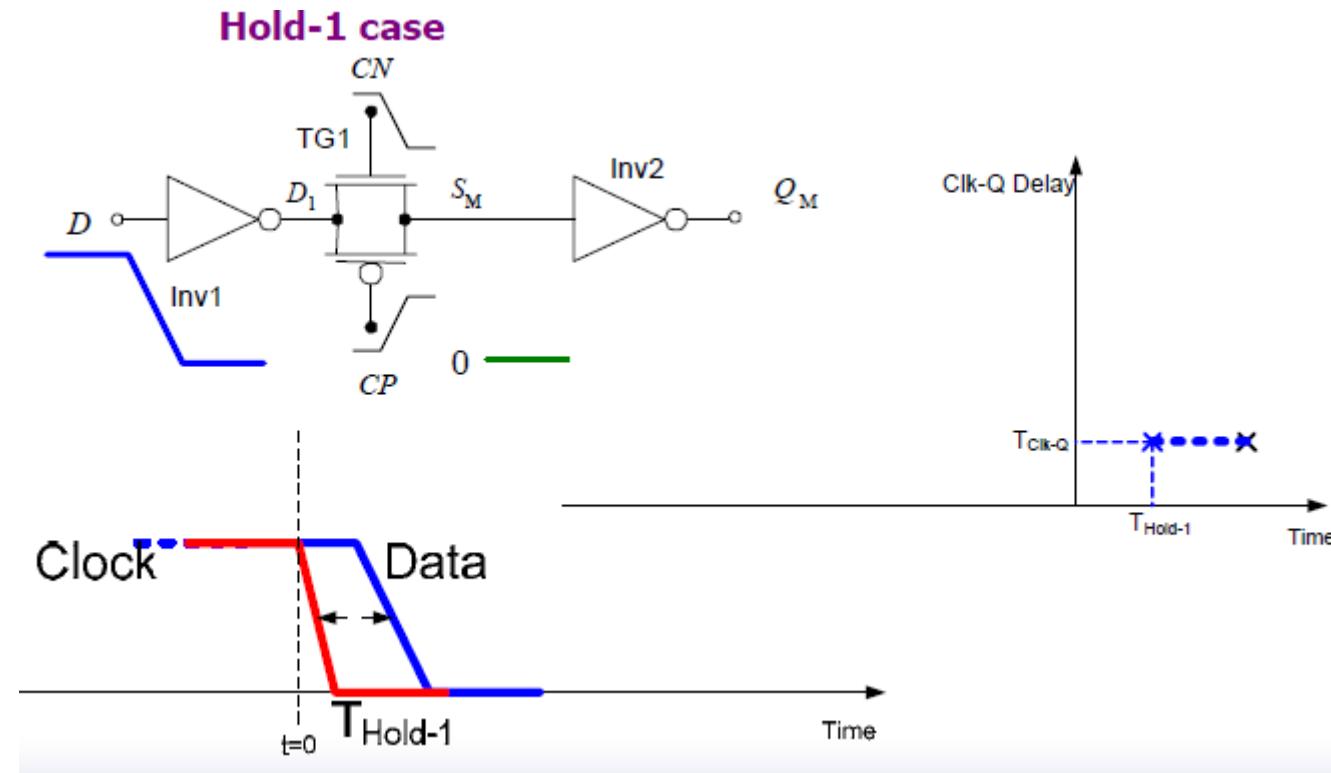
Timing

Hold Time Illustrations



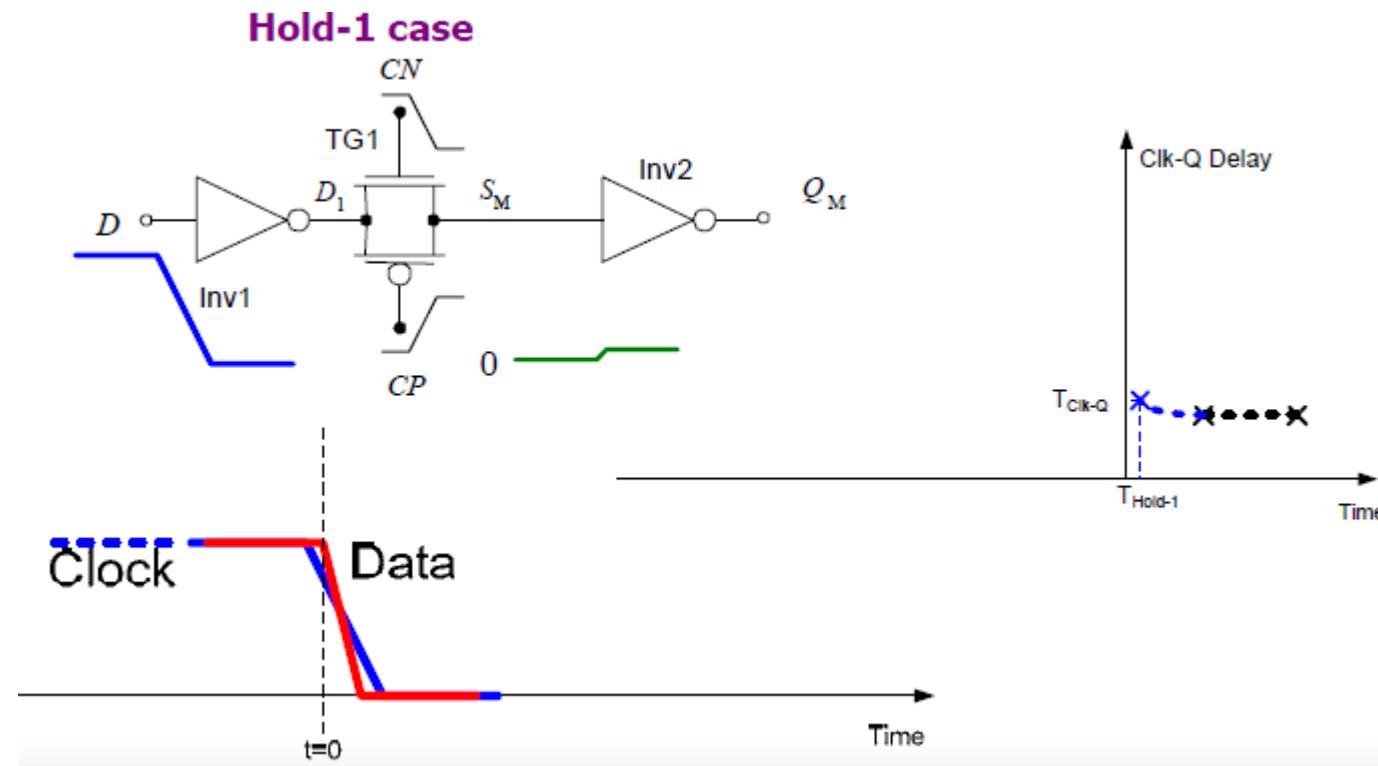
Timing

Hold Time Illustrations

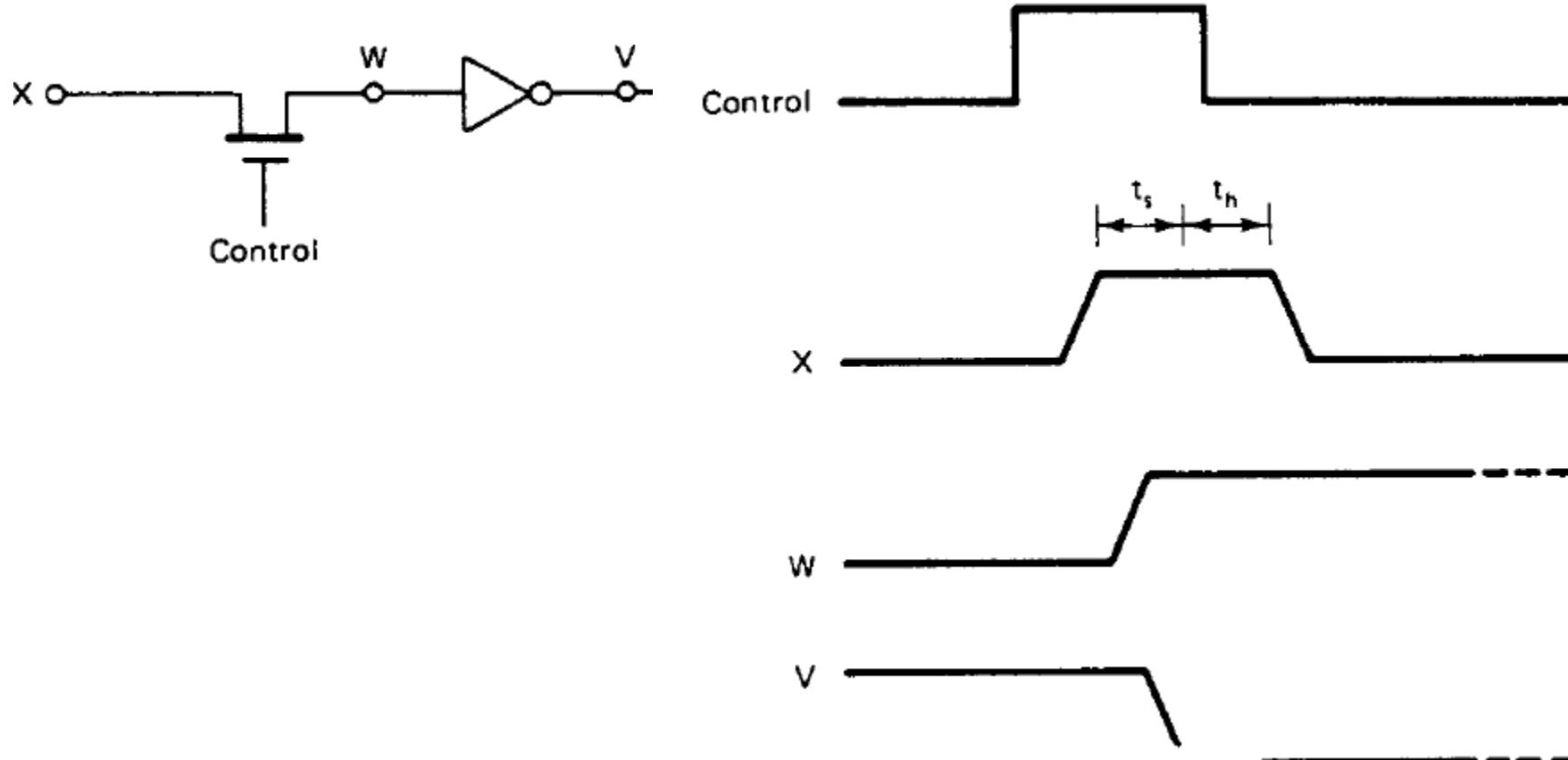


Timing

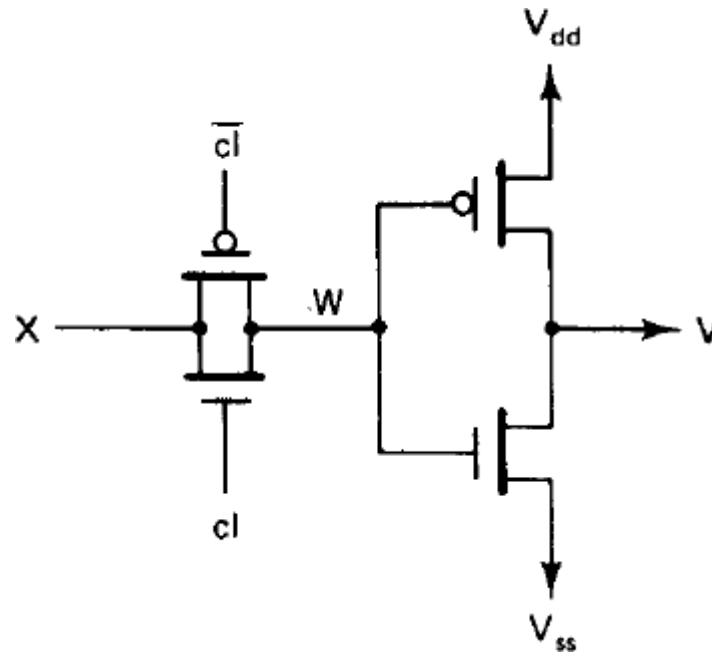
Hold Time Illustrations



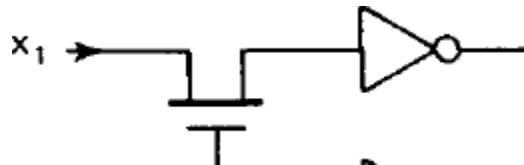
NMOS Dynamic latch



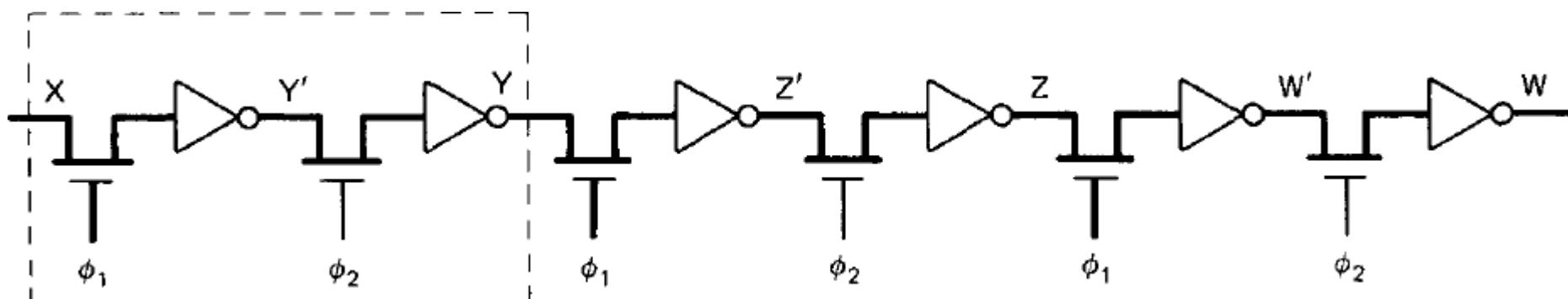
CMOS dynamic latch



Dynamic Shift register

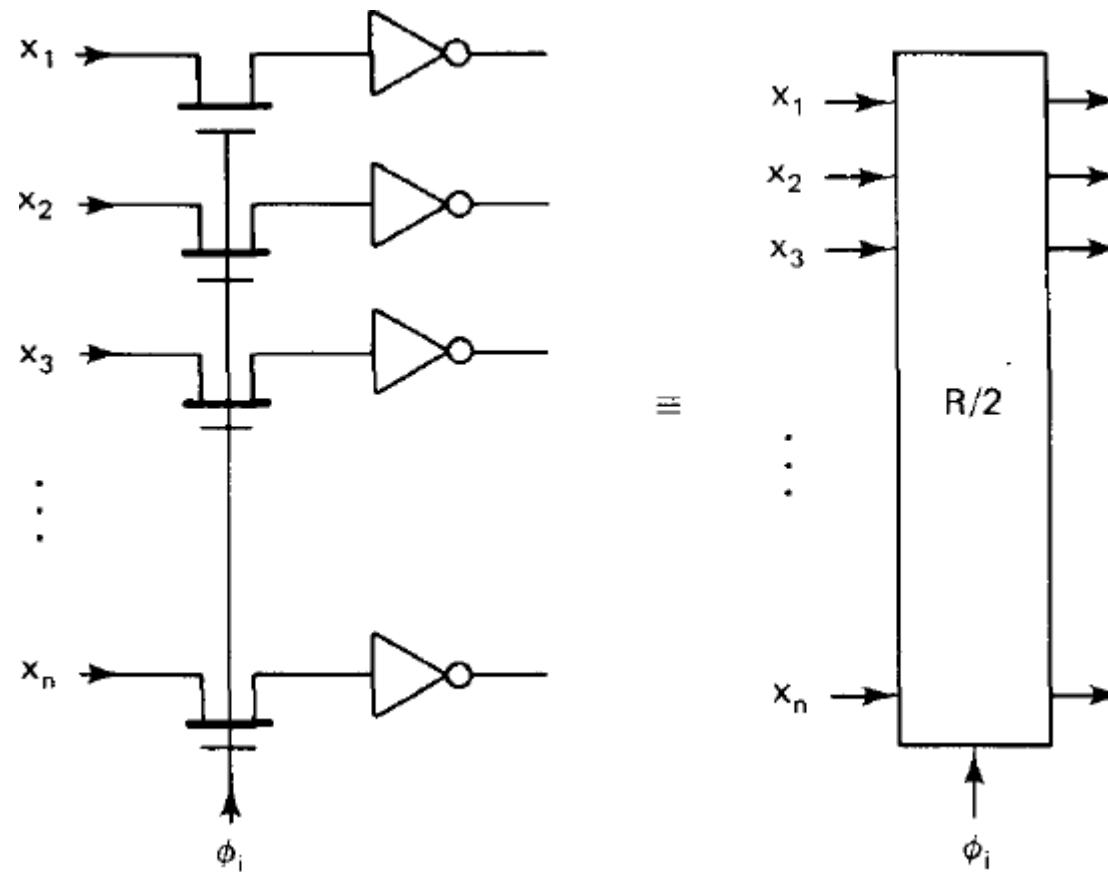


One bit Half-register stage

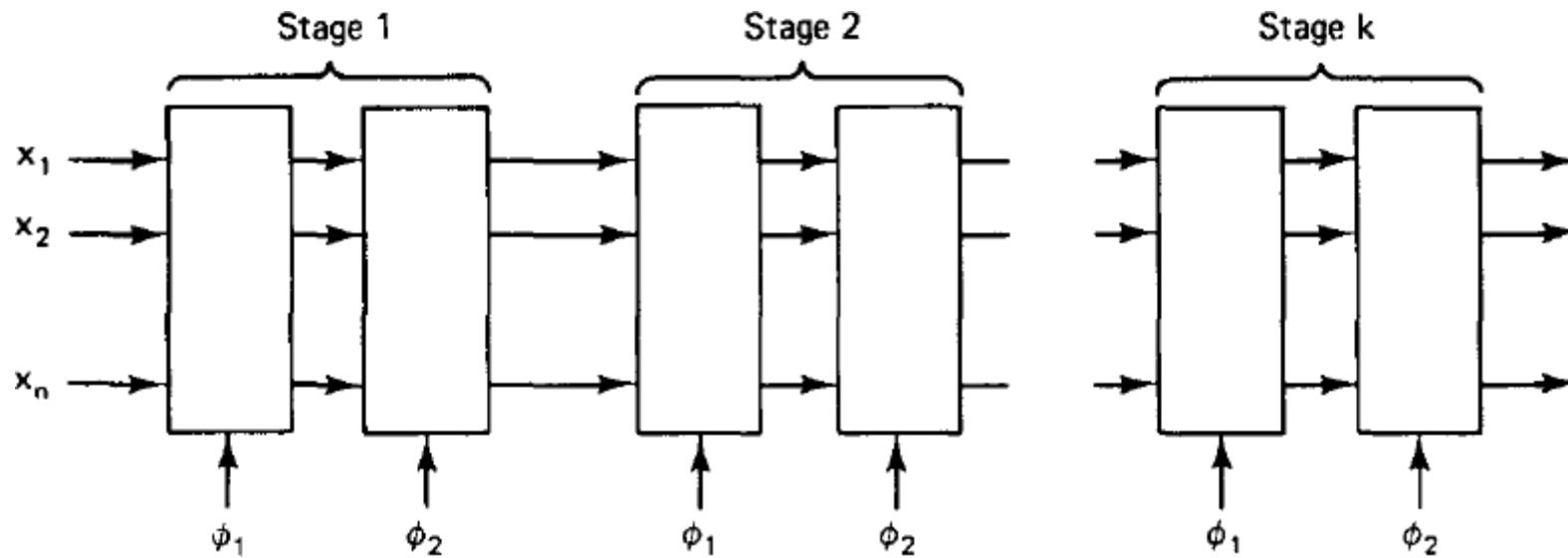


Dynamic shift register for one bit

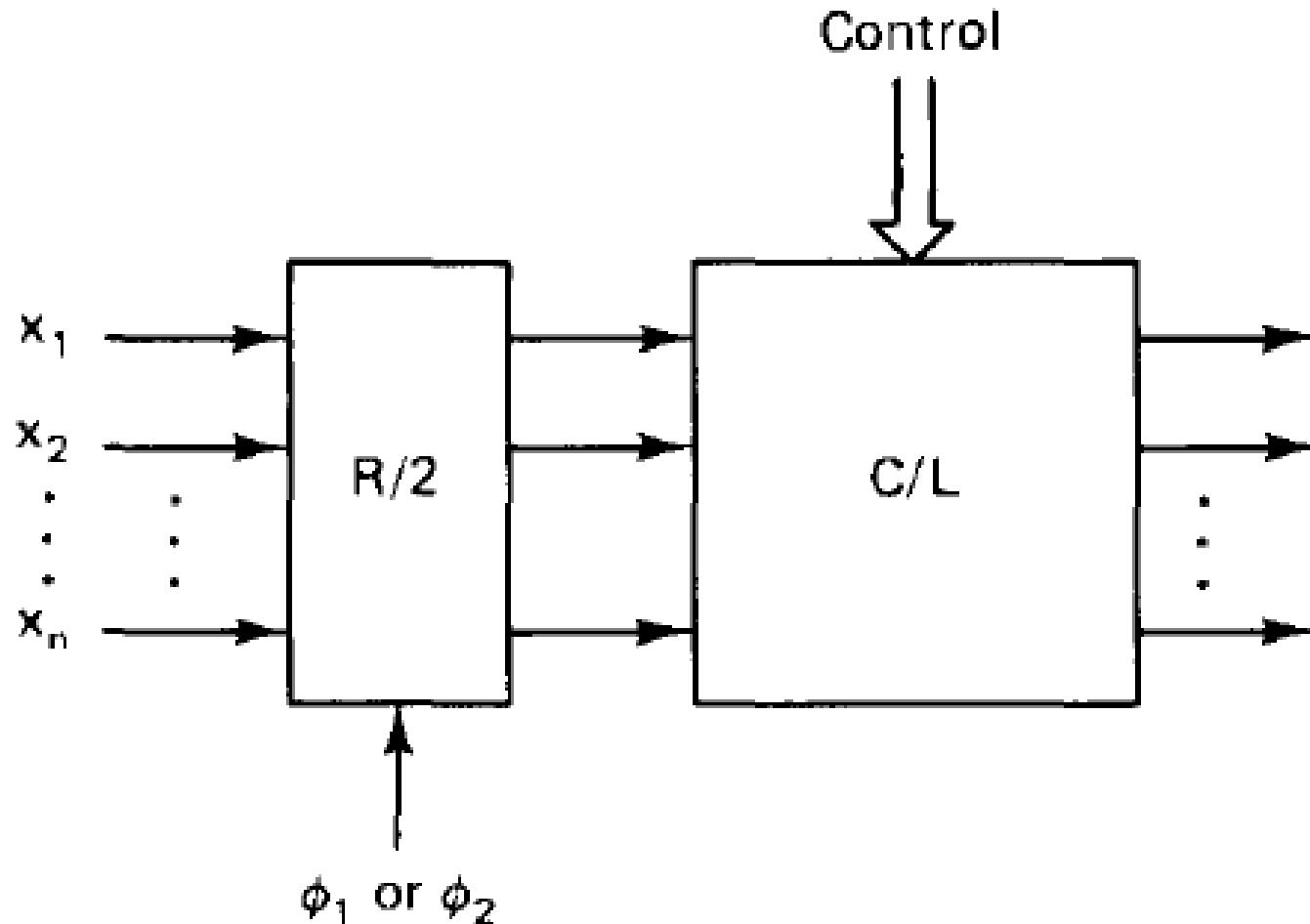
N-Bit Half Register Stage



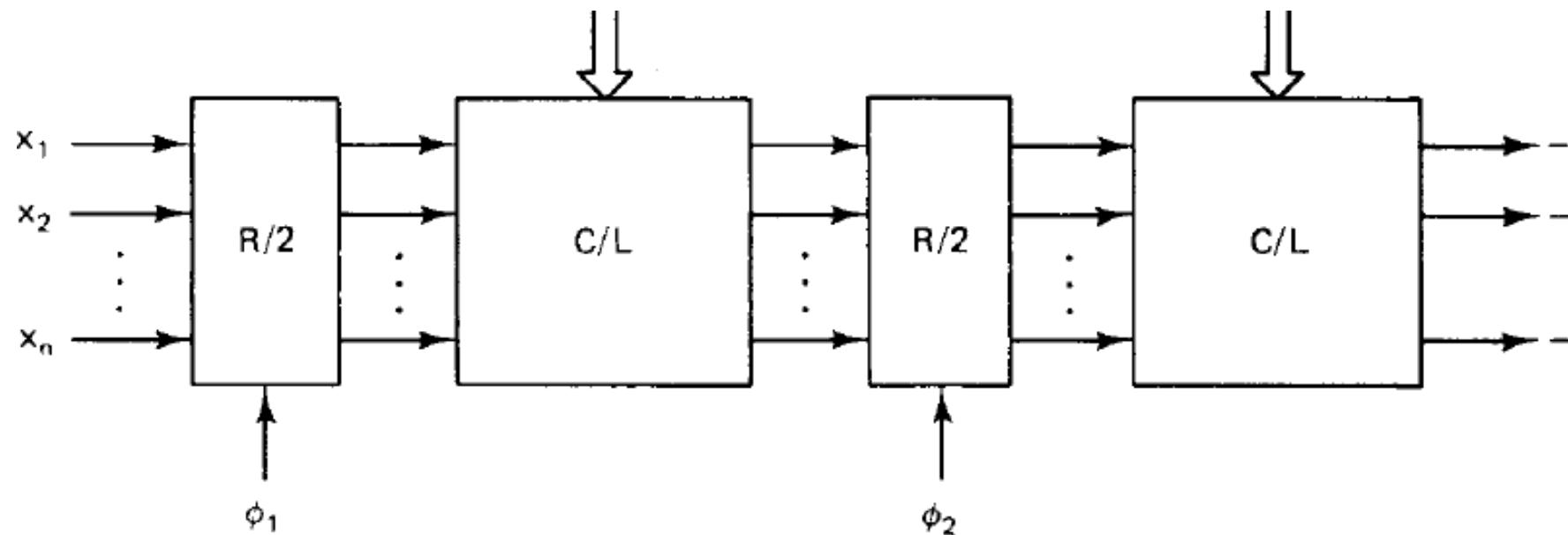
N-bit shift register K-Stage



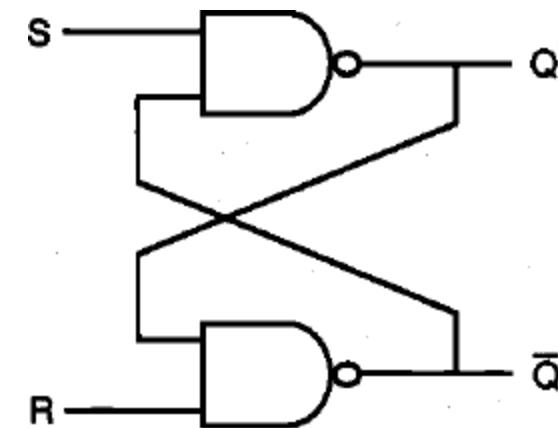
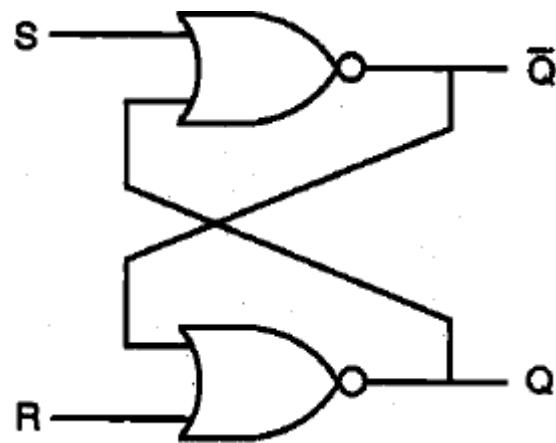
Half-stage of a data path.



An n MOS data path



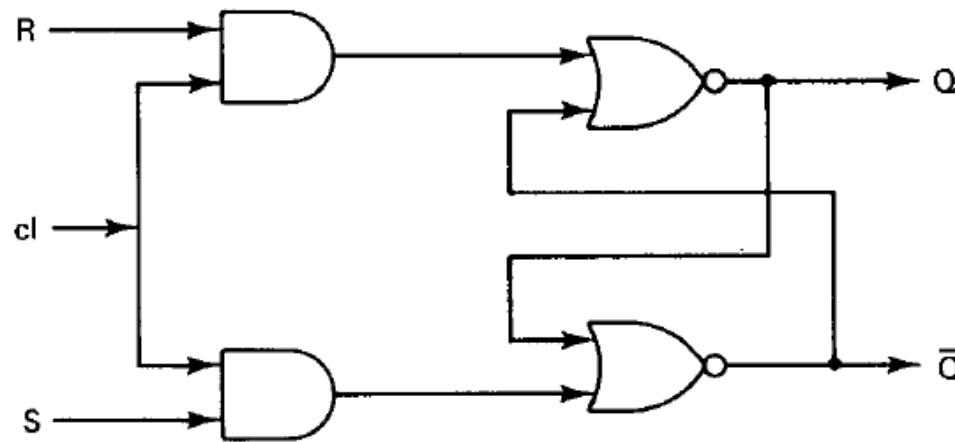
Static Storage Element



S	R	Q_{n+1}	\bar{Q}_{n+1}	Operation
0	0	Q_n	\bar{Q}_n	hold
1	0	1	0	set
0	1	0	1	reset
1	1	0	0	not allowed

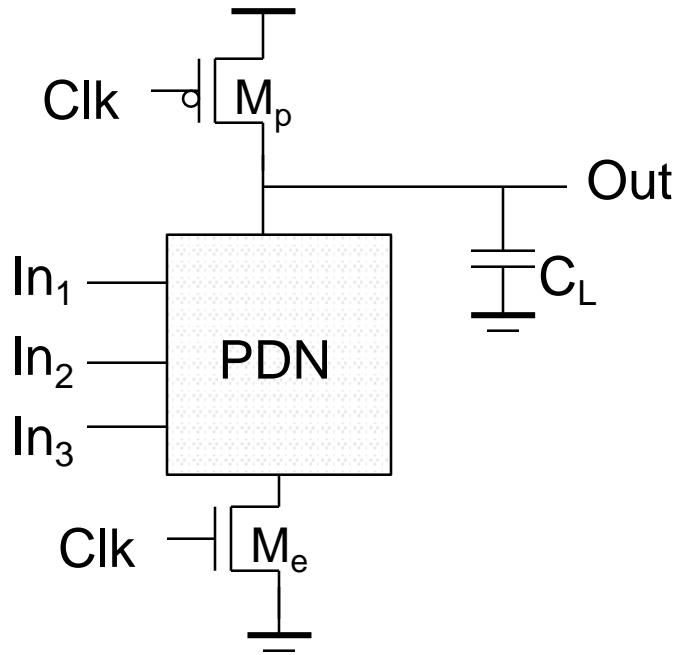
S	R	Q_{n+1}	\bar{Q}_{n+1}	Operation
0	0	1	1	not allowed
0	1	1	0	set
1	0	0	1	reset
1	1	Q_n	\bar{Q}_n	hold

Static Storage Element



$S(t)$	$R(t)$	$Q(t + 1)$	$\bar{Q}(t + 1)$
1	0	1	0
0	1	0	1
0	0	$Q(t)$	$\bar{Q}(t)$
1	1	Not allowed	

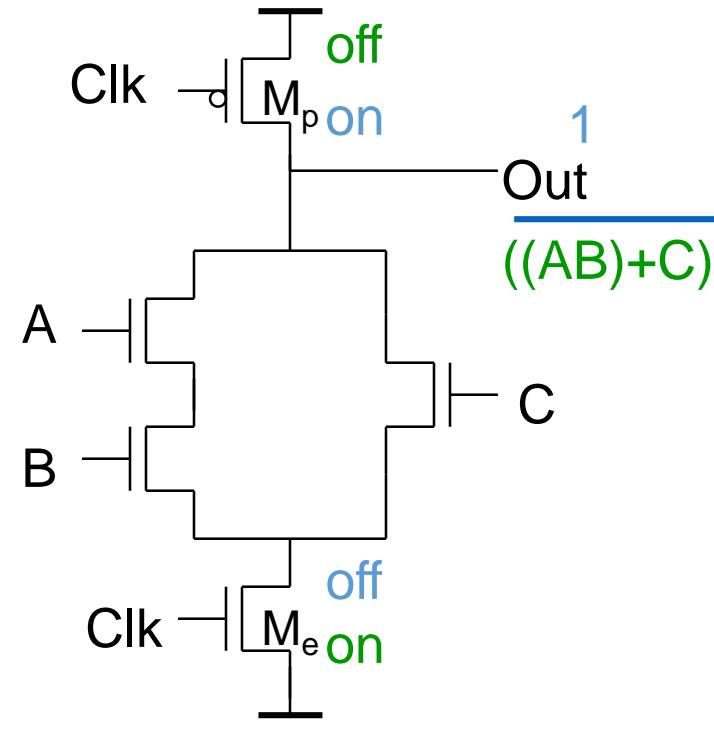
Dynamic CMOS Logic



Two phase operation

Precharge (Clk = 0)

Evaluate (Clk = 1)



Dynamic CMOS Logic

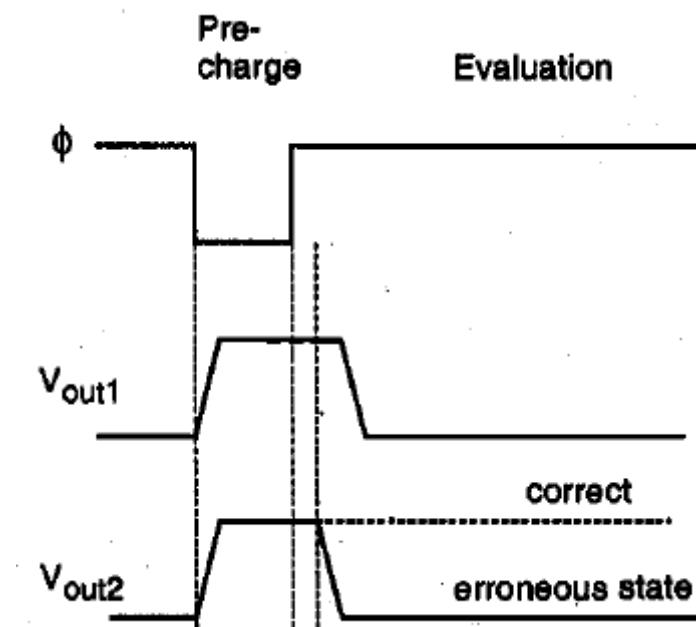
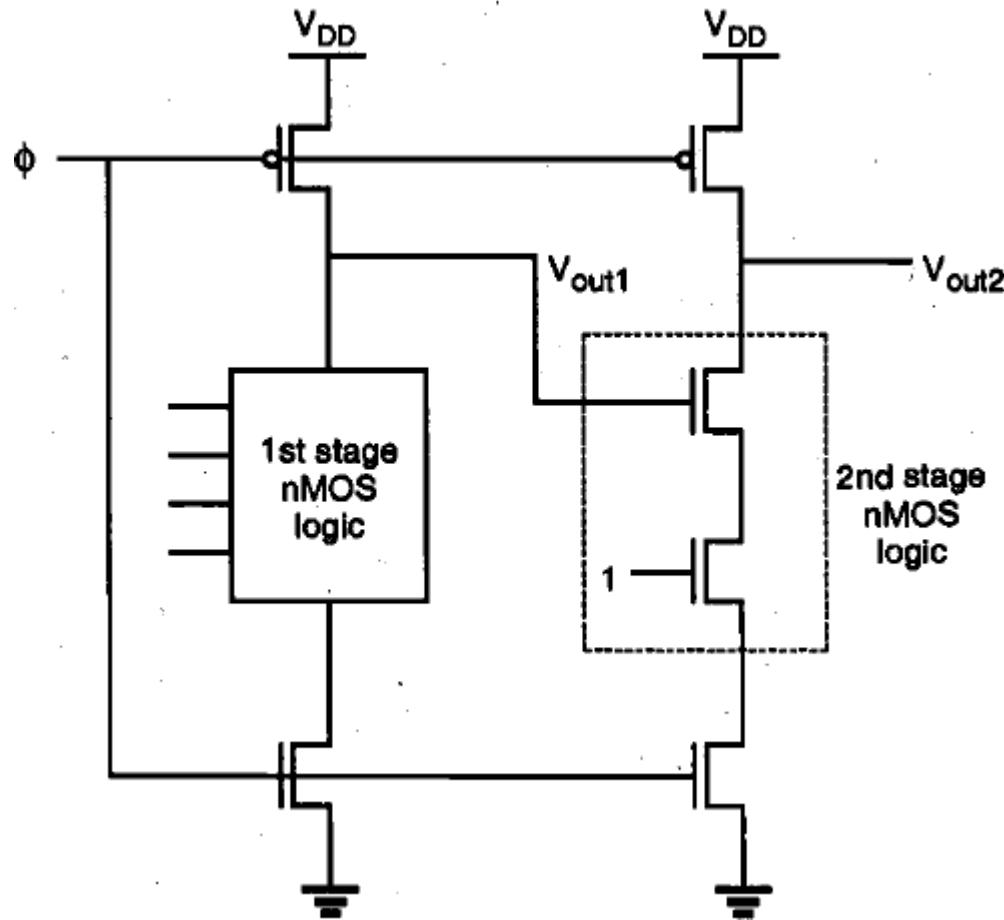
- In **static** circuits the output is connected to either GND or V_{DD} via a **low resistance path**.
 - fan-in of n requires $2n$ (n N-type + n P-type) devices
- **Dynamic** circuits use temporary storage of signal values on the **capacitance of high impedance nodes**.
 - requires on $n + 2$ ($n+1$ N-type + 1 P-type) transistors

Dynamic CMOS Logic

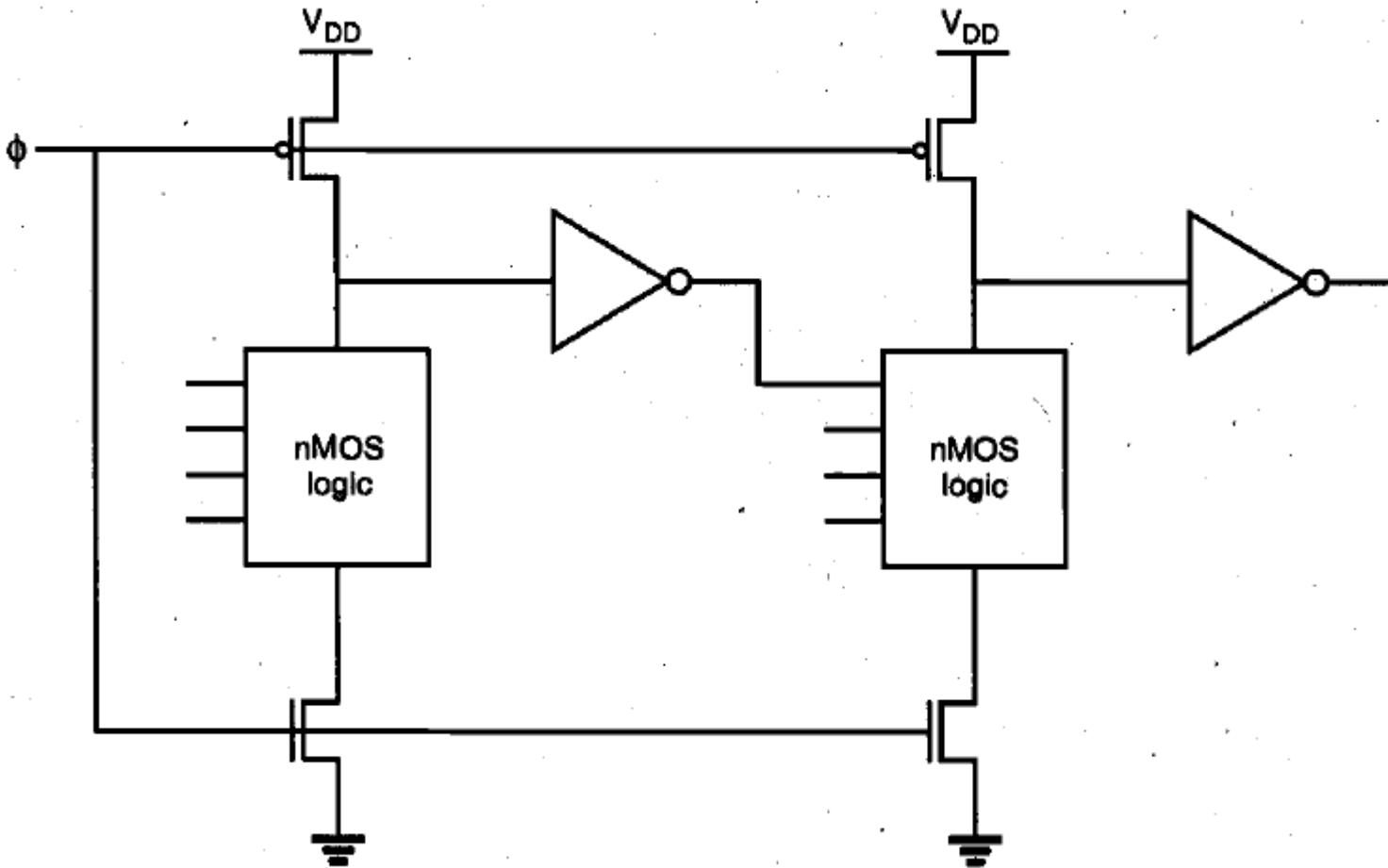
- Once the output of a dynamic gate is discharged, it **cannot be charged** again until the next precharge operation.
- Inputs to the gate can make **at most** one transition during evaluation.
- Output can be in the **high impedance state** during and after evaluation (PDN off), state is stored on C_L

Cascading Problem in Dynamic CMOS Logic

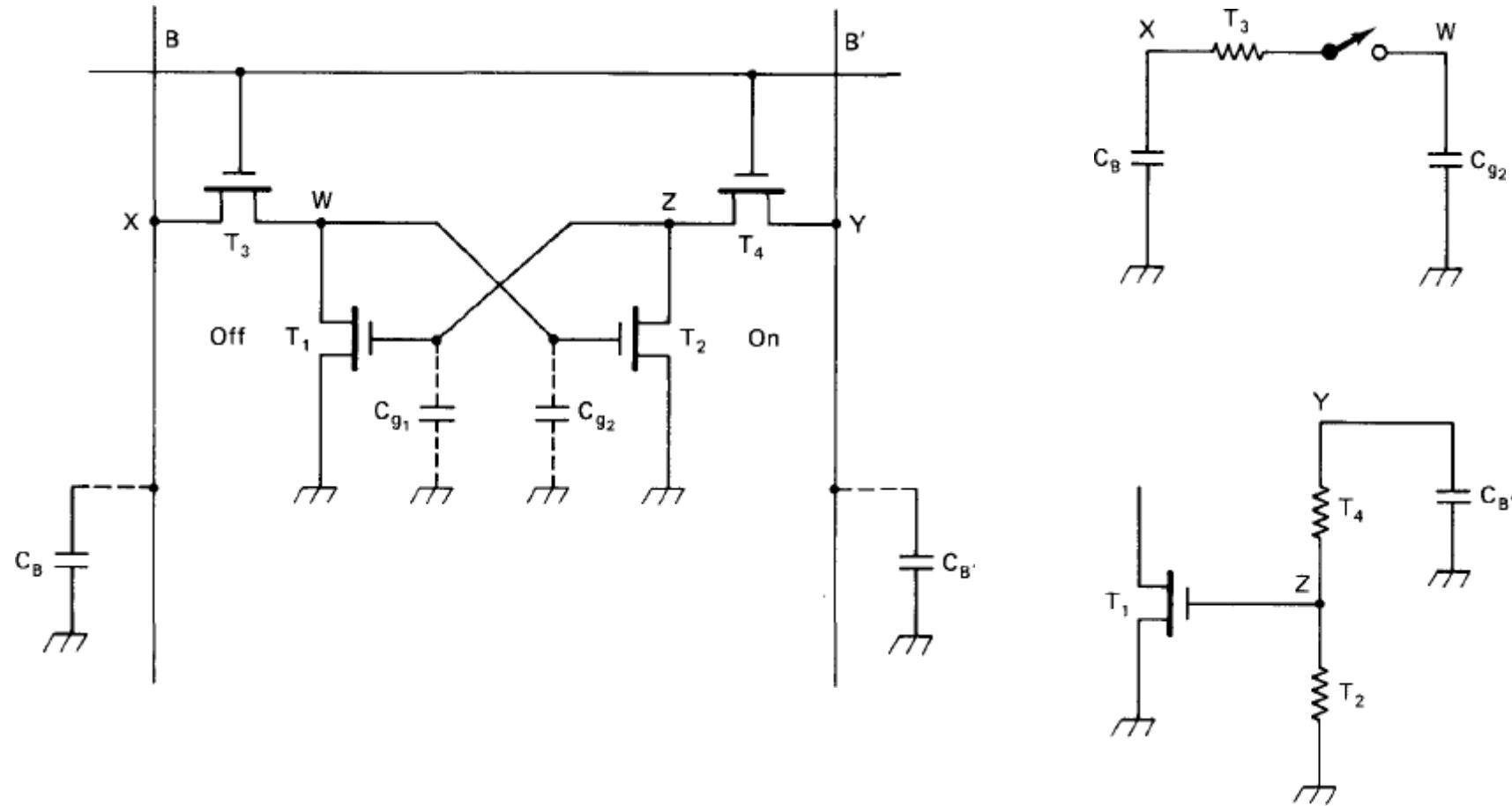
- If several stages of the previous CMOS dynamic logic circuit are cascaded together using the same clock ϕ , a problem in evaluation involving a built-in “race condition” will exist
- Consider the two stage dynamic logic circuit below:
 - During **pre-charge**, both V_{out1} and V_{out2} are pre-charged to V_{dd}
 - When ϕ goes high to begin **evaluate**, all inputs at stage 1 require some finite time to resolve, but during this time charge may erroneously be discharged from V_{out2}
 - e.g. assume that eventually the 1st stage NMOS logic tree conducts and fully discharges V_{out1} , but since all the inputs to the N-tree are not immediately resolved, it takes some time for the N-tree to finally discharge V_{out1} to GND.
 - If, during this time delay, the 2nd stage has the input condition shown with bottom NMOS transistor gate at a logic 1, then V_{out2} will start to fall and discharge its load capacitance until V_{out1} finally evaluates and turns off the top series NMOS transistor in stage 2
 - The result is an error in the output of the 2nd stage V_{out2}



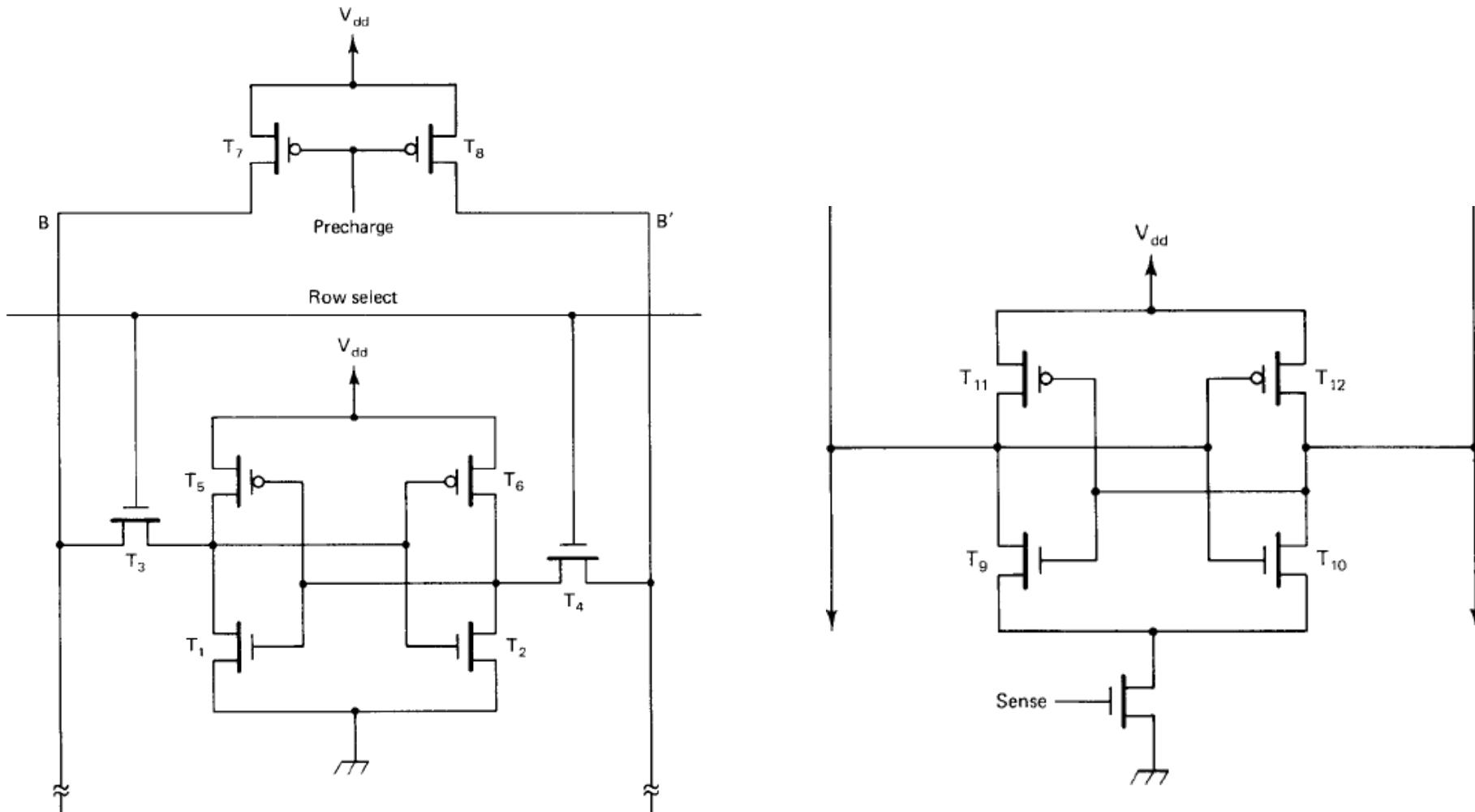
DOMINO CMOS Logic



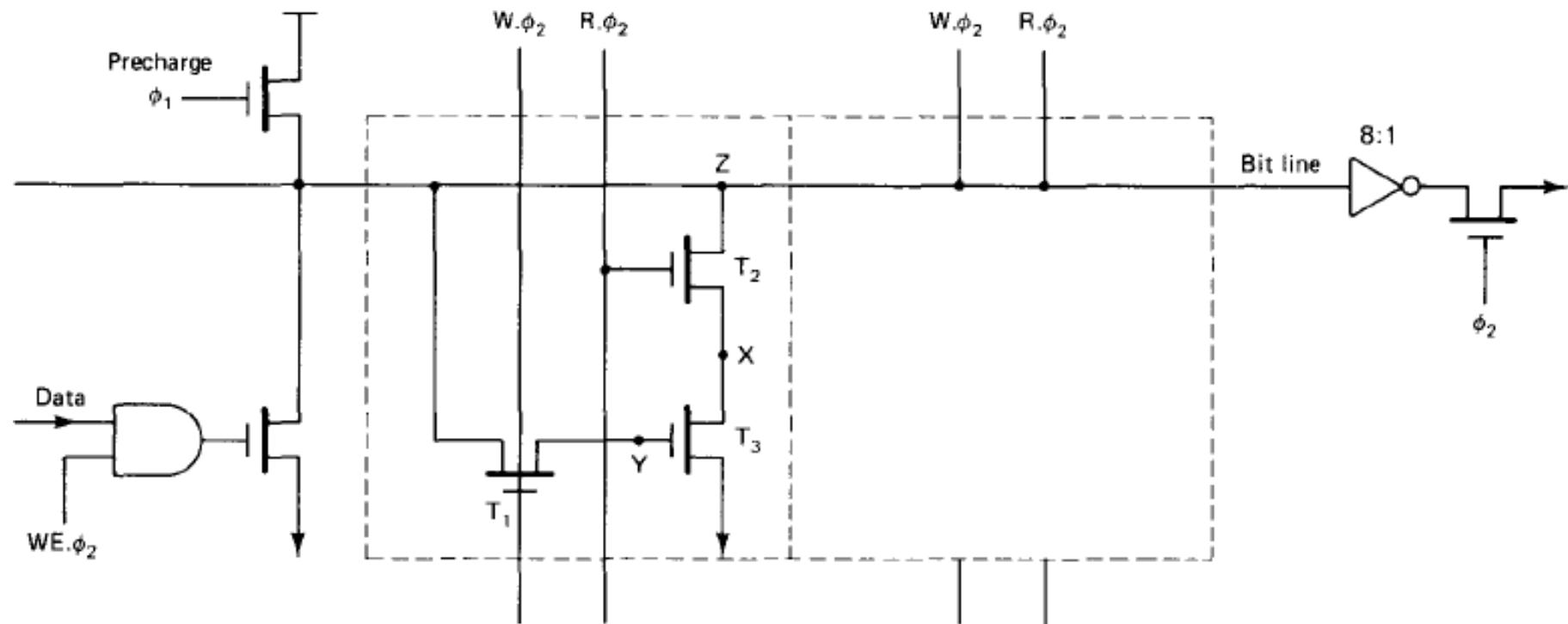
Four-Transistor Dynamic Memory Cell



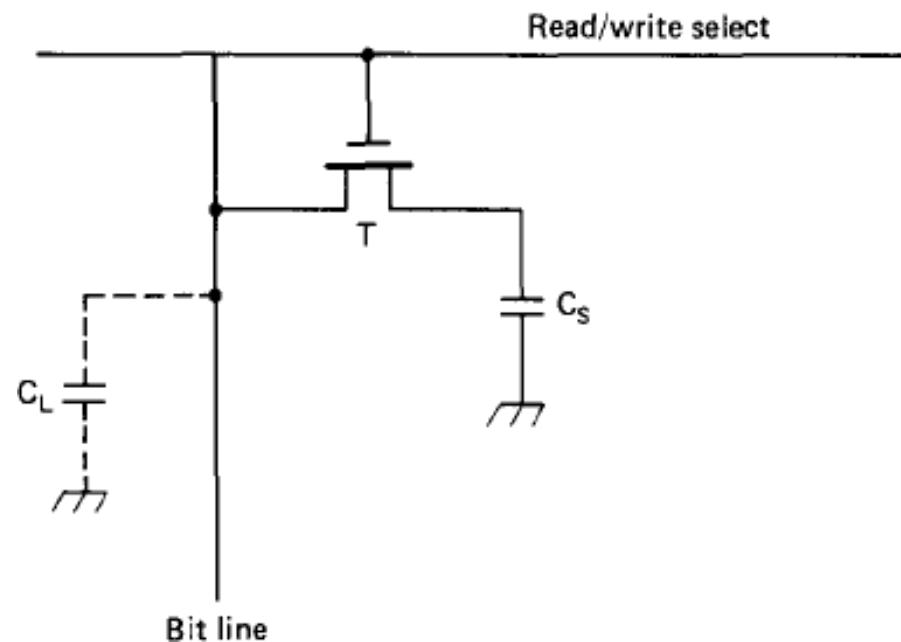
Six-Transistor Static Memory Cell

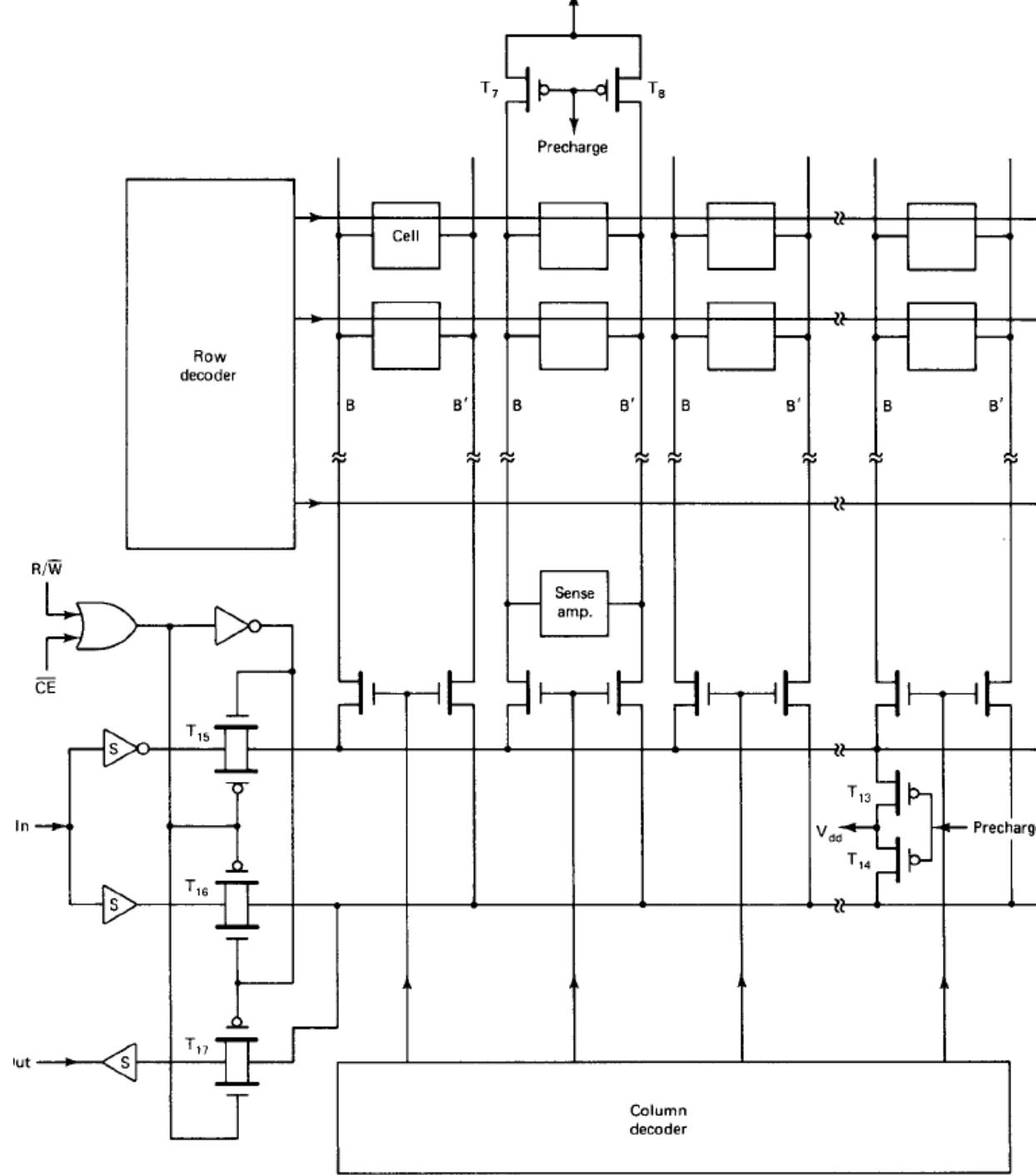


Three-Transistor Dynamic Memory Cell



One-Transistor Dynamic Memory Cell





Random Access Memory(1-T DRAM)



$$V_B = \frac{Q_B}{C_B}$$

$$V_C = \frac{Q_C}{C_C}$$

$$V_F = \frac{Q_B + Q_C}{C_B + C_C}$$

$$\Delta V_B = V_B - V_F$$

$$\Delta V_B = \frac{Q_B}{C_B} - \frac{Q_B + Q_C}{C_B + C_C}$$

$$\Delta V_B = \frac{Q_B(C_B + C_C) - C_B(Q_B + Q_C)}{C_B(C_B + C_C)}$$

$$\Delta V_B = \frac{Q_B \left(1 + \frac{C_C}{C_B}\right) - (Q_B + Q_C)}{(C_B + C_C)}$$

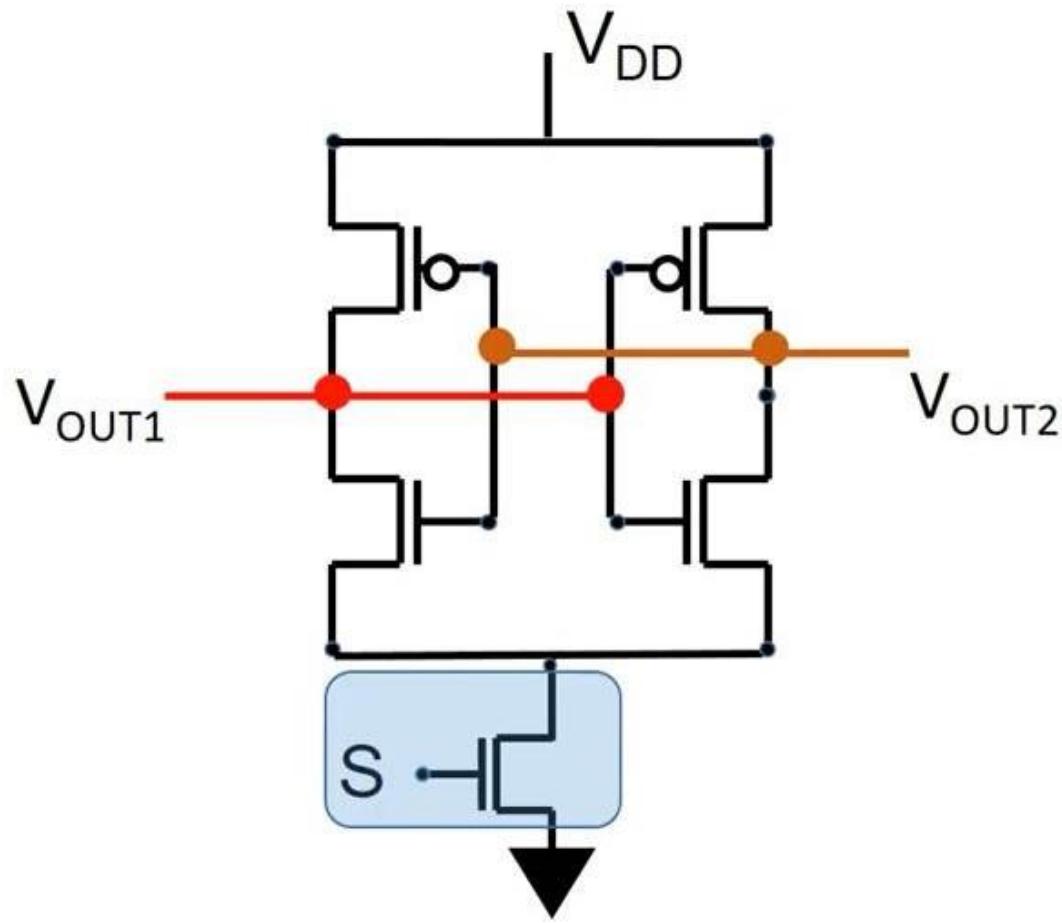
$$\Delta V_B = \frac{C_C(V_B - V_C)}{(C_B + C_C)}$$

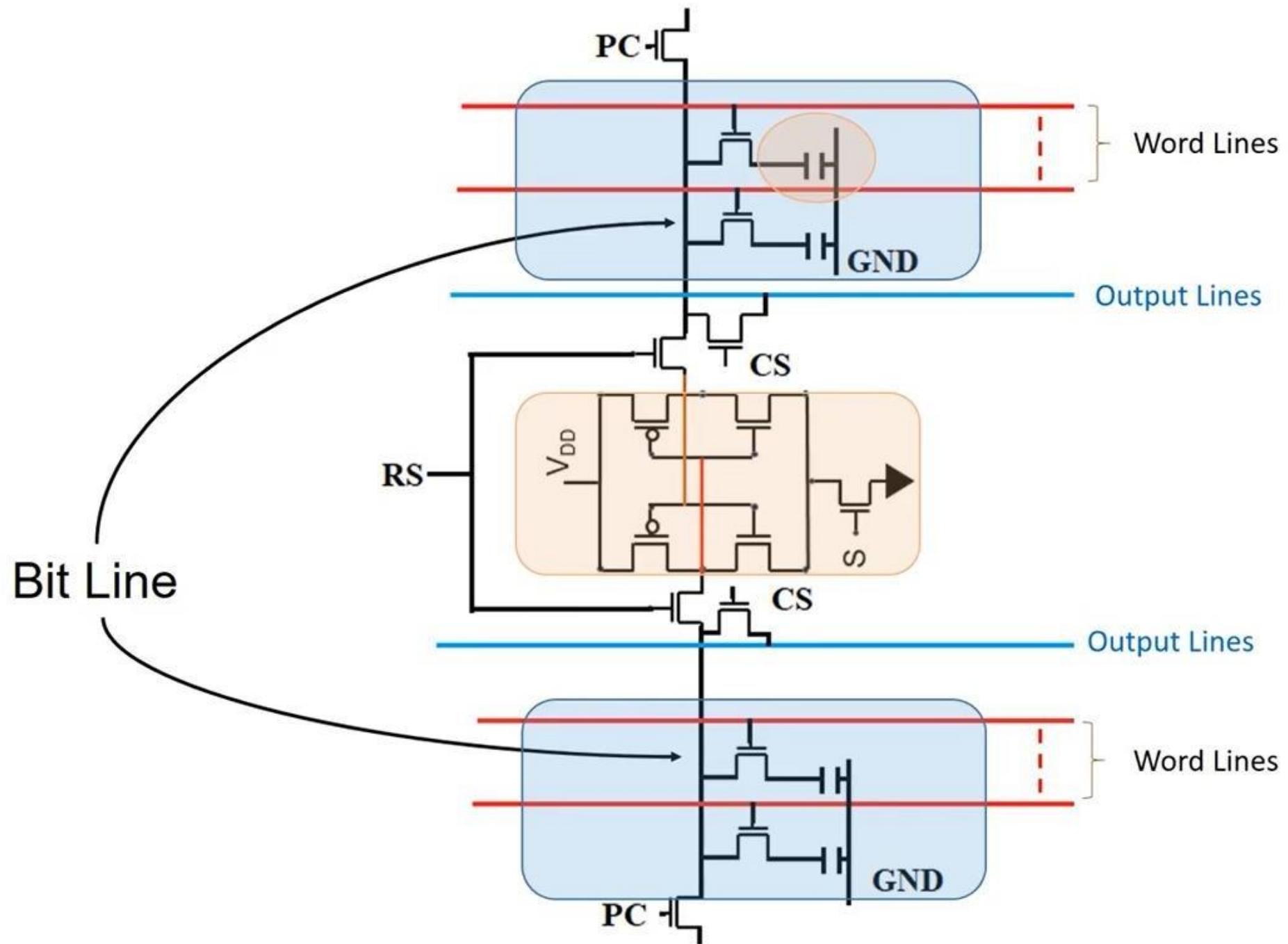
Assume $V_B=2.5V$, $V_C=5V$, $C_B=1pF$, and $C_C=50fF$,

$$\Delta V_B = -119mV$$

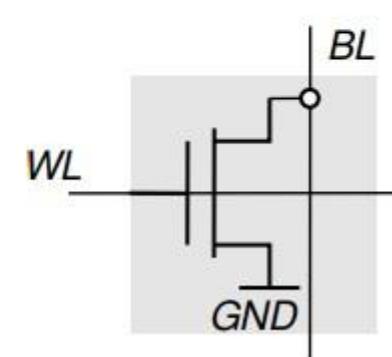
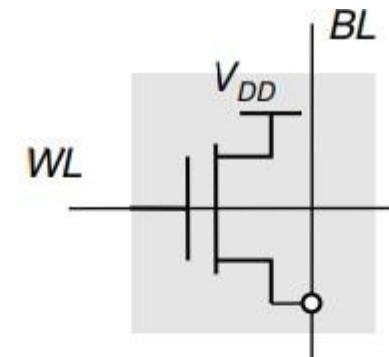
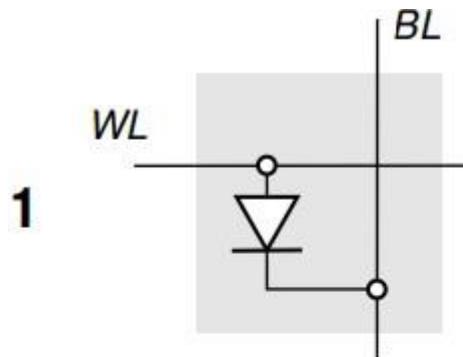


Random Access Memory(1-T DRAM)





Read-Only Memory Cells



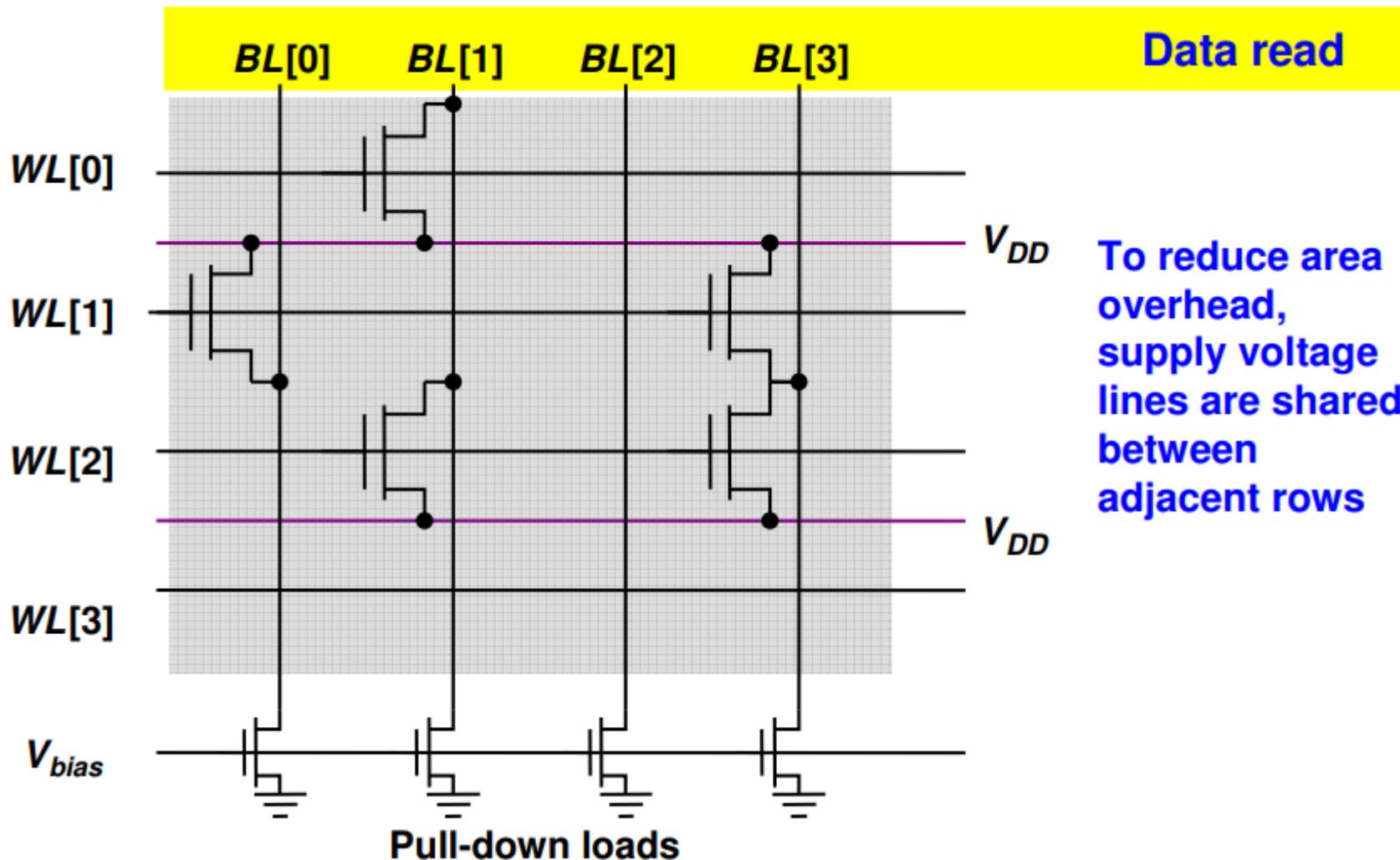
Diode ROM

MOS ROM 1

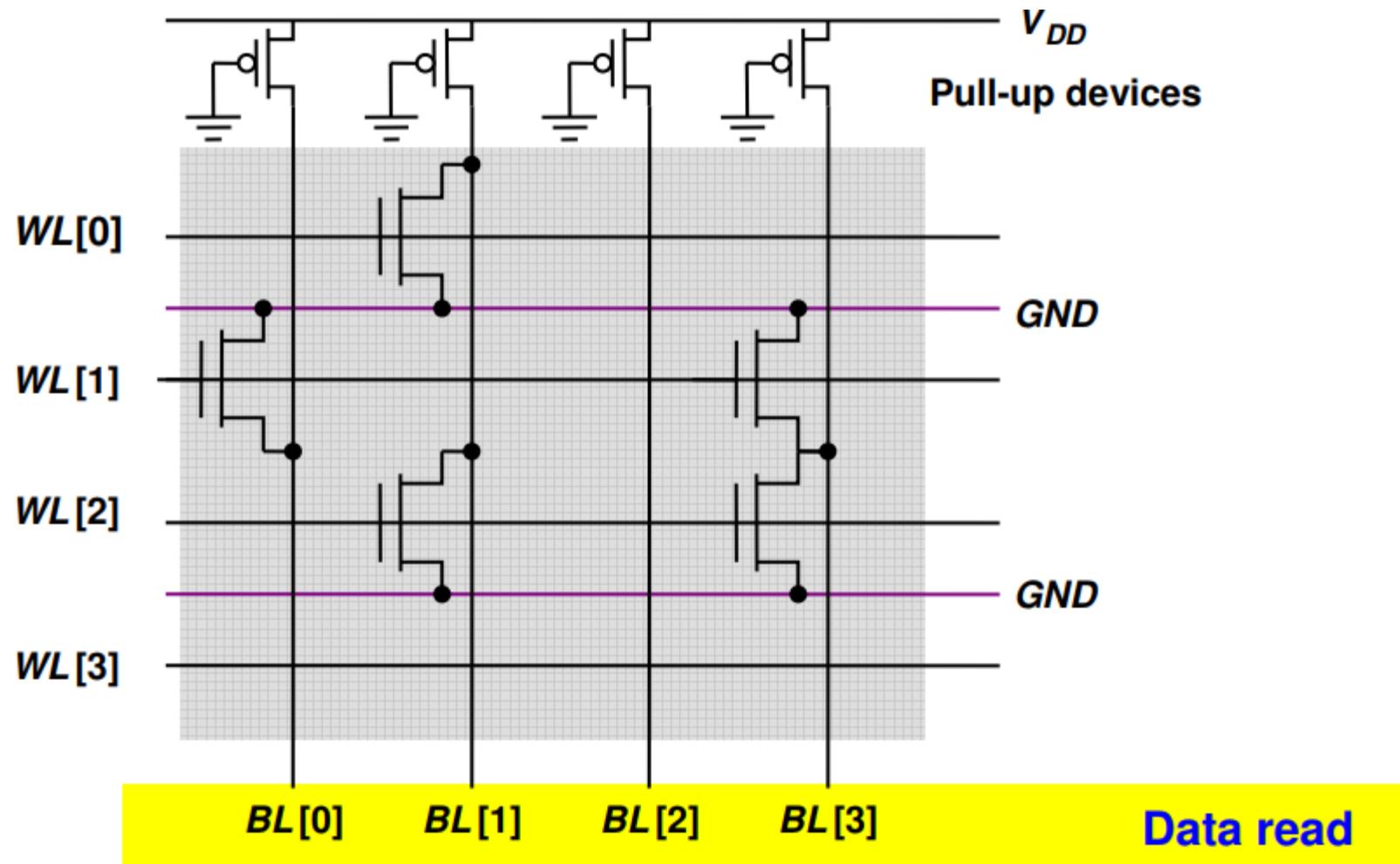
MOS ROM 2

MOS ROM1: BL must be resistively clamped to ground MOS
ROM2: BL must be resistively clamped to Vdd

MOS OR ROM

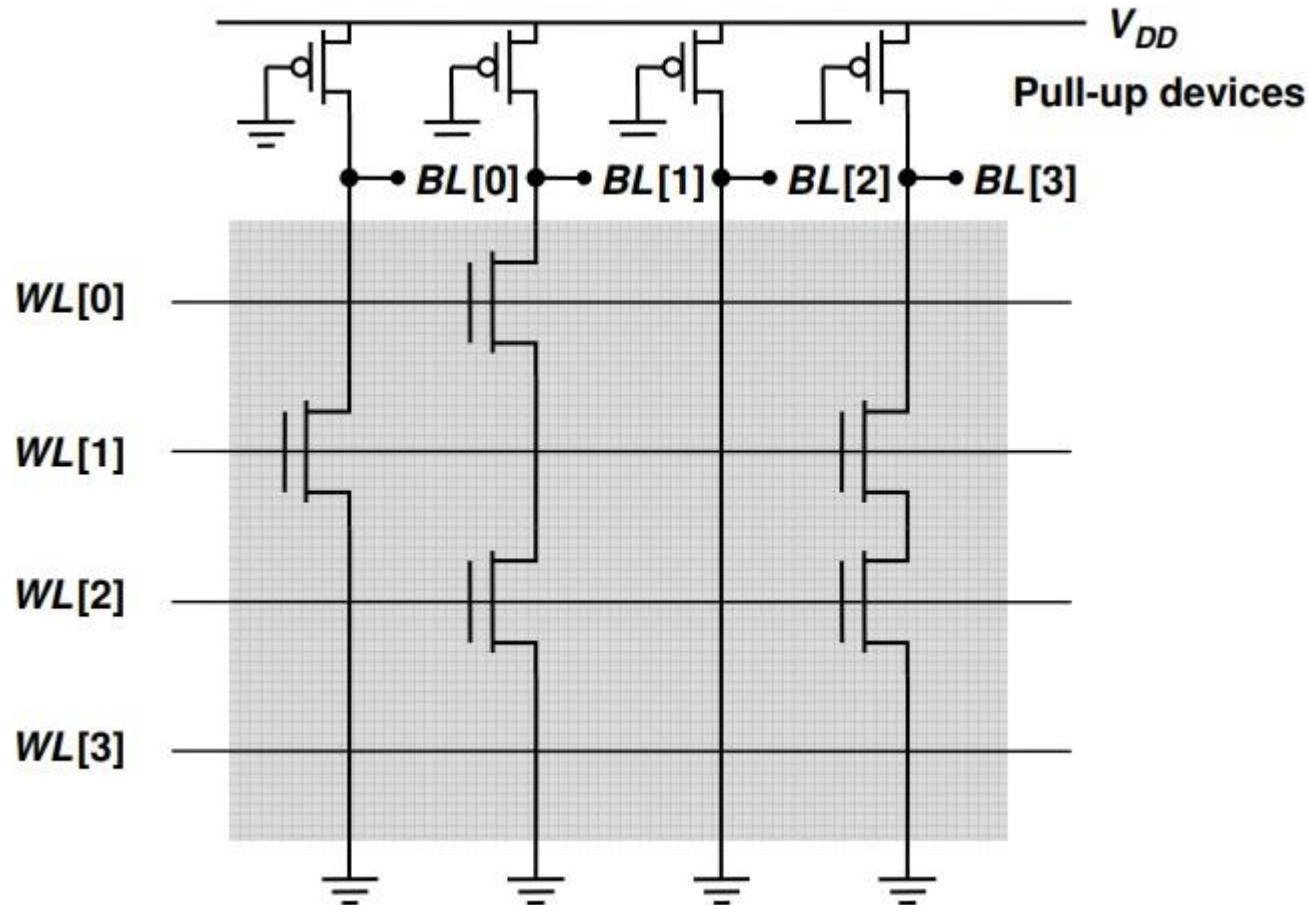


MOS NOR ROM



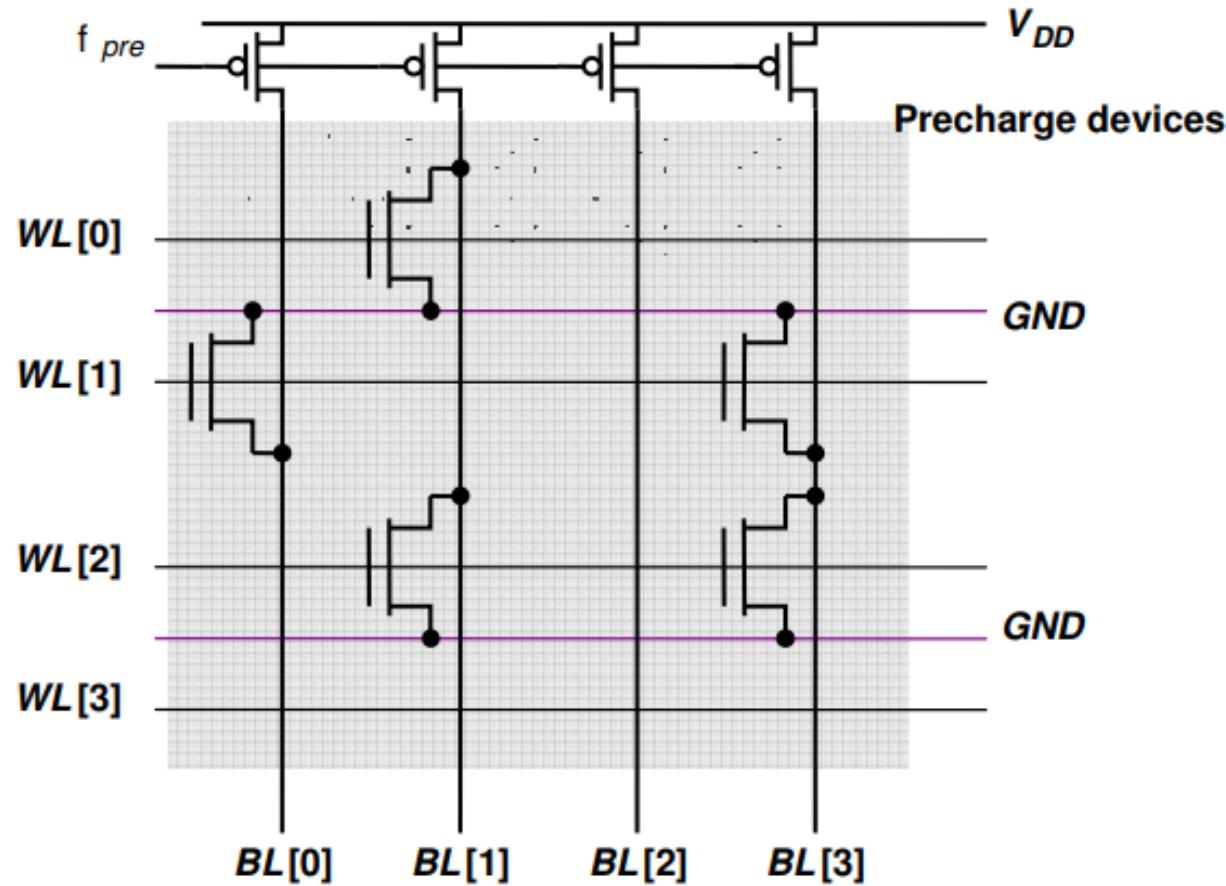
Each row is a pseudo-NMOS (WLs are the inputs)!

MOS NAND ROM



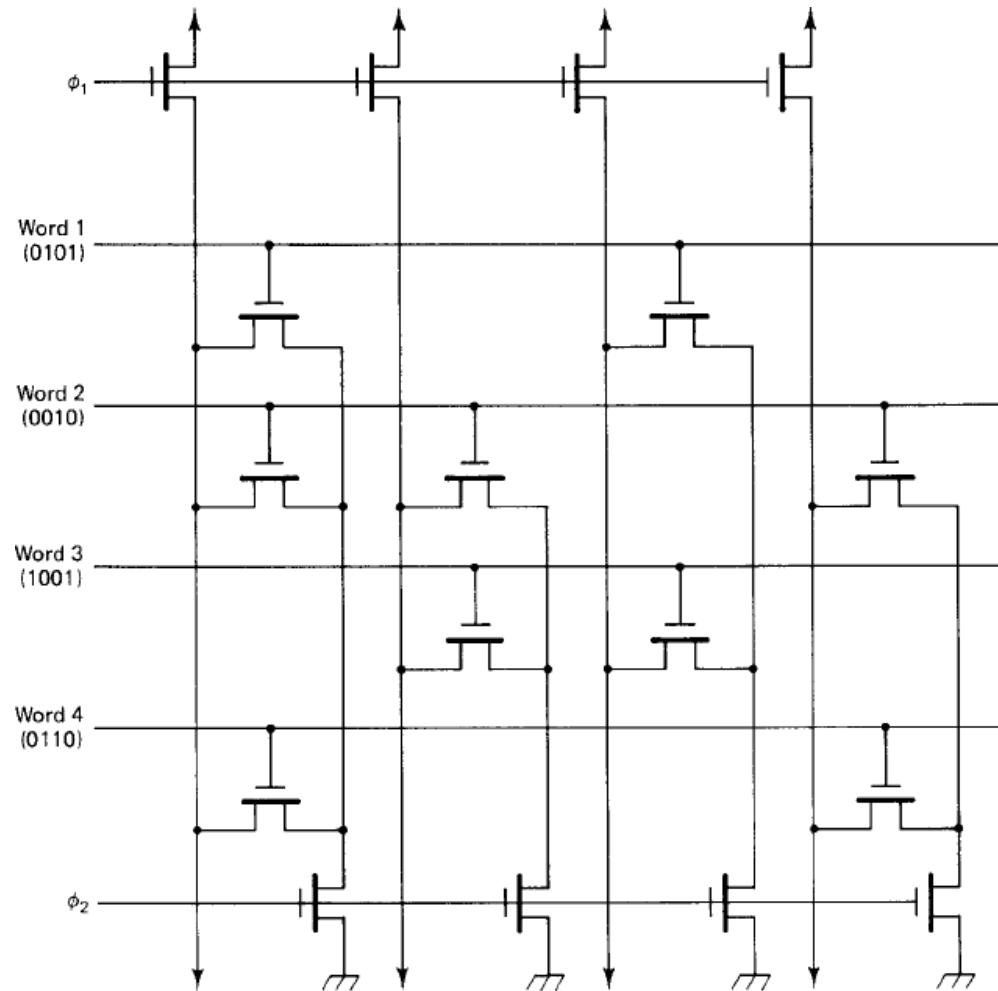
All word lines high by default with exception of selected row

Precharged MOS NOR ROM

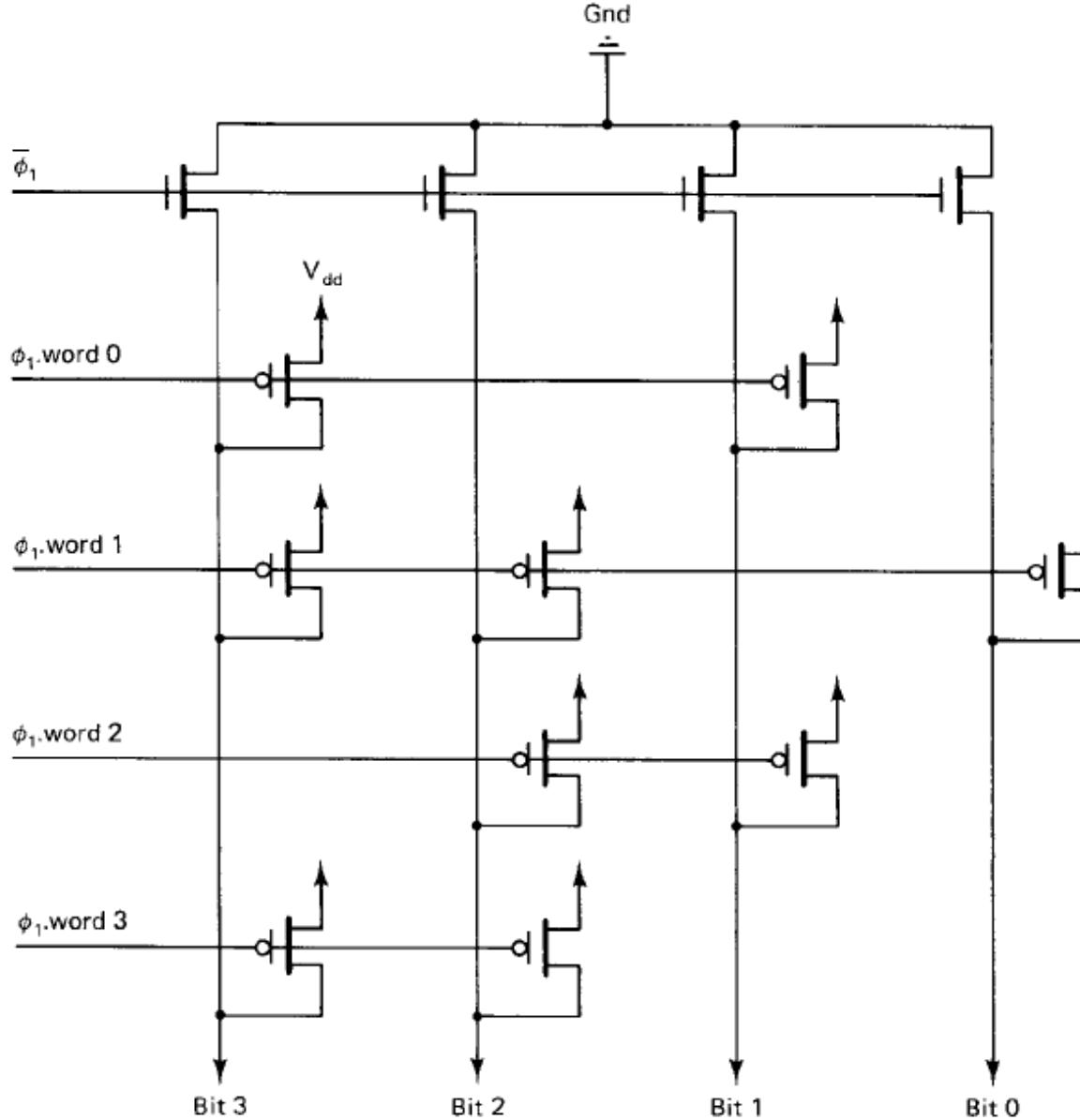


Read-Only Memory

3,9,10,11,12,13,17-19,21-
23,25,27-
29,31,33,36,37,38,39,41-
42,47,49,



An *n*MOS ROM.



Word 0: 1010
 Word 1: 1101
 Word 2: 0110
 Word 3: 1100

Figure 8.12 A CMOS ROM.

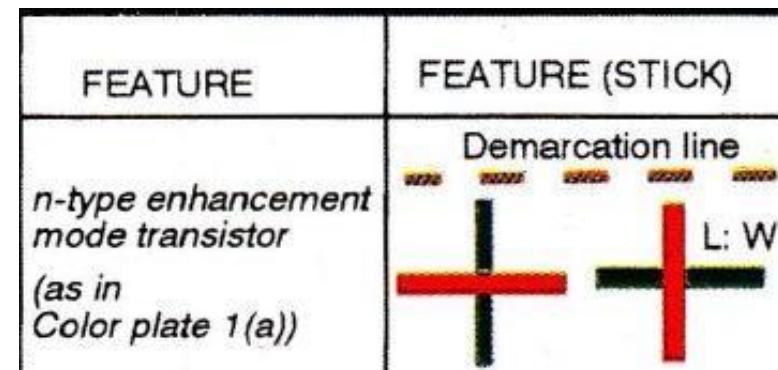
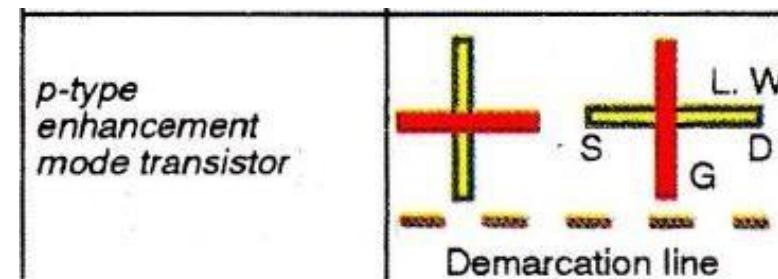
Stick Diagram

COLOR	STICK ENCODING	LAYERS
GREEN		n-diffusion (n ⁺ active) Thinox*
RED		Polysilicon
BLUE		Metal 1
BLACK		Contact cut
GRAY	NOT APPLICABLE	Overglass

nMOS ONLY YELLOW		Implant
nMOS ONLY BROWN		Buried contact

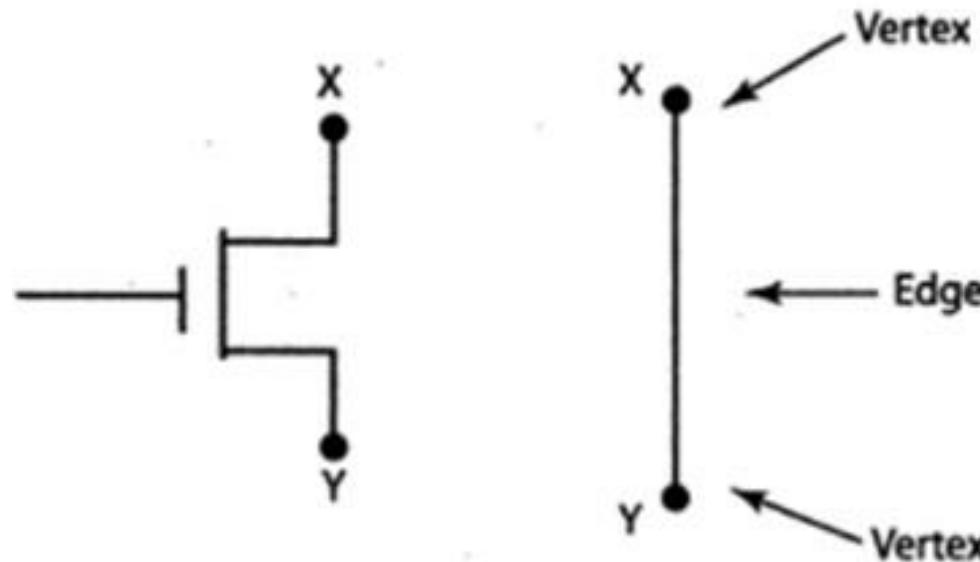
FEATURE	FEATURE (STICK)	n-type depletion mode transistor nMOS only
n-type enhancement mode transistor		

YELLOW (STICK)		p-diffusion (p+ active)
YELLOW	green outline here for clarity Not shown on diagram	p ⁺ mask
DARK BLUE OR PURPLE		Metal 2
BLACK		VIA
BROWN	Demarcation line p-well edge is shown as a demarcation line in stick diagrams	p-well
BLACK		V_{DD} or V_{SS} contact



Euler's Graph

- Diffusion without break
- Common Polysilicon running between P and N diffusion layer
- Less Number of Contacts
- The order of edge visited for NMOS and PMOS must be same

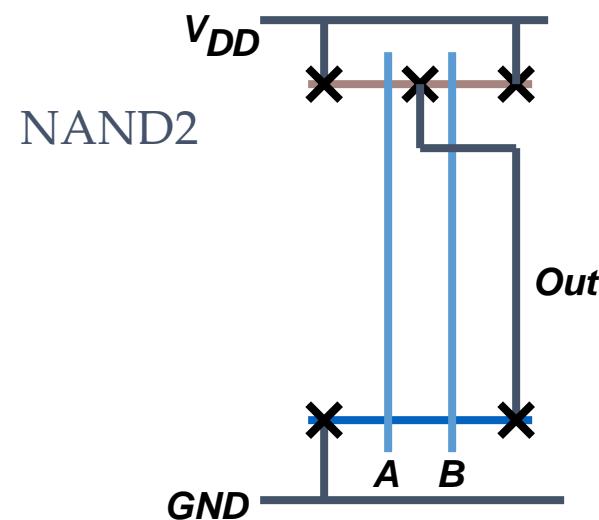
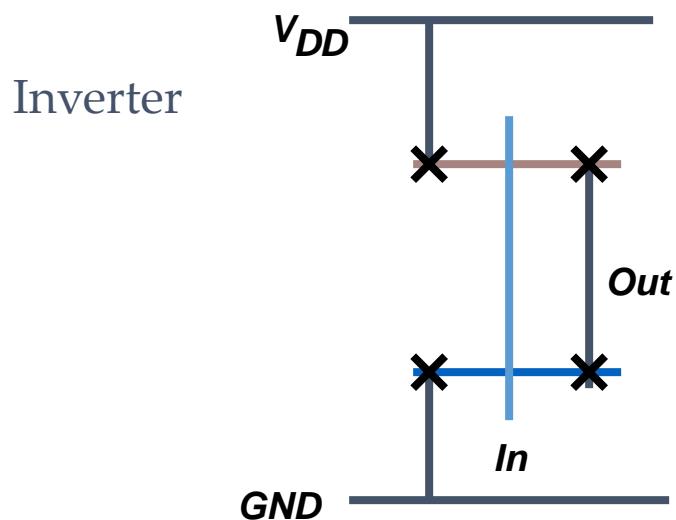


Euler path: a path through all nodes in the graph such that each edge is visited once and only once, Nodes can be repeated.

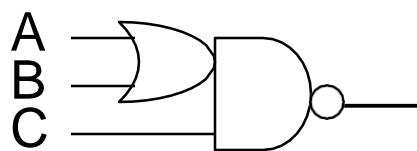
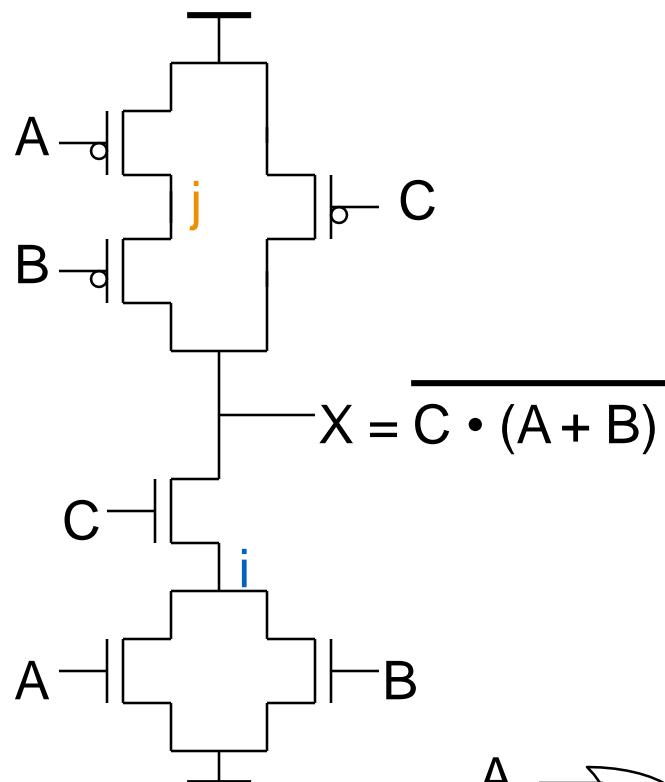
Stick Diagrams

Contains no dimensions

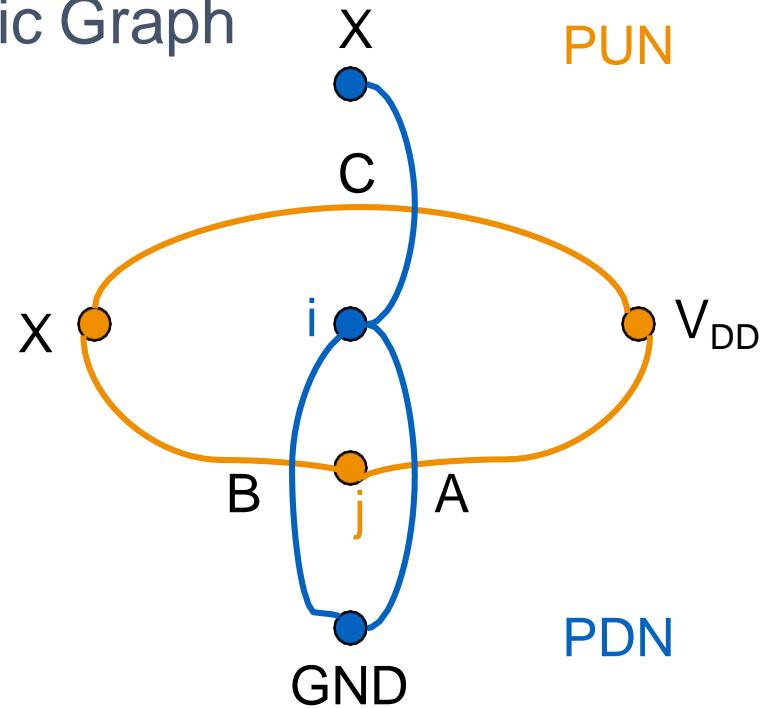
Represents relative positions of transistors



Stick Diagrams



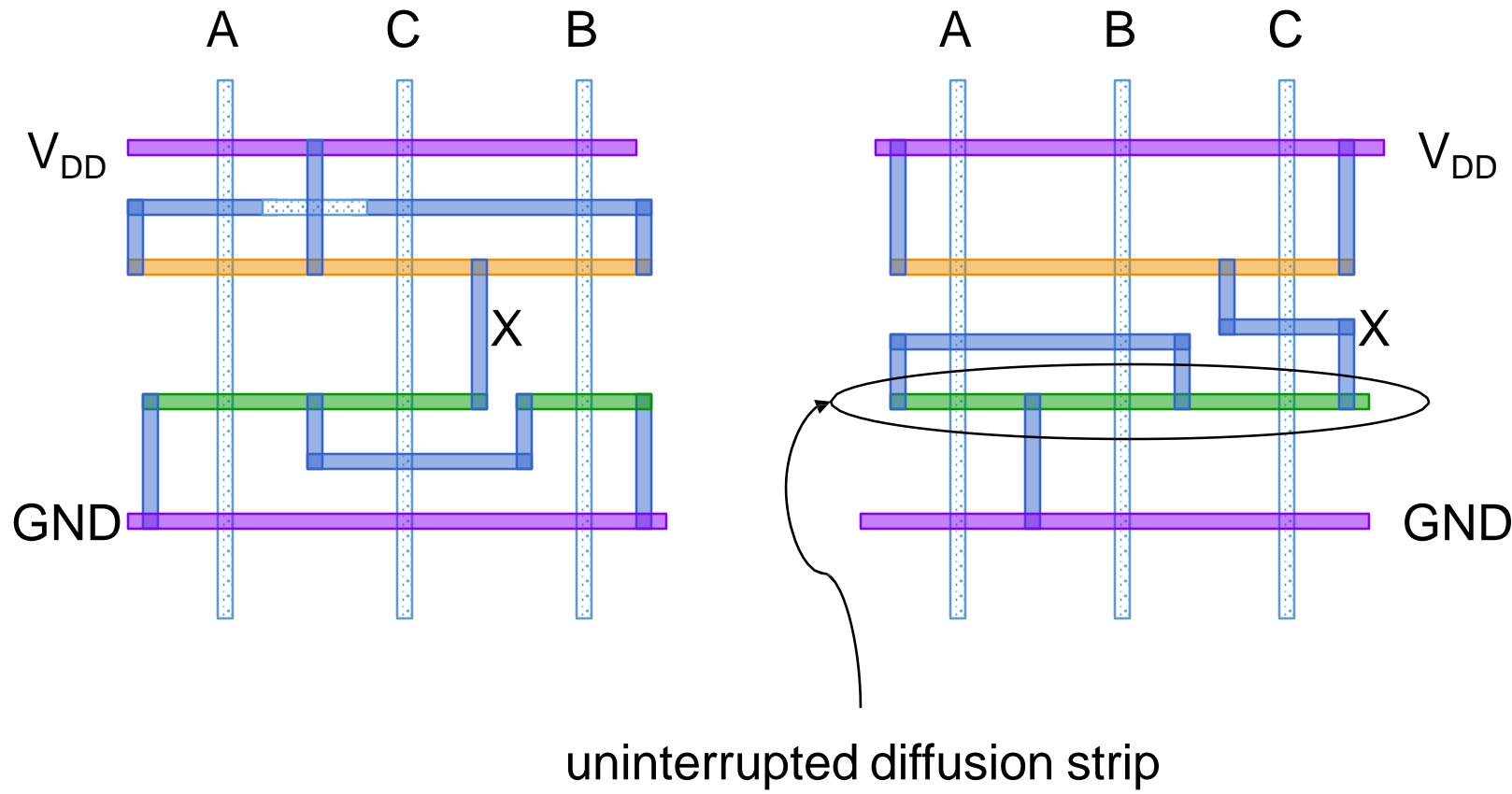
Logic Graph



PUN

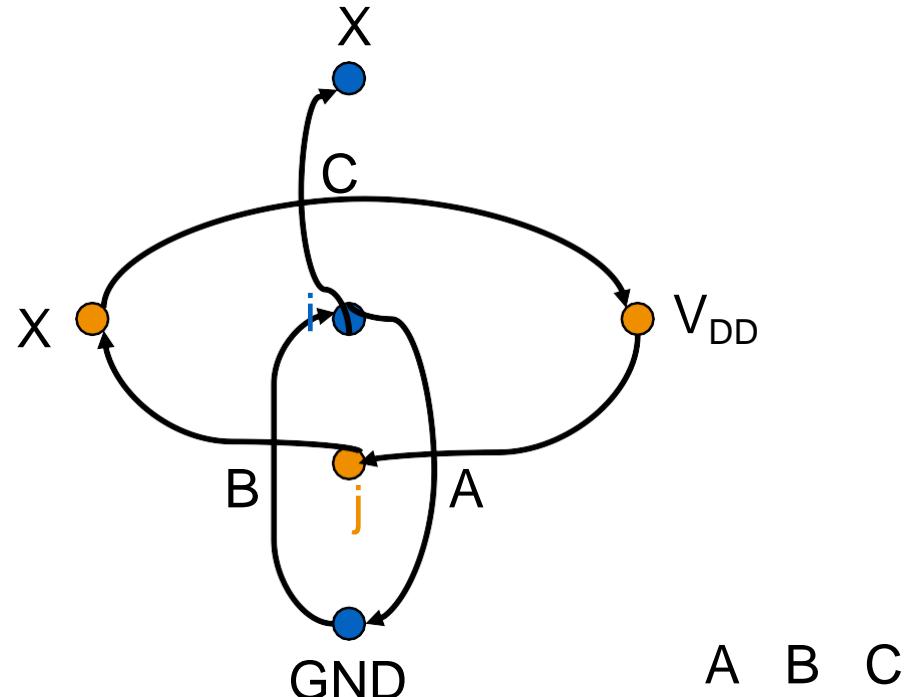
PDN

Two Versions of $C \bullet (A + B)$

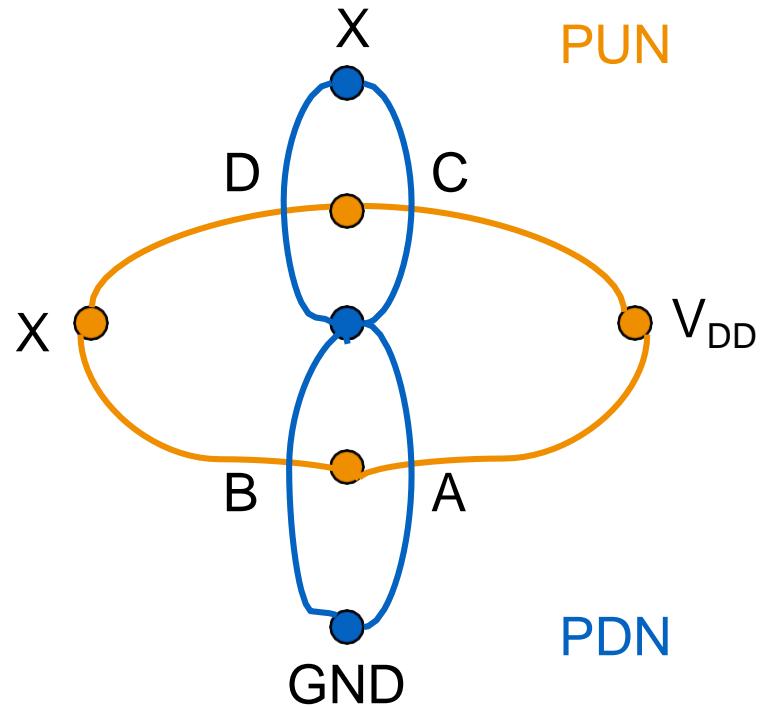
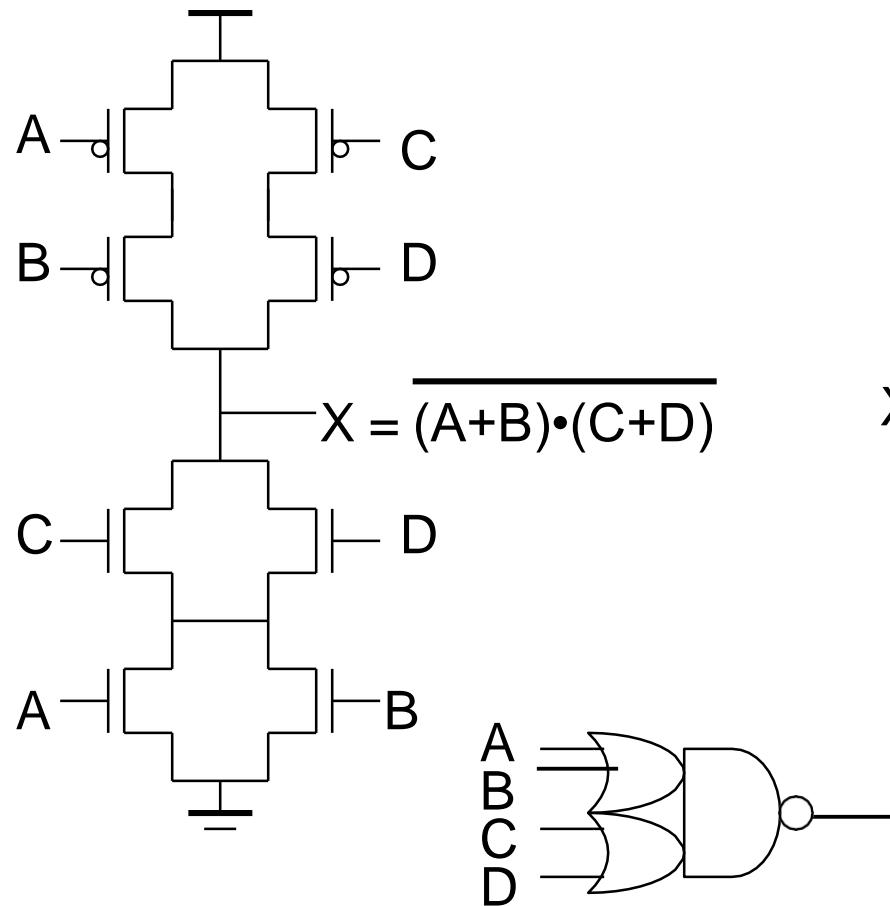


Consistent Euler Path

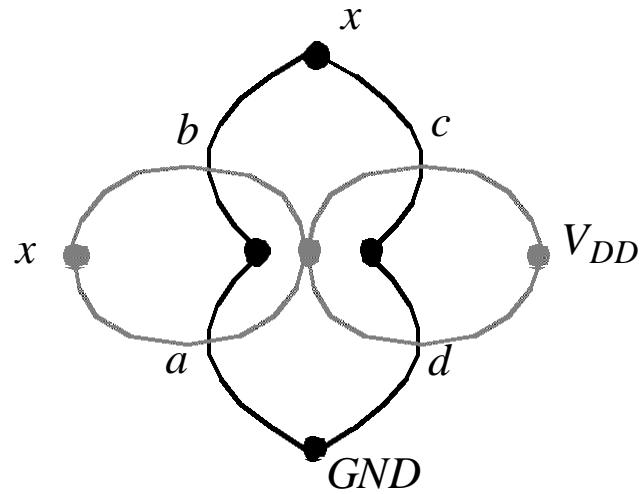
- An uninterrupted diffusion strip is possible only if there exists a Euler path in the logic graph
 - Euler path: a path through all nodes in the graph such that each edge is visited once and only once.



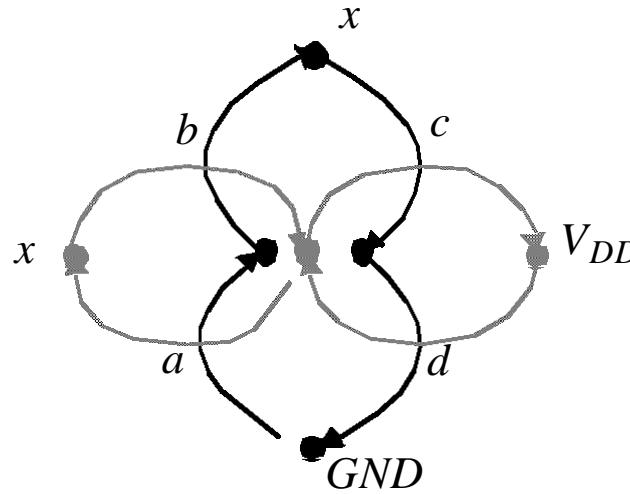
OAI22 Logic Graph



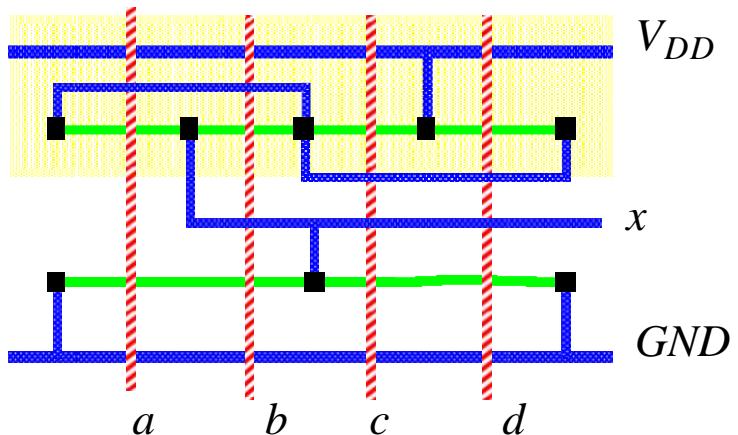
Example: $x = ab+cd$



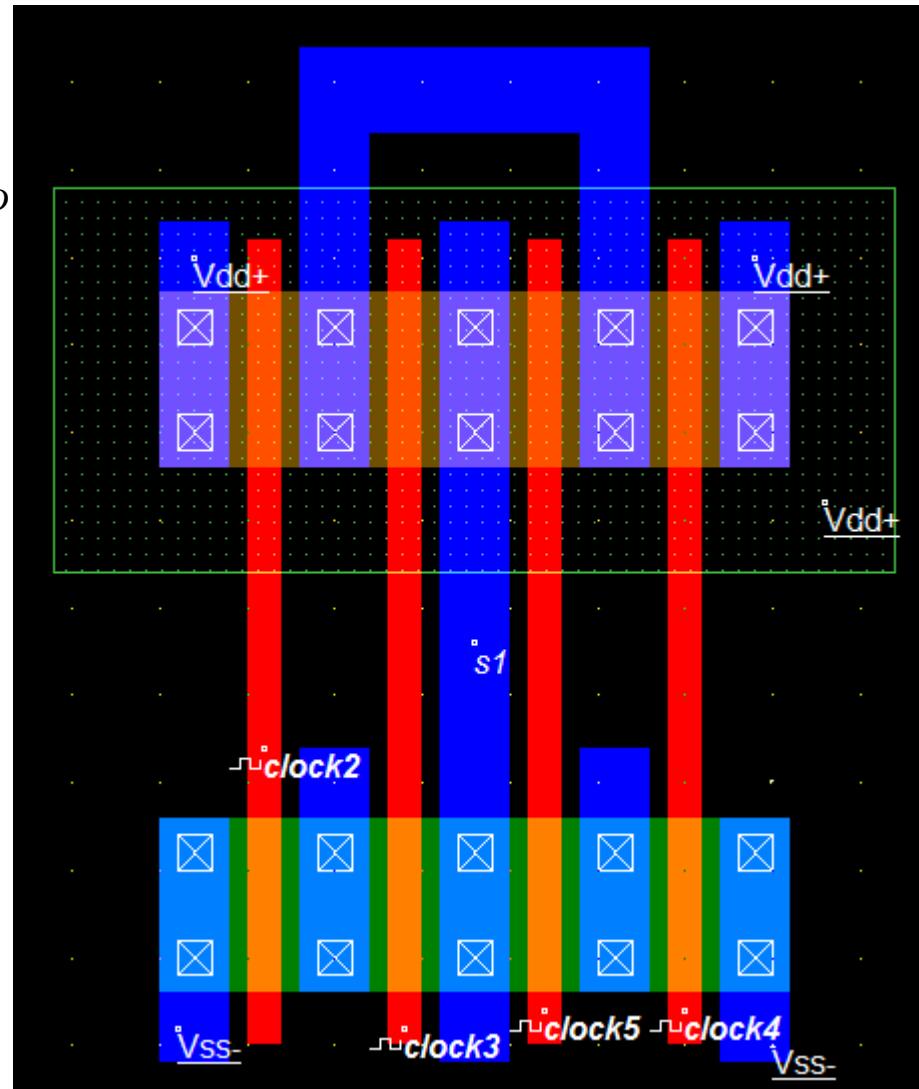
(a) Logic graphs for $(\overline{ab} + \overline{cd})$



(b) Euler Paths $\{a \ b \ c \ d\}$



(c) stick diagram for ordering $\{a \ b \ c \ d\}$



LAYOUT

Scaling

- The process of **reduction of the feature size (L)** and **line widths (W)** is called scaling.
- The effect of scaling must be studied on certain parameters that effect the performance.
- The **figure of merit** of VLSI technology is
 1. Minimum Feature Size
 2. Number of gates on one chip
 3. Power dissipation
 4. Maximum operational frequency
 5. Die size
 6. Production cost

What is expected ?

Complexity $\longrightarrow \infty$

Delay $\longrightarrow \text{Min}$

Power $\longrightarrow \text{Min}$

Cost $\longrightarrow \text{Min}$

Design Time $\longrightarrow \text{Min}$

Size $\longrightarrow \text{Min}$

Scaling Model

There are three types of scaling models used

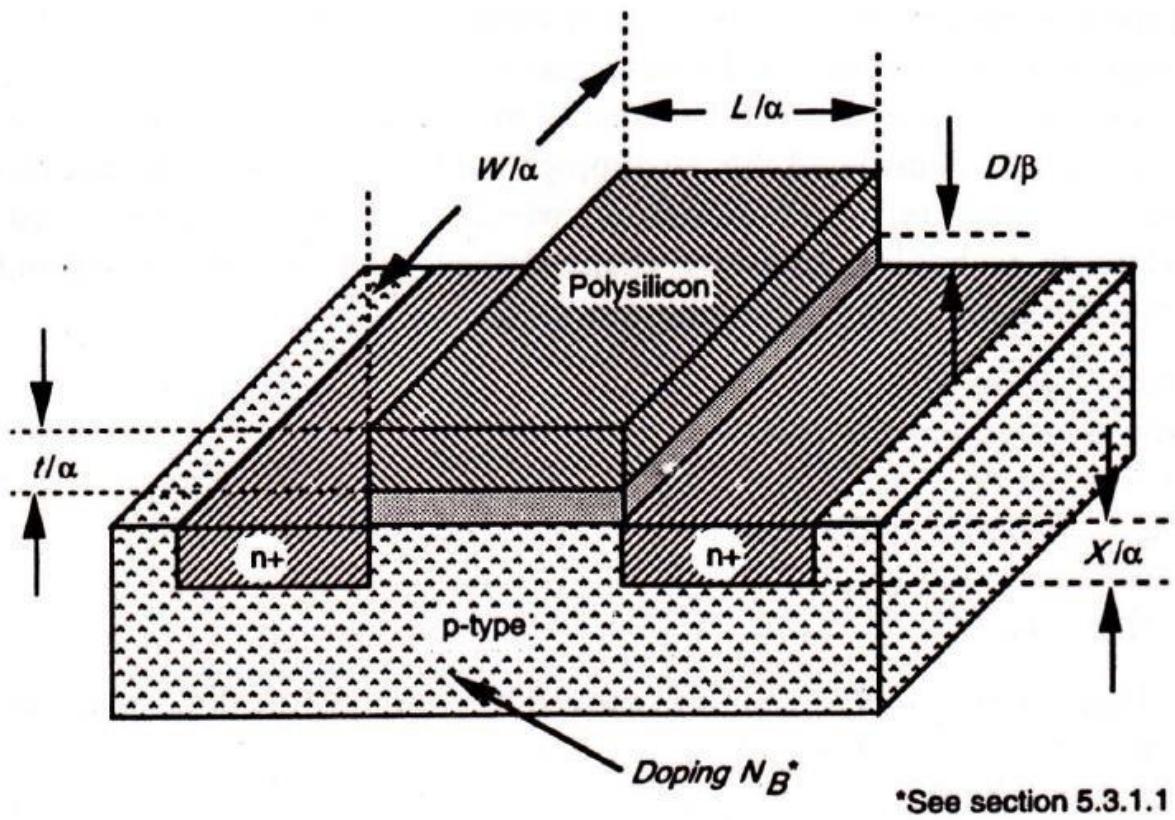
1. Constant electric field (E) scaling model
 2. Constant voltage (V) scaling model
 3. Combined voltage (V) and Dimension (D) scaling model
- 
- Simplified view
and consider only
first order effect.

Scaling Factors

- Here we have considered 2 scaling factors, α and β .
- $1/\beta$ is the scaling factor for supply voltage (**VDD**) and gate oxide thickness (**D**).
- $1/\alpha$ is scaling factor for all other linear dimensions (both vertical and horizontal to chip surface).
- For the constant electrical field model and the constant voltage model, $\beta=\alpha$ and $\beta = 1$ respectively are applied.

- In constant E field scaling model, the model dimensions (L , W) and voltages (V_{dd} , V_{th}) are scaled together by the same scale factor $1/a$.
- In constant V scaling model, only dimensions are scaled by $1/a$, voltages are unscaled.
- In combined V and D scaling model, the model dimensions are scaled by $1/a$ and the voltages are scaled by $1/\beta$.

Scaling



Scaling

Gate Area A_g

$$A_g = L \cdot W.$$

Scaling Factor= $1/\alpha^2$

Gate Capacitance Per Unit Area C_0 or C_{ox}

$$C_0 = \frac{\epsilon_{ox}}{D}$$

$$\text{Scaling Factor} = \frac{1}{1/\beta} = \beta$$

Scaling Factor of constant is 1

Scaling

Gate Capacitance C_g

$$C_g = C_0 L \cdot W.$$

Thus C_g is scaled by $\beta \frac{1}{\alpha^2} = \frac{\beta}{\alpha^2}$

Parasitic Capacitance C_x

C_x is proportional to $\frac{A_x}{d}$

where d is the depletion width around source or drain which is scaled by $1/\alpha$, and A_x is the area of the depletion region around source or drain which is scaled by $1/\alpha^2$.

Thus C_x is scaled by $\frac{1}{\alpha^2} \cdot \frac{1}{1/\alpha} = \frac{1}{\alpha}$

Carrier Density in Channel Q_{on}

$$Q_{on} = C_0 \cdot V_{gs}$$

Thus Q_{on} is scaled by 1

Channel Resistance R_{on}

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

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

Scaling Factor of constant is 1

Gate Delay T_d

T_d is proportional to $R_{on} \cdot C_g$

Thus T_d is scaled by $\frac{1 \cdot \beta}{\alpha^2} = \frac{\beta}{\alpha^2}$

Maximum Operating Frequency f_0

$$f_0 = \frac{W}{L} \frac{\mu C_0 V_{DD}}{C_g}$$

Thus f_0 is scaled by $\frac{1}{\beta/\alpha^2} = \frac{\alpha^2}{\beta}$

Saturation Current I_{dss}

$$I_{dss} = \frac{C_0 \mu}{2} \frac{W}{L} (V_{gs} - V_t)^2$$

noting that both V_{gs} and V_t are scaled by $1/\beta$, we have

I_{dss} is scaled by $\beta(1/\beta)^2 = 1/\beta$

Current Density J

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

J is scaled by $\frac{1/\beta}{1/\alpha^2} = \frac{\alpha^2}{\beta}$

Switching Energy Per Gate E_g

$$E_g = \frac{1}{2} C_g (V_{DD})^2$$

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

Power Dissipation Per Gate P_g

P_g comprises two components such that

$$P_g = P_{gs} + P_{gd}$$

where the static component

$$P_{gs} = \frac{(V_{DD})^2}{R_{on}}$$

and the dynamic component

$$P_{gd} = E_g f_0$$

It will be seen that both P_{gs} and P_{gd} are scaled by $1/\beta^2$

So, P_g is scaled by $1/\beta^2$

Power Dissipation Per Unit Area P_a

$$P_a = \frac{P_g}{A_g}$$

P_a is scaled by $\frac{1/\beta^2}{1/\alpha^2} = \alpha^2/\beta^2$

Power-speed Product P_T

$$P_T = P_g \cdot T_d$$

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

<i>Parameters</i>		<i>Combined V and D</i>	<i>Constant E</i>	<i>Constant V</i>
V_{DD}	Supply voltage	$1/\beta$	$1/\alpha$	1
L	Channel length	$1/\alpha$	$1/\alpha$	$1/\alpha$
W	Channel width	$1/\alpha$	$1/\alpha$	$1/\alpha$
D	Gate oxide thickness	$1/\beta$	$1/\alpha$	1
A_g	Gate area	$1/\alpha^2$	$1/\alpha^2$	$1/\alpha^2$
C_0 (or C_{ox})	Gate C per unit area	β	α	1
C_g	Gate capacitance	β/α^2	$1/\alpha$	$1/\alpha^2$
C_x	Parasitic capacitance	$1/\alpha$	$1/\alpha$	$1/\alpha$
Q_{on}	Carrier density	1	1	1
R_{on}	Channel resistance	1	1	1
I_{dss}	Saturation current	$1/\beta$	$1/\alpha$	1
A_c	Conductor X-section area	$1/\alpha^2$	$1/\alpha^2$	$1/\alpha^2$
I	Current density	α^2/β	α	α^2
V_g	Logic 1 level	$1/\beta$	$1/\alpha$	1
E_g	Switching energy	$1/\alpha^2 \cdot \beta$	$1/\alpha^3$	$1/\alpha^2$
P_g	Power dispn per gate	$1/\beta^2$	$1/\alpha^2$	1
N	Gates per unit area	α^2	α^2	α^2
P_a	Power dispn per unit area	α^2/β^2	1	α^2
T_d	Gate delay	β/α^2	$1/\alpha$	$1/\alpha^2$
f_0	Max. operating frequency	α^2/β	α	α^2
P_T	Power-speed product	$1/\alpha^2 \cdot \beta$	$1/\alpha^3$	$1/\alpha^2$

Constant E: $\beta = \alpha$; Constant V: $\beta = 1$

BASIC CIRCUIT CONCEPTS

Concepts such as sheet resistance R_s , and a standard unit of capacitance $\square Cg$, help greatly in evaluating the effects of wiring and input and output capacitances. Further, the delays associated with wiring, with inverters and with other circuitry may be conveniently evaluated in terms of a delay unit τ .

RESISTANCE ESTIMATION

The concept of sheet resistance is being used to know the resistive behavior of the layers that go into formation of the MOS device. Let us consider a uniform slab of conducting material of the following characteristics .

Resistivity- ρ

Width - W

Thickness - t

Length between faces – L as shown next

Sheet Resistance

$$R_{AB} = \frac{\rho L}{A} \text{ ohm}$$

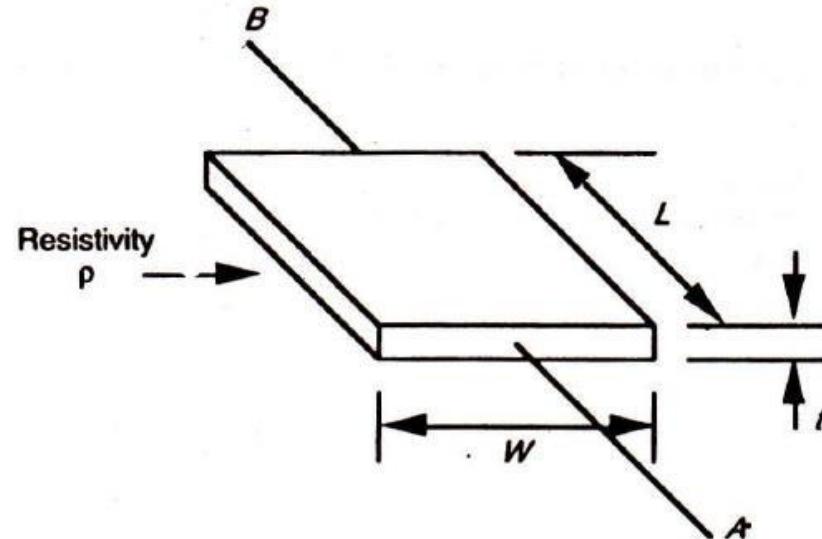
A = cross-section area

$$R_{AB} = \frac{\rho L}{tW} \text{ ohm}$$

If $L=W$ then

$$R_{AB} = \frac{\rho}{t} = R_s$$

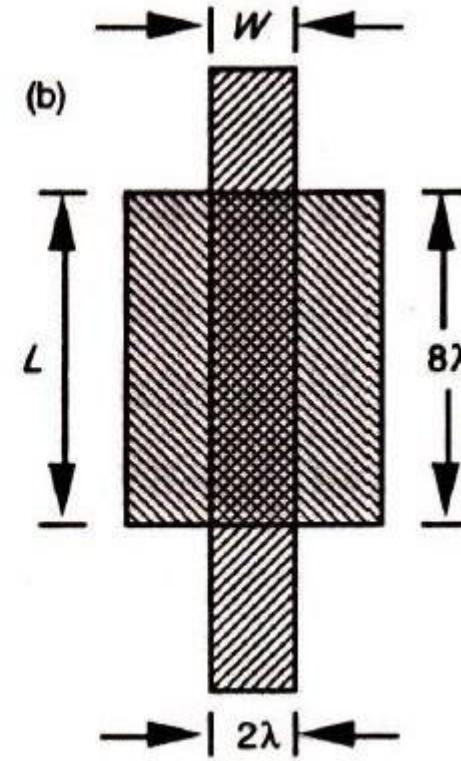
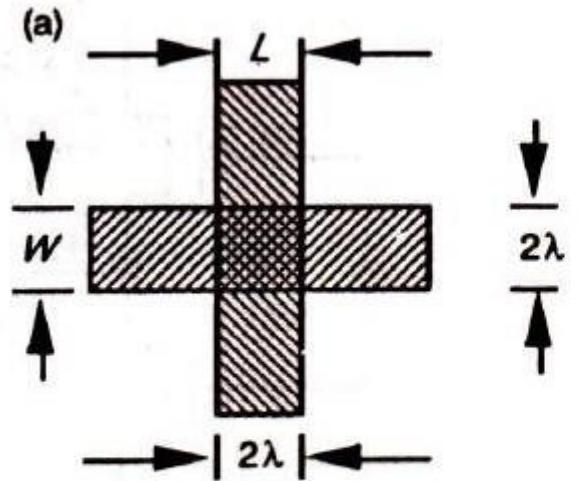
R_s = ohm per square or sheet resistance



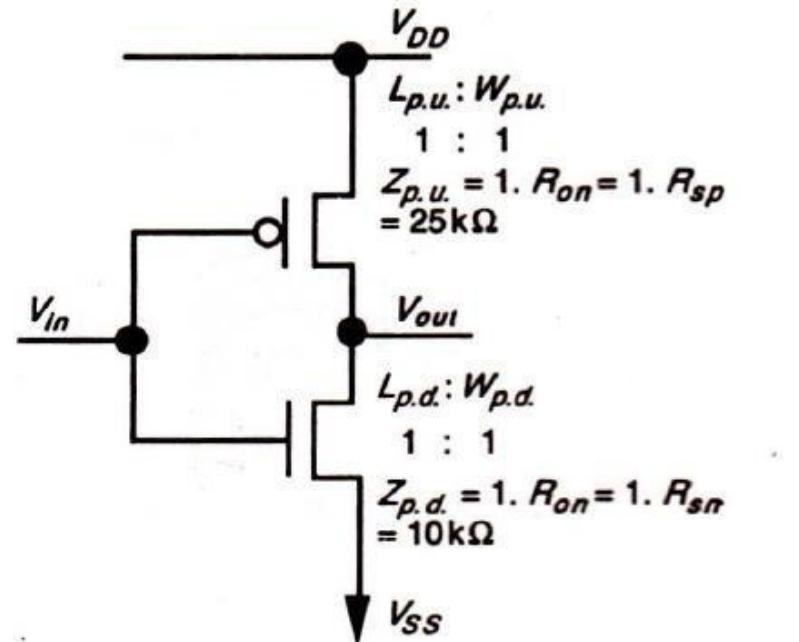
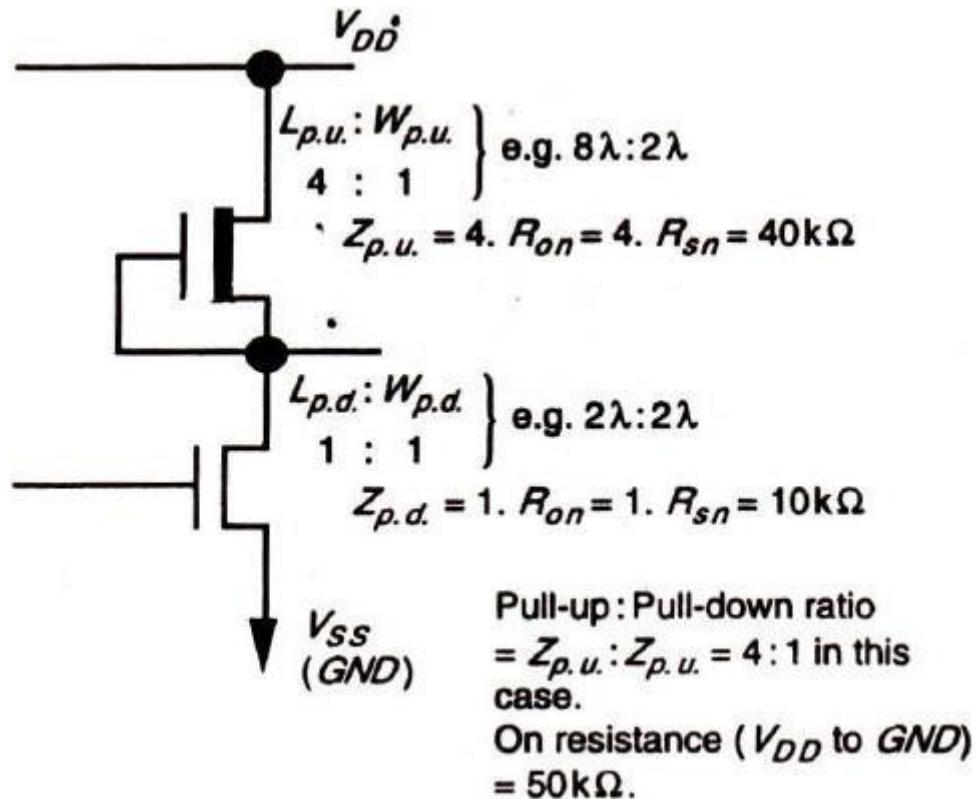
Sheet Resistance

Layer	<i>R_s ohm per square</i>		
	5 μm	2μm Orbit	Orbit 1.2 μm
Metal	0.03	0.04	0.04
Diffusion (or active)**	10→50	20→45	20→45
Silicide	2→4	—	—
Polysilicon	15→100	15→30	15→30
n-transistor channel	10 ^{4†}	2 × 10 ^{4†}	2 × 10 ^{4†}
p-transistor channel	2.5 × 10 ^{4†}	4.5 × 10 ^{4†}	4.5 × 10 ^{4†}

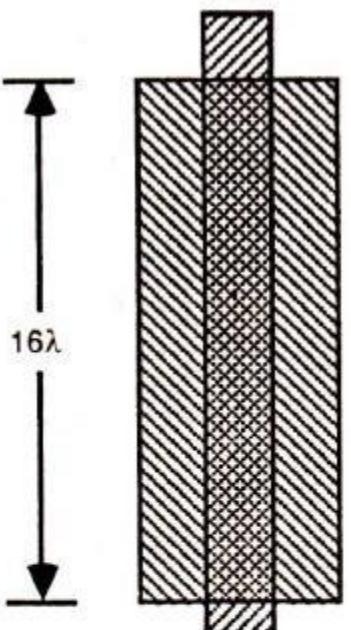
Sheet Resistance



Sheet Resistance



A ratio rule does not apply and there is no static resistance from V_{DD} to V_{SS} .



$$Z_{p.u.} = L_{p.u.}/W_{p.u.} = 8$$

$$R_{p.u.} = Z_{p.u.} \times R_s = 80 \text{ k}\Omega \text{ (nMOS)}$$

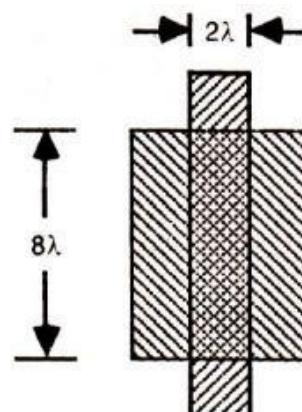
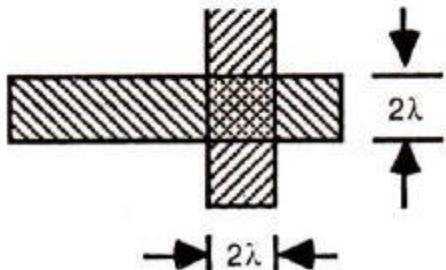
Similarly,

$$R_{p.d.} = Z_{p.d.} \times R_s = 10 \text{ k}\Omega$$

$$\text{Power dissipation (on)} P_d = \frac{V^2}{R_{p.u.} + R_{p.d.}}$$

$$= 0.28 \text{ mW}$$

$$\text{Input capacitance} = 1 \square C_g$$



$$Z_{p.u.} = L_{p.u.}/W_{p.u.} = 4$$

$$R_{p.u.} = Z_{p.u.} \times R_s = 40 \text{ k}\Omega \text{ (nMOS)}$$

Similarly,

$$R_{p.d.} = Z_{p.d.} \times R_s = 5 \text{ k}\Omega$$

$$\text{Power dissipation (on)} P_d = \frac{V^2}{R_{p.u.} + R_{p.d.}}$$

$$= 0.56 \text{ mW}$$

$$\text{Input capacitance} = 2 \square C_g$$

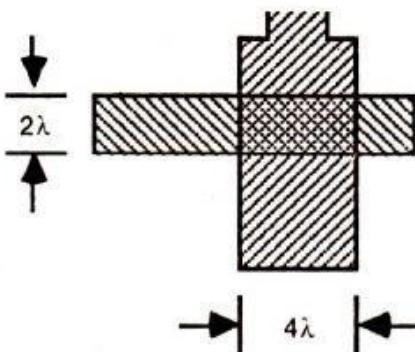


FIGURE 6.4 8:1 nMOS inverter (

Area capacitance of layers

$$C = \frac{\epsilon_{ins} \epsilon_0 A}{D}$$

D = thickness of silicon dioxide

A = area of plates

(and it is assumed that ϵ_0 , A , and D are in compatible units, for example, ϵ_0 in farads/cm,
 A in cm^2 , D in cm).

ϵ_{ins} = relative permittivity of $\text{SiO}_2 \approx 4.0$

$\epsilon_0 = 8.85 \times 10^{-14}$ F/cm (permittivity of free space)

A normal approach is to give layer area capacitances in $\text{pF}/\mu\text{m}^2$ (where μm = micron
= 10^{-6} meter = 10^{-4} cm). The appropriate figure may be calculated as follows:

$$C \left(\frac{\text{pF}}{\mu\text{m}^2} \right) = \frac{\epsilon_0 \epsilon_{ins}}{D} \frac{\text{F}}{\text{cm}^2} \times \frac{10^{12} \text{pF}}{\text{F}} \times \frac{\text{cm}^2}{10^8 \mu\text{m}^2}$$

TABLE 4.2 Typical area capacitance values for MOS circuits

<i>Capacitance</i>	<i>Value in pF × 10⁻⁴/μm² (Relative values in brackets)</i>		
	<i>5 μm</i>	<i>2 μm</i>	<i>1.2 μm</i>
Gate to channel	4 (1.0)	8 (1.0)	16 (1.0)
Diffusion (active)	1 (0.25)	1.75 (0.22)	3.75 (0.23)
Polysilicon* to substrate	0.4 (0.1)	0.6 (0.075)	0.6 (0.038)
Metal 1 to substrate	0.3 (0.075)	0.33 (0.04)	0.33 (0.02)
Metal 2 to substrate	0.2 (0.05)	0.17 (0.02)	0.17 (0.01)
Metal 2 to metal 1	0.4 (0.1)	0.5 (0.06)	0.5 (0.03)
Metal 2 to polysilicon	0.3 (0.075)	0.3 (0.038)	0.3 (0.018)

Notes: Relative value = specified value/gate to channel value for that technology.

*Poly. 1 and Poly. 2 are similar (also silicides where used).

Standard unit of capacitance

Standard unit of capacitance:

The unit is denoted as $\square C_g$ and is defined the gate-to-channel capacitance

Of a MOS transistor having $W = L = \text{feature size}$, that is, a 'standard' or 'feature size' square as for lambda-based rules.

Standard unit of capacitance

$\square C_g$ may be evaluated for any MOS process. For example, for 5 μm MOS circuits:

Area/standard square = 5 $\mu\text{m} \times 5 \mu\text{m} = 25 \mu\text{m}^2$ (= area of minimum size transistor)

Capacitance value (from Table 4.2) = $4 \times 10^{-4} \text{ pF}/\mu\text{m}^2$

Thus, standard value $\square C_g = 25 \mu\text{m}^2 \times 4 \times 10^{-4} \text{ pF}/\mu\text{m}^2 = .01 \text{ pF}$

for 2 μm MOS circuits (Orbit):

Area/standard square = 2 $\mu\text{m} \times 2 \mu\text{m} = 4 \mu\text{m}^2$

Gate capacitance value (from Table 4.2) = $8 \times 10^{-4} \text{ pF}/\mu\text{m}^2$

Thus, standard value $\square C_g = 4 \mu\text{m}^2 \times 8 \times 10^{-4} \text{ pF}/\mu\text{m}^2 = .0032 \text{ pF}$

for 1.2 μm MOS circuits (Orbit):

Area/standard square = 1.2 $\mu\text{m} \times 1.2 \mu\text{m} = 1.44 \mu\text{m}^2$

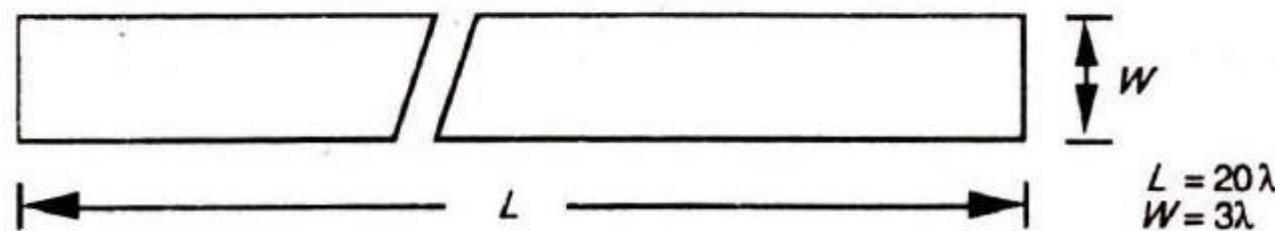
Gate capacitance value (from Table 4.2) = $16 \times 10^{-4} \text{ pF}/\mu\text{m}^2$

Thus, standard value $\square C_g = 1.44 \mu\text{m}^2 \times 16 \times 10^{-4} \text{ pF}/\mu\text{m}^2 = .0023 \text{ pF}$

Standard unit of capacitance

The calculation of capacitance values can be done by establishing the ratio between the area of interest and the area of standard (feature size square) gate ($2\lambda \times 2\lambda$. for λ -based rules) and multiplying this ratio by the appropriate relative C value from Table 4.2. The product will give the required capacitance in $\square C_g$ units.

$$\text{Relative area} = \frac{20\lambda \times 3\lambda}{2\lambda \times 2\lambda} = 15$$



1. Consider the area in metal 1.

Capacitance to substrate = relative area \times relative C value

$$= 15 \times 0.0750 \square C_g$$

$$= 1.125 \square C_g$$

Standard unit of capacitance

2. Consider the same area in polysilicon.

$$\begin{aligned}\text{Capacitance to substrate} &= 15 \times 0.1 \square C_g \\ &= 1.5 \square C_g\end{aligned}$$

3. Consider the same area in n-type diffusion.

$$\begin{aligned}\text{Capacitance to substrate} &= 15 \times 0.25 \square C_g \\ &= 3.75 \square C_g^*\end{aligned}$$

Standard unit of capacitance

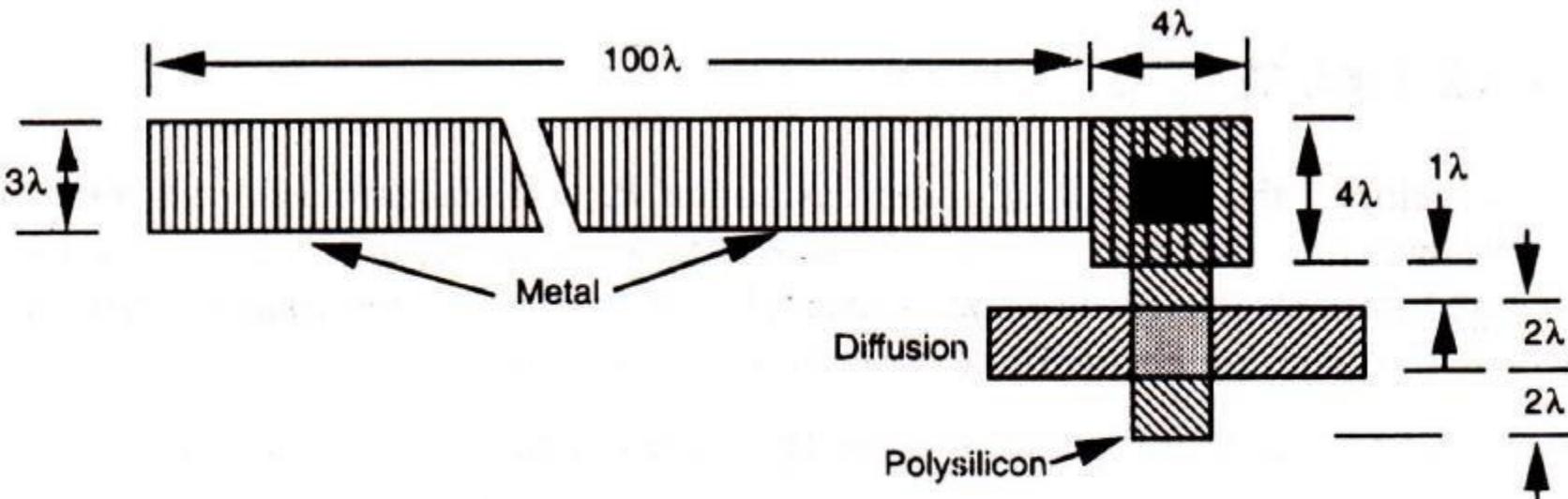


FIGURE 4.5 Capacitance calculation (multilayer).

Consider the metal area (less the contact region where the metal is connected to polysilicon and shielded from the substrate)

Standard unit of capacitance

$$\text{Ratio} = \frac{\text{Metal area}}{\text{Standard gate area}} = \frac{100\lambda \times 3\lambda}{4\lambda^2} = 75$$

$$\text{Metal capacitance } C_m = 75 \times 0.075 = 5.625 \square C_g$$

Consider the polysilicon area (excluding the gate region)

$$\text{Polysilicon area} = 4\lambda \times 4\lambda + 3\lambda \times 2\lambda = 22\lambda^2$$

Therefore

$$\text{Polysilicon capacitance } C_p = \frac{22}{4} \times 0.1 = .55 \square C_g$$

For the transistor,

$$\text{Gate capacitance } C_g = 1 \square C_g$$

$$\text{Total capacitance } C_T = C_m + C_p + C_g = 7.20 \square C_g$$

For example, if the metal width is increased to 4λ in Figure 4.5, the capacitance C_m is increased to $7.5 C_g$ and the capacitance of the complete structure will increase to about $9 C_s$.

Delay Calculation

Time constant: It is time required to charge one standard (feature size square) gate area capacitance through one feature size square of n channel resistance (that is, through R_s for an nMOS pass transistor channel),

$$\text{Time constant } \tau = (1R_s \text{ (n channel)} \times 1 \square C_g) \text{ seconds}$$

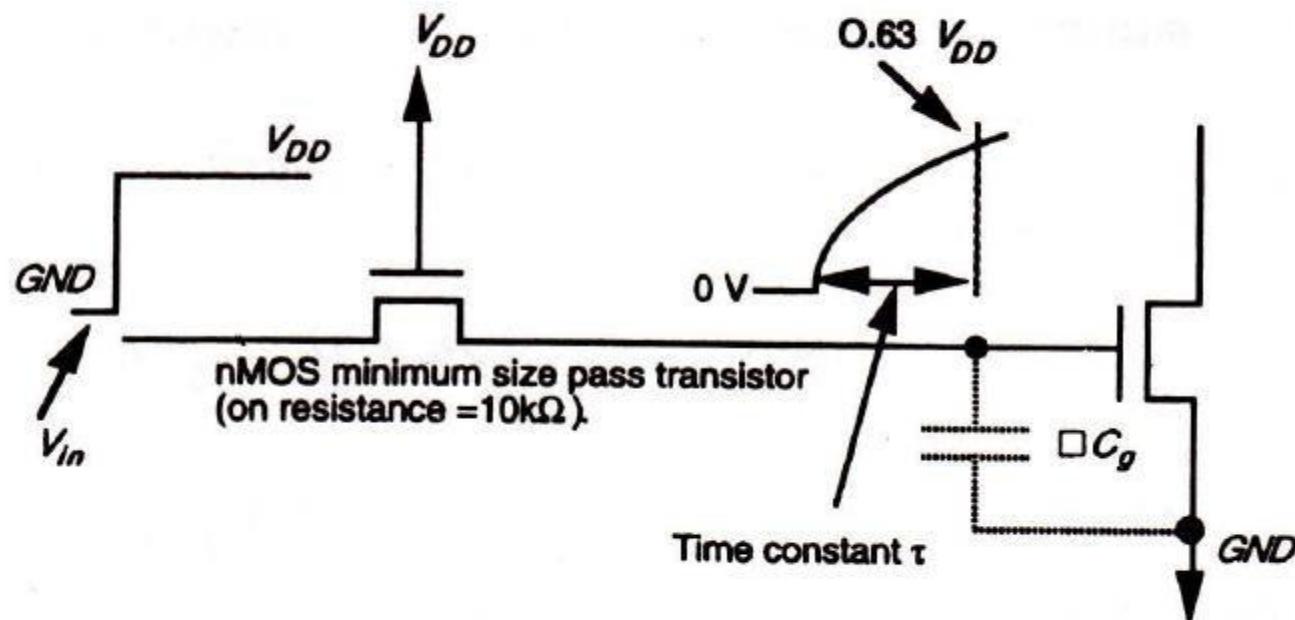


FIGURE 4.6 Model for derivation of τ .

Delay Calculation

This can be evaluated for any technology and for 5 μm technology,

$$\tau = 10^4 \text{ ohm} \times 0.01 \text{ pF} = 0.1 \text{ nsec}$$

and for 2 μm (Orbit) technology,

$$\tau = 2 \times 10^4 \text{ ohm} \times 0.0032 \text{ pF} = 0.064 \text{ nsec}$$

and for 1.2 μm (Orbit) technology,

$$\tau = 2 \times 10^4 \text{ ohm} \times 0.0023 \text{ pF} = 0.046 \text{ nsec}$$

However, in practice, circuit wiring and parasitic capacitances must be allowed for so that the figure taken for τ is often increased by a factor of two or three so that for 5 μm circuit

$\tau = 0.2$ to 0.3 nsec is a typical design figure used in assessing likely worst case delays.

Note that τ thus obtained is not much different from transit time τ_{sd} calculated from equation (2.2).

$$\tau_{sd} = \frac{L^2}{\mu_n V_{ds}}$$

Note that V_{ds} varies as C_g charges from 0 volts to 63% of V_{DD} in period τ in Figure 4.6, so that an appropriate value for V_{ds} is the average value = 3 volts. For 5 μm technology, then,

$$\begin{aligned}\tau_{sd} &= \frac{25 \mu\text{m}^2 \text{ V sec}}{650 \text{ cm}^2 \text{ 3V}} \times \frac{10^9 \text{ nsec cm}^2}{10^8 \mu\text{m}^2} \\ &= 0.13 \text{ nsec}\end{aligned}$$

This is very close to the theoretical time constant τ calculated above.

Delay Calculation

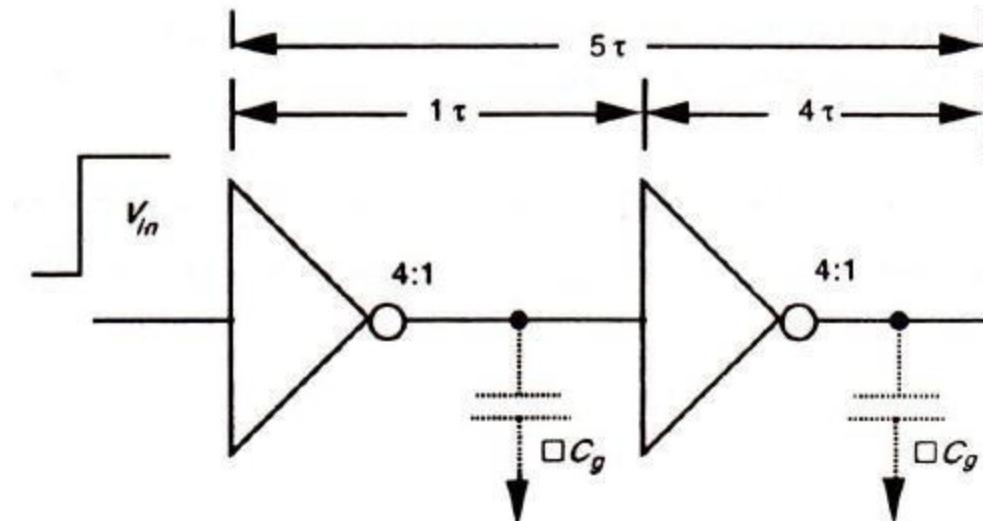


FIGURE 4.7 nMOS inverter pair delay.

$$T_d = (1 + Z_{p.u.}/Z_{p.d.})\tau$$

Delay Calculation

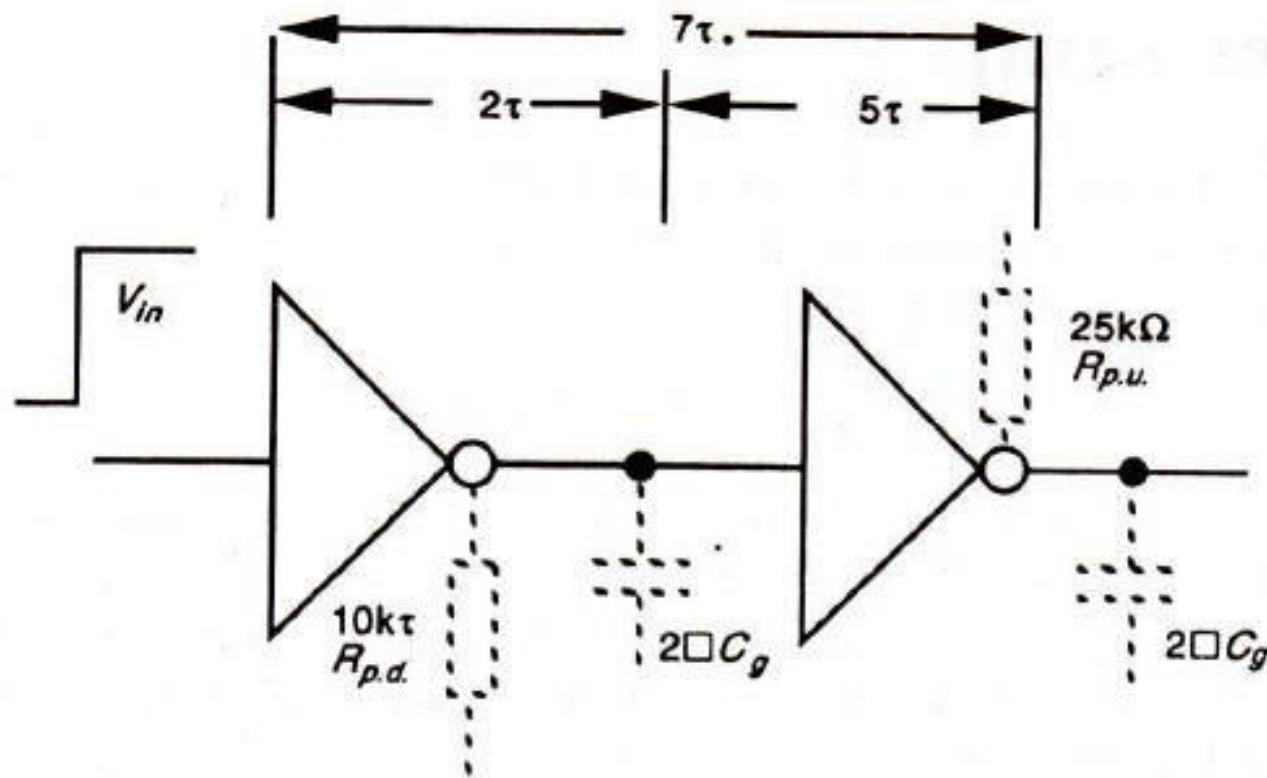


FIGURE 4.8 Minimum size CMOS inverter pair delay.

**1. minimum size lambda-based geometries this would result
in the inverter having an input capacitance of**

$$1\square C_g \text{ (n-device)} + 2.5\square C_g \text{ (p-device)} = 3.5\square C_g$$

For $\beta_n = \beta_p$.

Formal Estimation of CMOS Inverter Delay

Rise-time estimation

$$I_{dsp} = \frac{\beta_p (V_{gs} - |V_{tp}|)^2}{2}$$

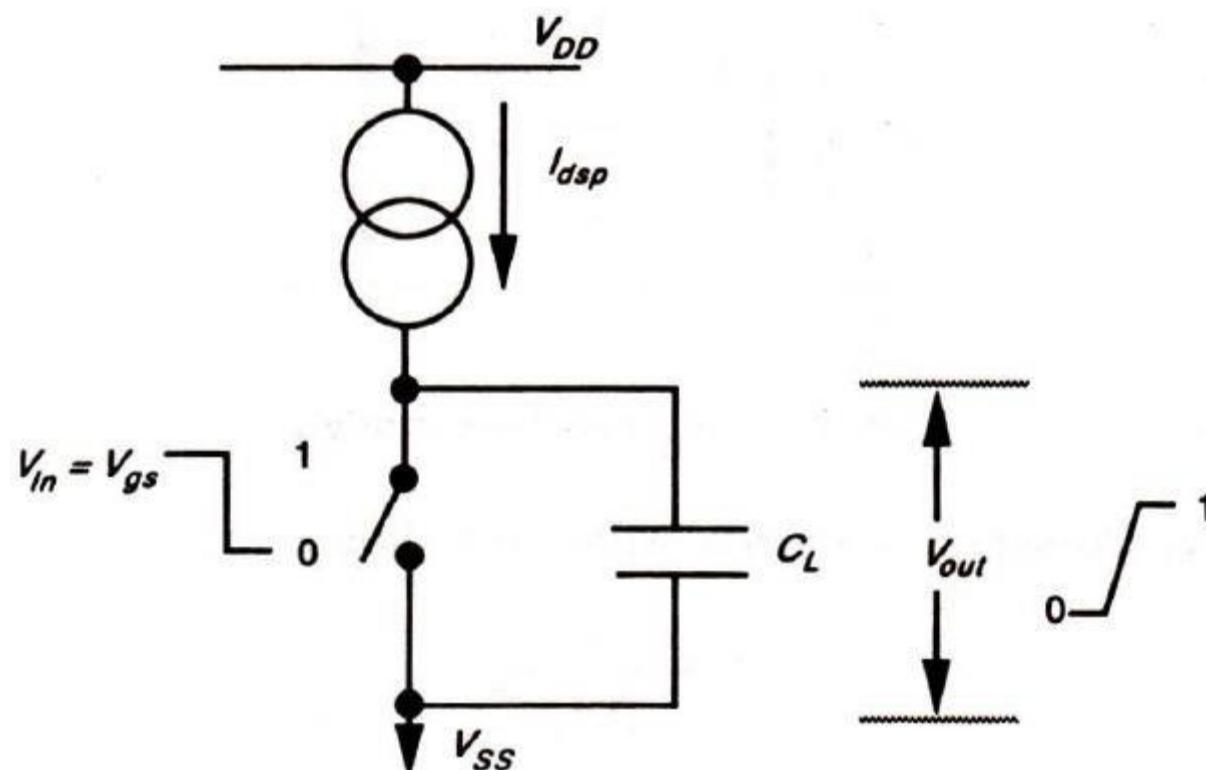


FIGURE 4.9 Rise-time model.

This current charges C_L and, since its magnitude is approximately constant, we have

$$V_{out} = \frac{I_{dsp}t}{C_L}$$

Substituting for I_{dsp} and rearranging we have

$$t = \frac{2C_L V_{out}}{\beta_p (V_{gs} - |V_{tp}|)^2}$$

We now assume that $t = \tau_r$ when $V_{out} = +V_{DD}$, so that

$$\tau_r = \frac{2V_{DD}C_L}{\beta_p (V_{DD} - |V_{tp}|)^2}$$

with $|V_{tp}| = 0.2V_{DD}$, then

$$\tau_r = \frac{3C_L}{\beta_p V_{DD}}$$

Fall-time estimation

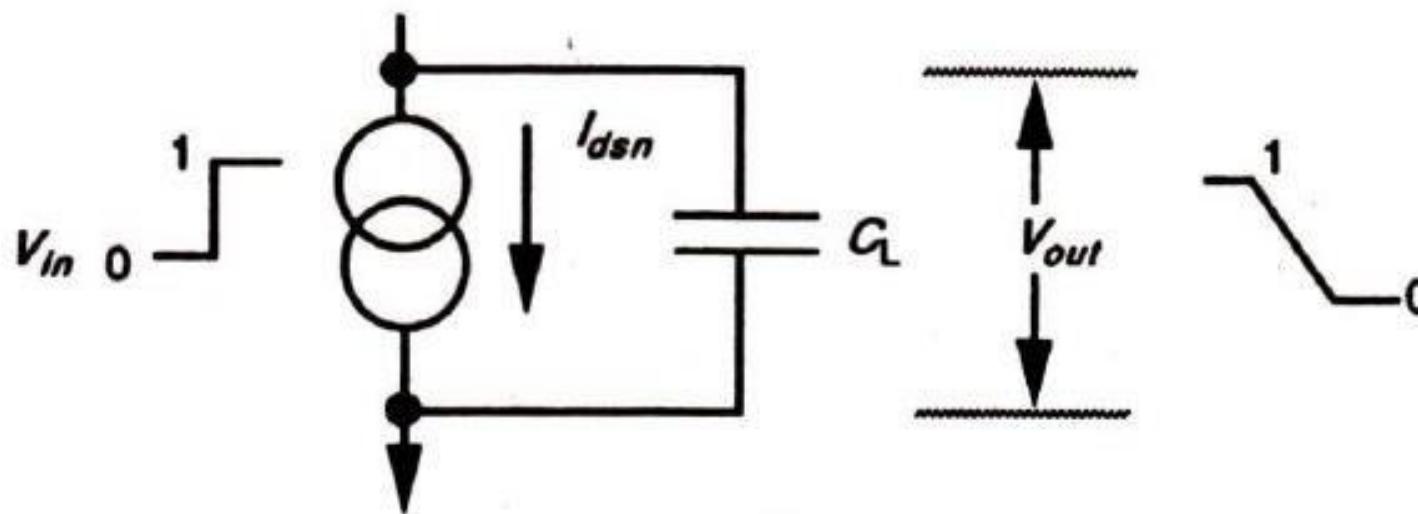


FIGURE 4.10 Fall-time model.

$$\tau_f = \frac{3C_L}{\beta_n V_{DD}}$$

Summary of CMOS rise and fall factors

Using these expressions we may deduce that:

$$\frac{\tau_r}{\tau_f} = \frac{\beta_n}{\beta_p}$$

But $\mu_n = 2.5 \mu_p$ and hence $\beta_n = 2.5\beta_p$, so that the rise-time is slower by a factor of 2.5 when using minimum size devices for both 'n' and 'p'.

In order to achieve symmetrical operation using minimum channel length, we would need to make $W_p = 2.5W_n$ and for minimum size lambda-based geometries this would result in the inverter having an input capacitance of $1\square C_g$ (n-device) + $2.5\square C_g$ (p-device) = $3.5\square C_g$ in total.

1. τ_r and τ_f are proportional to $1/V_{DD}$;
2. τ_r and τ_f are proportional to C_L ;
3. $\tau_r = 2.5\tau_f$ for equal n- and p-transistor geometries.

DRIVING LARGE CAPACITIVE LOADS

The problem of driving comparatively large capacitive loads arises when signals must be propagated from the chip to off chip destinations. Generally, typical off chip capacitances may be several orders higher than on chip $\square C_g$ values. For example, if the off chip load is denoted C_L then

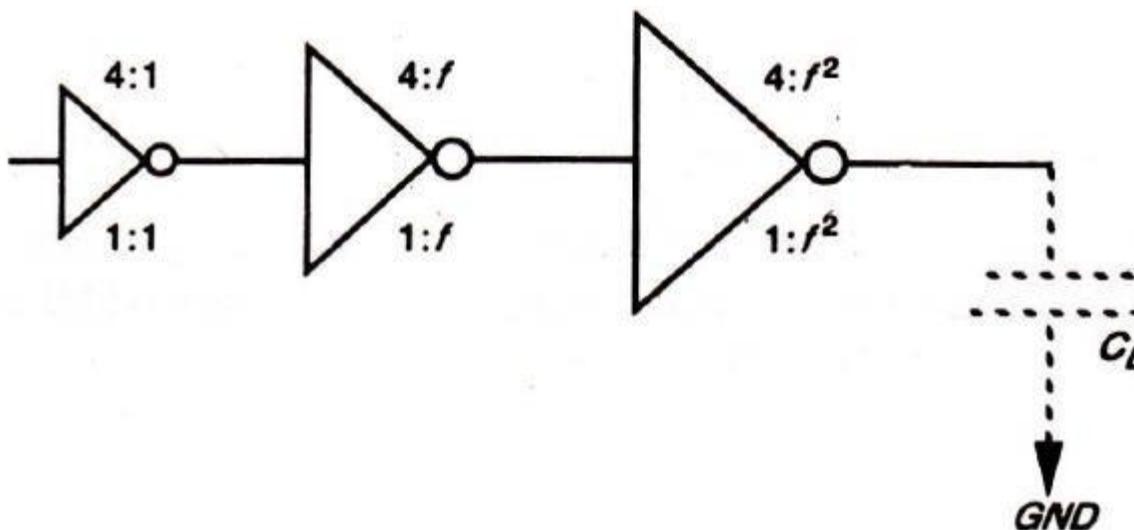
$$C_L \geq 10^4 \square C_g \text{ (typically)}$$

Clearly capacitances of this order must be driven through low resistances, otherwise excessively long delays will occur.

Cascaded Inverters as Drivers

- for MOS circuits, low resistance values for $Z_{p.d.}$ and $Z_{p.u.}$ imply low $L: W$ ratios;
- So channels must be made very wide to reduce resistance value
- Then gate region area $L \times W$ becomes significant and a comparatively large capacitance is presented at the input, which in turn slows down the rates of change of voltage which can take place at the input.

- The remedy is to use N cascaded inverters, each one of which is larger than the preceding stage by a width factor f as shown in Figure below.



Driving large capacitive loads

With large f , N decreases but delay per stage increases. For 4:1 nMOS inverters

$$\left. \begin{array}{l} \text{delay per stage} = f\tau \text{ for } \Delta V_{in} \\ \text{or} = 4f\tau \text{ for } \nabla V_{in} \end{array} \right\} \quad \begin{array}{l} \text{where } \Delta V_{in} \text{ indicates logic 0 to 1} \\ \text{transition and } \nabla V_{in} \text{ indicates} \\ \text{logic 1 to 0 transition of } V_{in} \end{array}$$

Therefore, total delay per nMOS pair = $5f\tau$. A similar treatment yields delay per CMOS pair = $7f\tau$. Now let

$$y = \frac{C_L}{\square C_g} = f^N$$

We now need to determine the value of f which will minimize the overall delay for a given value of y and from the definition of y

$$\ln(y) = N \ln(f)$$

$$\text{i.e } N = \frac{\ln(y)}{\ln(f)}$$

Thus, for N even total delay = $\frac{N}{2} 5f\tau = 2.5 Nf\tau$ (nMOS)

or = $\frac{N}{2} 7f\tau = 3.5 Nf\tau$ (CMOS)

Thus, in all cases delay $\propto Nf\tau = \frac{\ln(y)}{\ln(f)} f\tau$

For minimum delay Thus, assuming that $f = e$, we have

and overall delay t_d

N even: $t_d = 2.5eN \tau$ (nMOS)

or $t_d = 3.5eN \tau$ (CMOS)

N odd: $t_d = [2.5(N - 1) + 1]e\tau$ (nMOS) }
 or $t_d = [3.5(N - 1) + 2]e\tau$ (CMOS) } for ΔV_{in}

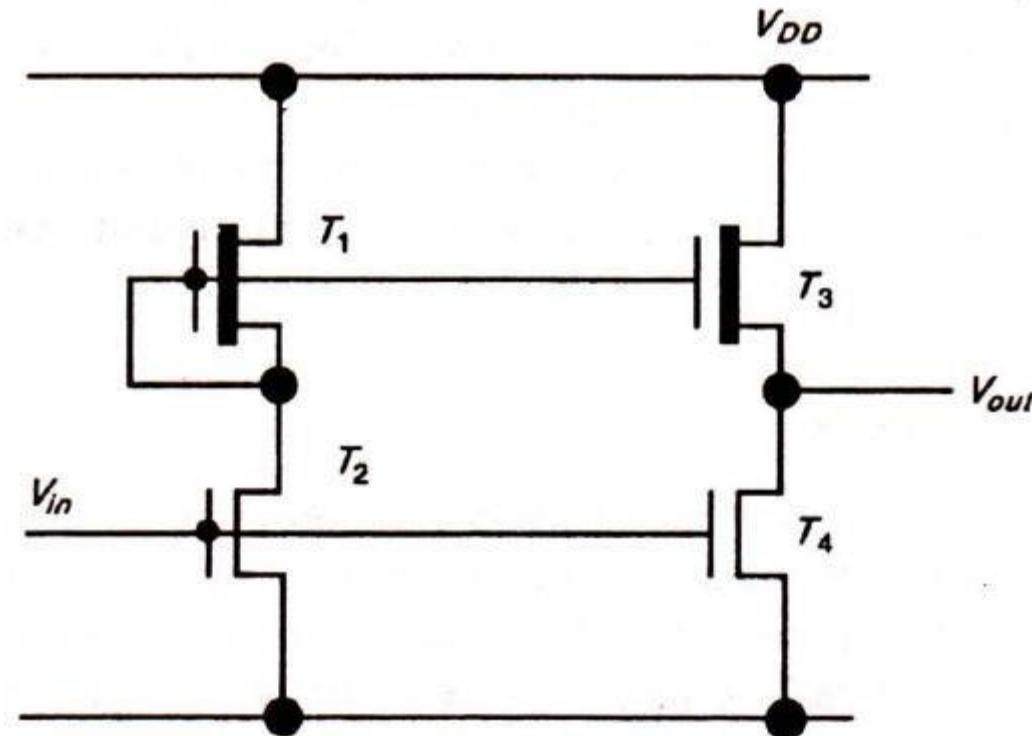
or

$t_d = [2.5(N - 1) + 4]e\tau$ (nMOS) }
 or $t_d = [3.5(N - 1) + 5]e\tau$ (CMOS) } for ∇V_{in}

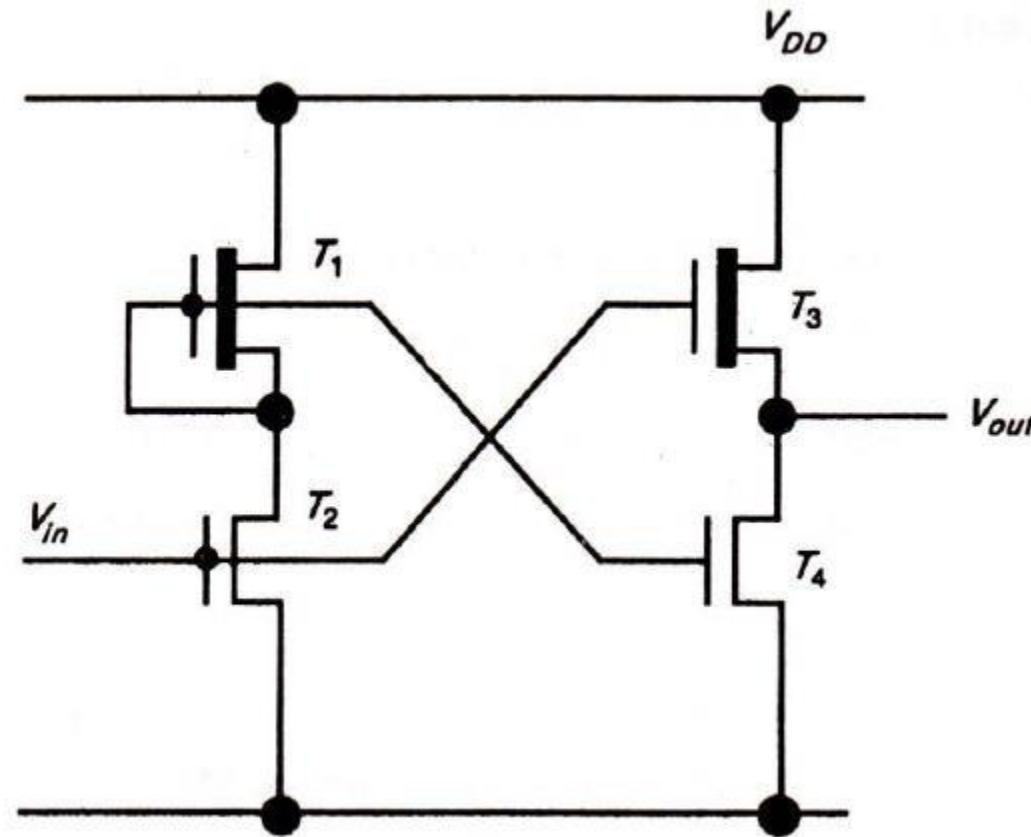
Super Buffers

- The asymmetry of the conventional inverter is clearly undesirable, and gives rise to significant delay problems when an inverter is used to drive more significant capacitive loads.
- A common approach used in nMOS technology to alleviate this effect is to make use of super buffers.
- **Inverting type nMOS super buffer**
- **Non-Inverting type nMOS super buffer.**

Inverting type nMOS super buffer.



Non-Inverting type nMOS super buffer



Subsystem design

Design description domains and design strategies

1. Define the requirements (properly and carefully).
2. Partition the overall architecture into appropriate subsystems.
3. Consider communication paths carefully in order to develop sensible interrelationships between subsystems.
4. Draw a floor plan of how the system is to map onto the silicon (and alternate between 2, 3 and 4 as necessary).
5. Aim for regular structures so that design is largely a matter of replication.
6. Draw suitable (stick or symbolic) diagrams of the leaf-cells of the subsystems.
7. Convert each cell to a layout.
8. Carefully and thoroughly carry out a design rule check on each cell.
9. Simulate the performance of each cell/subsystem.

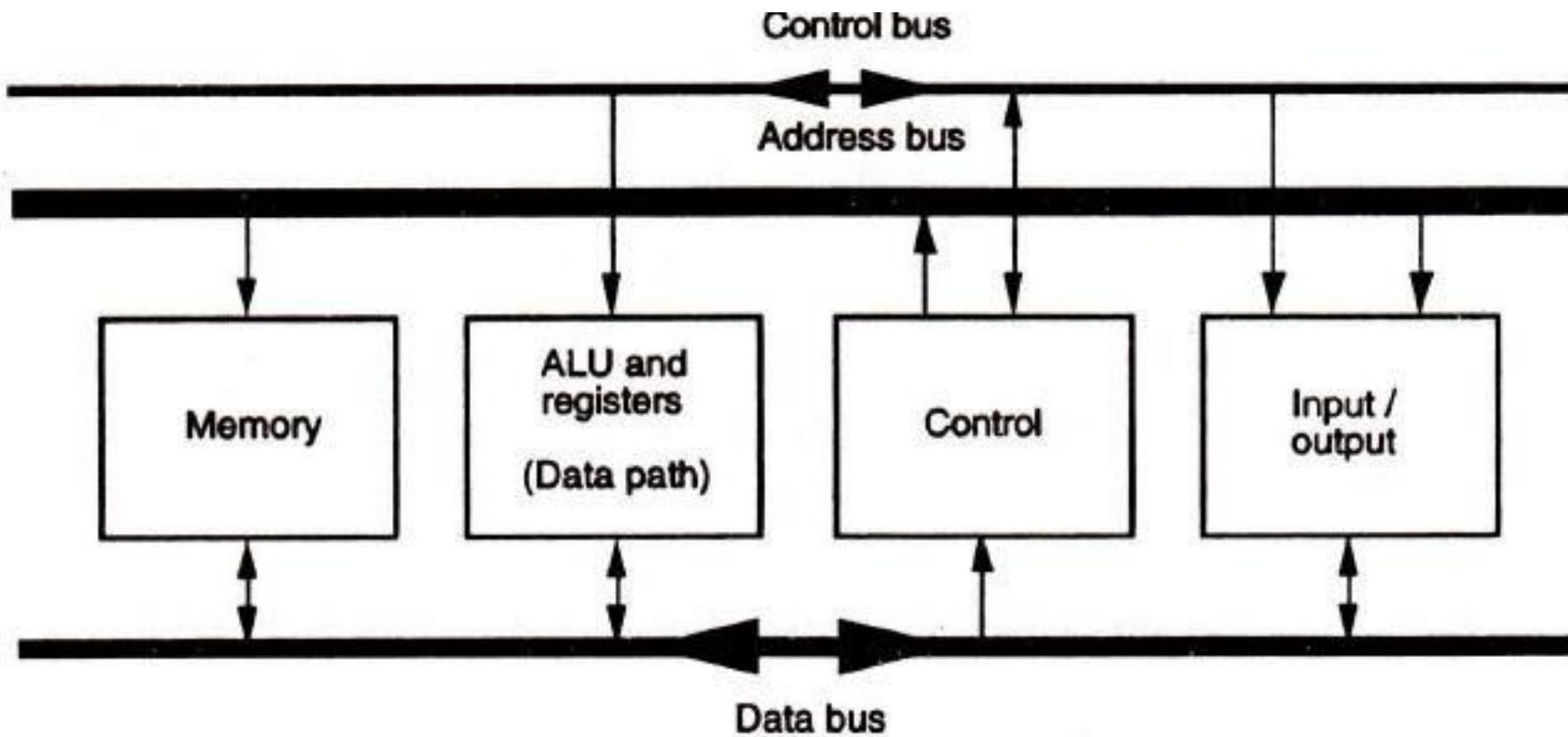


FIGURE 7.1 Basic digital processor structure.

Datapath- in detail

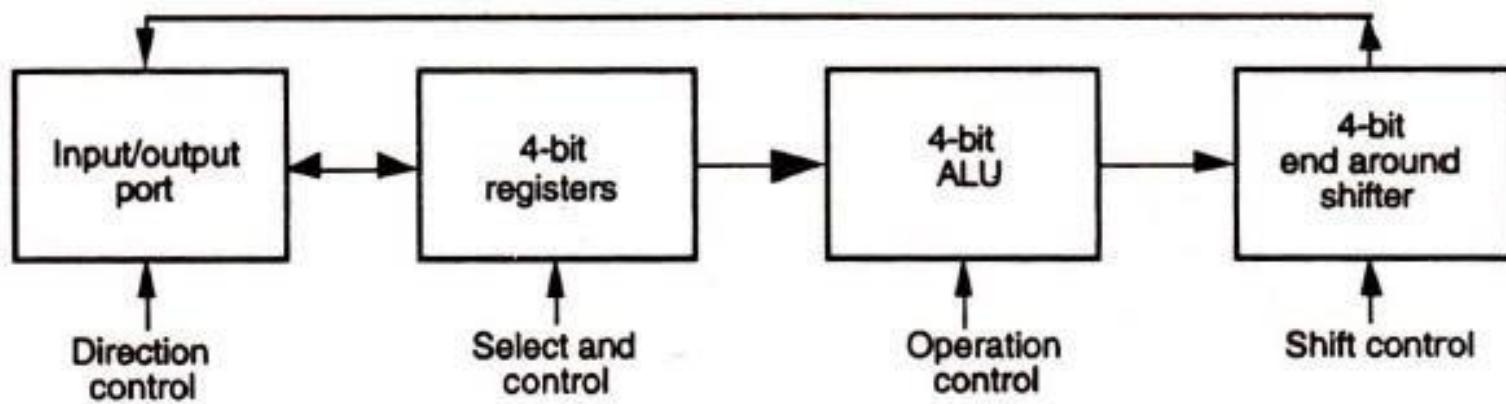


FIGURE 7.3 Subunits and basic interconnections for data path.

Shifter design-4X4 Crossbar Switch

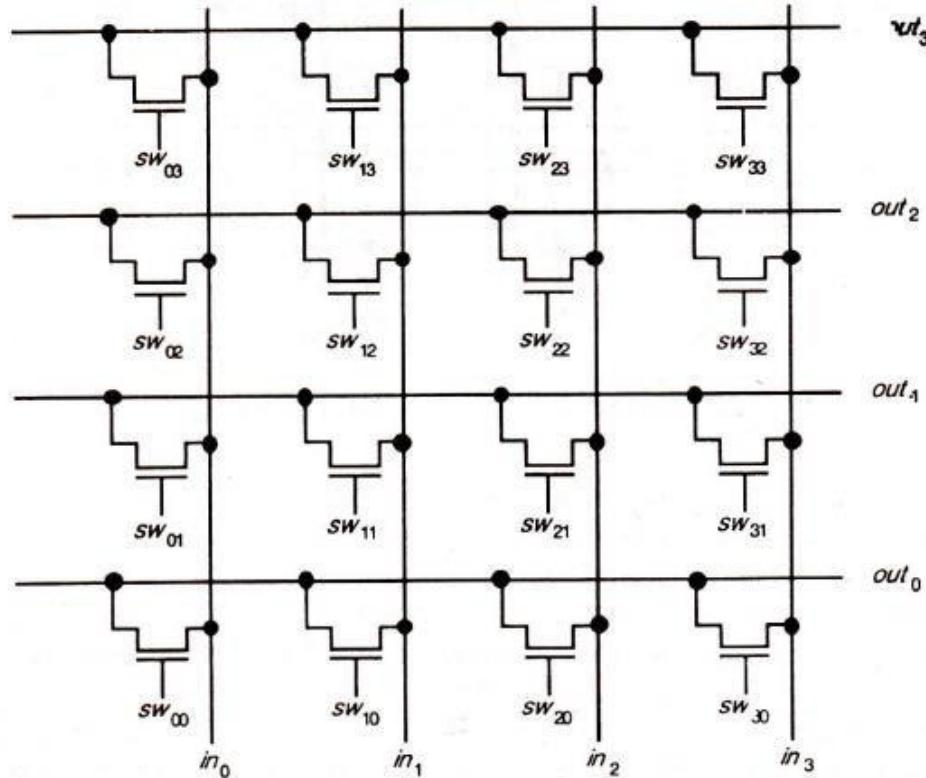


FIGURE 7.6 4×4 crossbar switch.

$$\text{Regularity} = \frac{\text{Total number of transistors on the chip}}{\text{Number of transistor circuits that must be designed in detail}}$$

Barrel Shifter

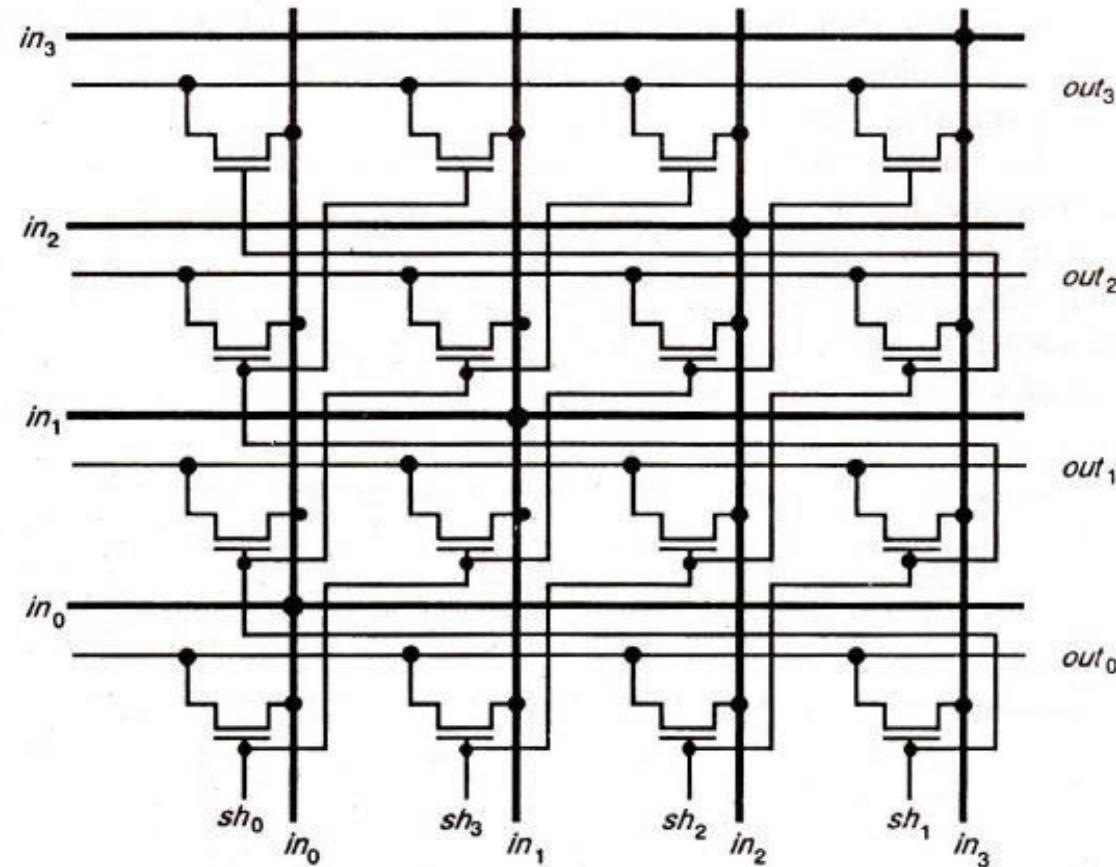
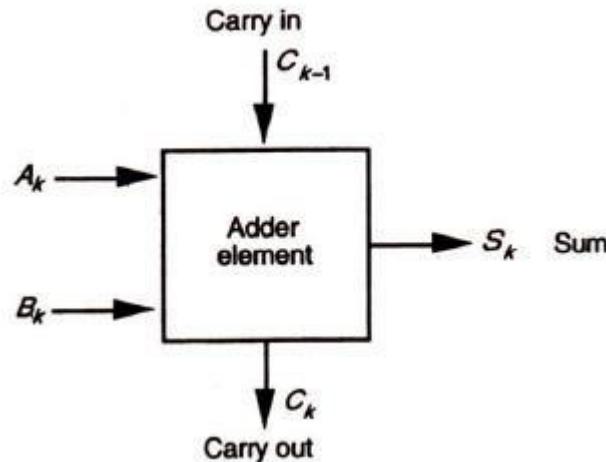


FIGURE 7.7 4×4 barrel shifter.

Basic Building Block of ALU



Sum
$$S_k = H_k \bar{C}_{k-1} + \bar{H}_k C_{k-1}$$

New carry
$$C_k = A_k B_k + H_k C_{k-1}$$

Half sum
$$H_k = \bar{A}_k B_k + A_k \bar{B}_k$$

Consider, first, the *Sum* output if C_{k-1} is held at logical 0, then

$$S_k = H_k \cdot 1 + \bar{H}_k \cdot 0 = H_k$$

that is

$$S_k = H_k = A_k B_k + A_k \bar{B}_k — \text{An Exclusive-Or operation}$$

Now, hold C_{k-1} at logical 1, then

$$S_k = H_k \cdot 0 + \bar{H}_k \cdot 1 = \bar{H}_k$$

that is

$$S_k = \bar{H}_k = \bar{A}_k \bar{B}_k + A_k B_k — \text{An Exclusive-Nor (Equality) operation}$$

$$C_k = A_k \cdot B_k + H_k \cdot 0 = A_k \cdot B_k — \text{An And operation}$$

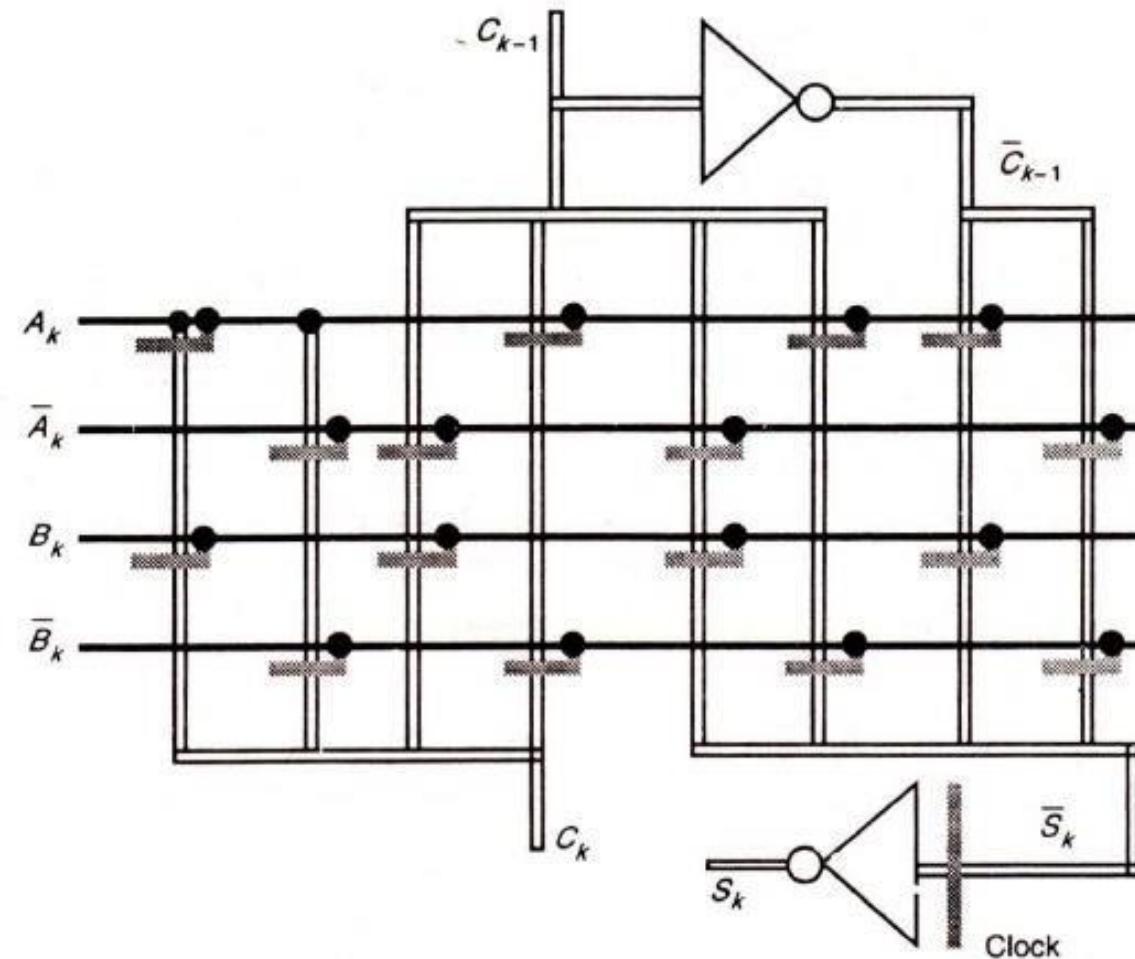
Now, if C_{k-1} is held at logical 1, then

$$C_k = A_k \cdot B_k + H_k \cdot 1 = A_k \cdot B_k + \bar{A}_k \cdot B_k + A_k \cdot \bar{B}_k$$

Therefore

$$C_k = A_k + B_k — \text{An Or operation}$$

Stick Diagram Representation of Full adder



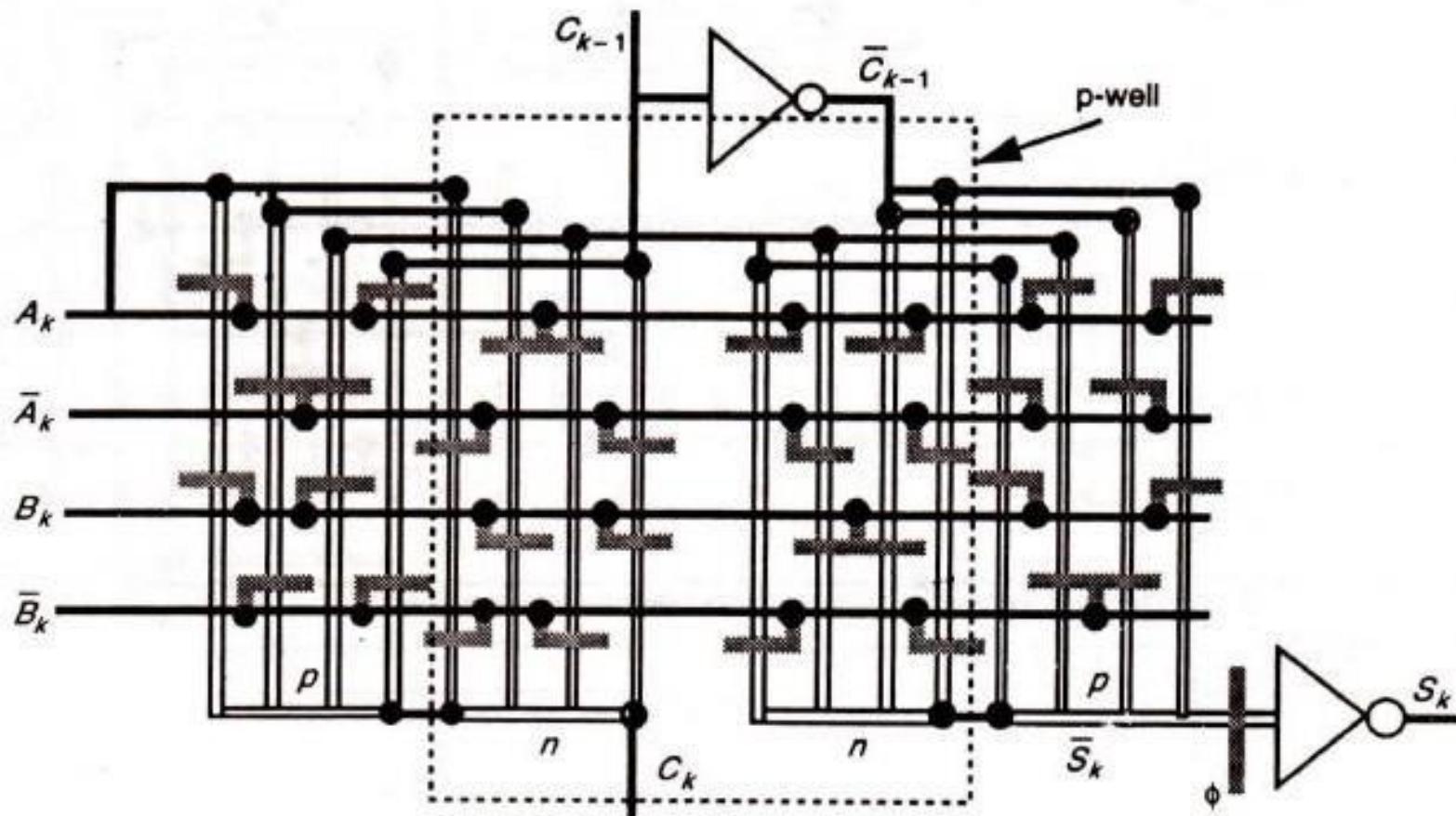


FIGURE 8.4 CMOS version of adder logic.

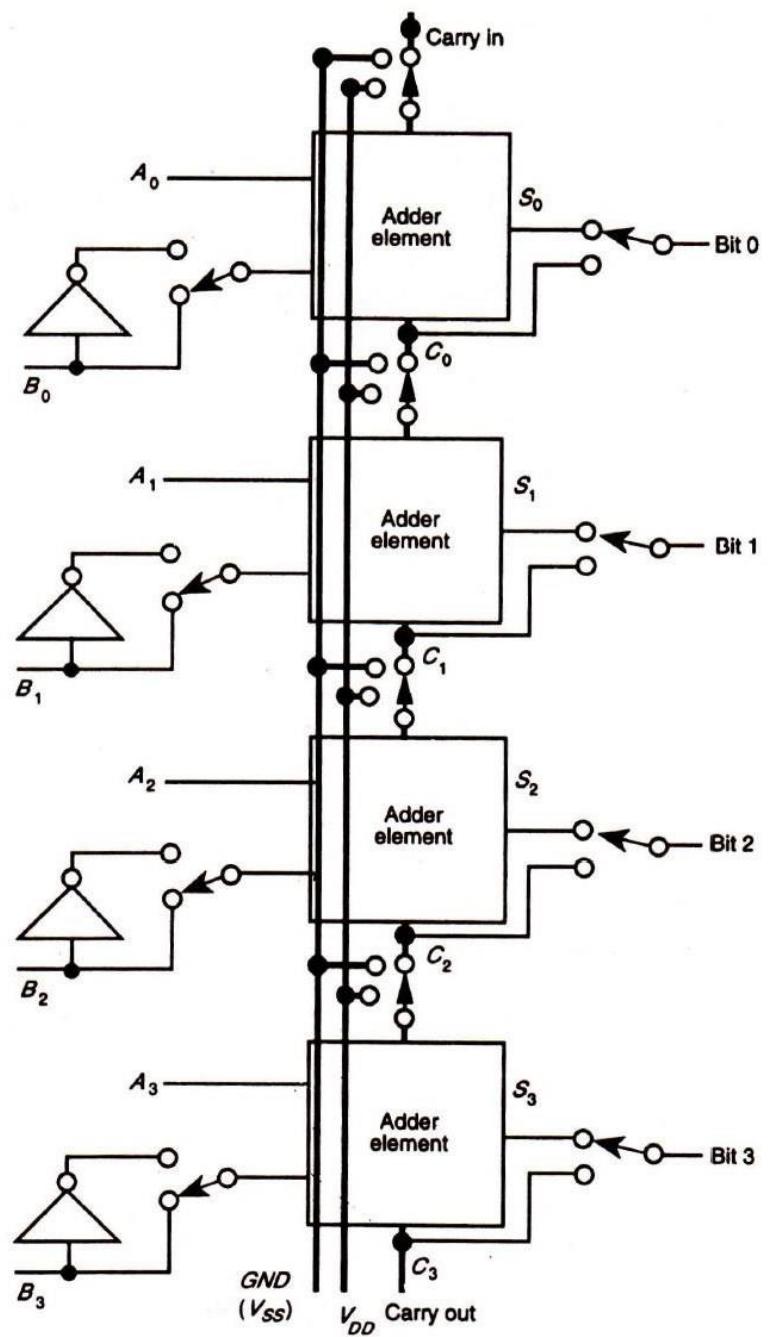
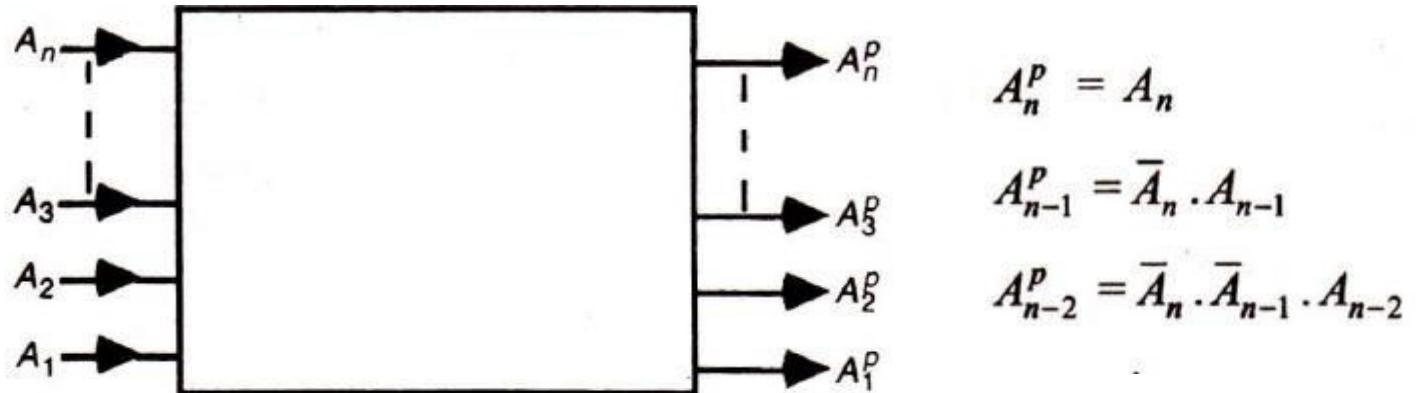


FIGURE 8.12 4-bit ALU.

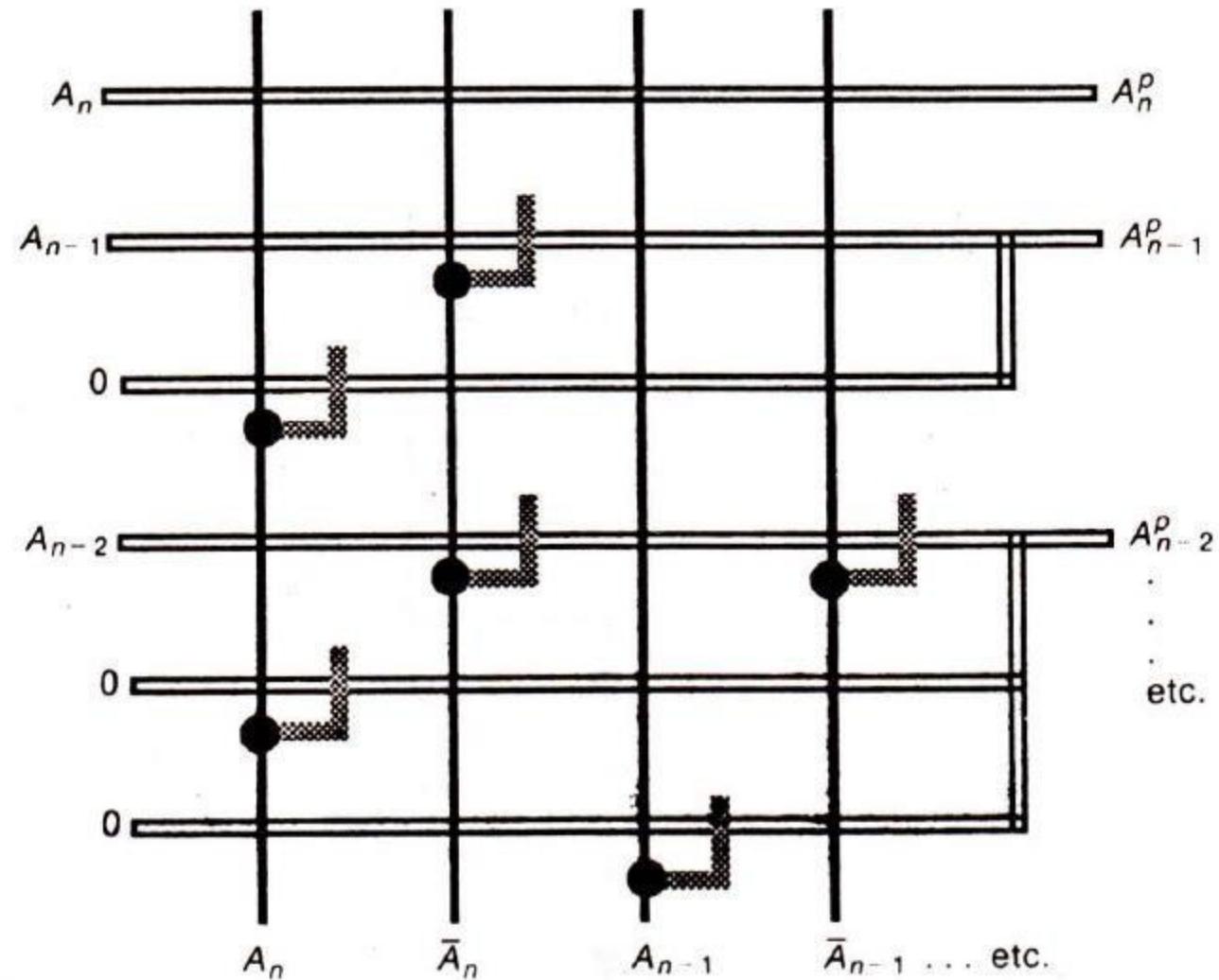
Bus arbitration Logic



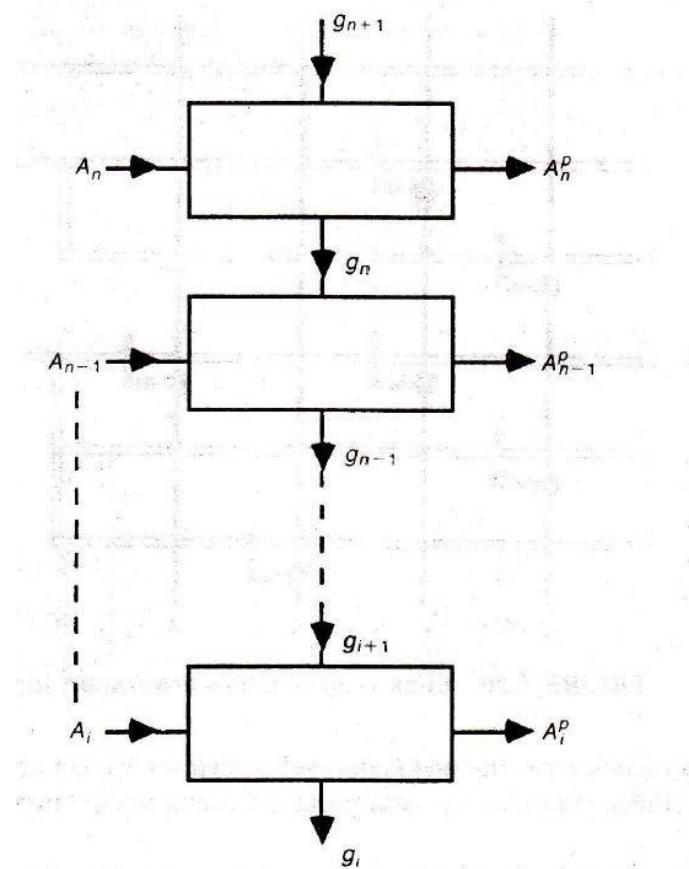
Truth table

A_n	...	A_3	A_2	A_1	A_n^P	...	A_3^P	A_2^P	A_1^P
0	...	0	0	0	0	...	0	0	0
0		0	0	1	0	...	0	0	1
0		0	1	X	0		0	1	0
0		1	X	X	0		1	0	0
.
.
.
1	...	X	X	X	1	...	0	0	0

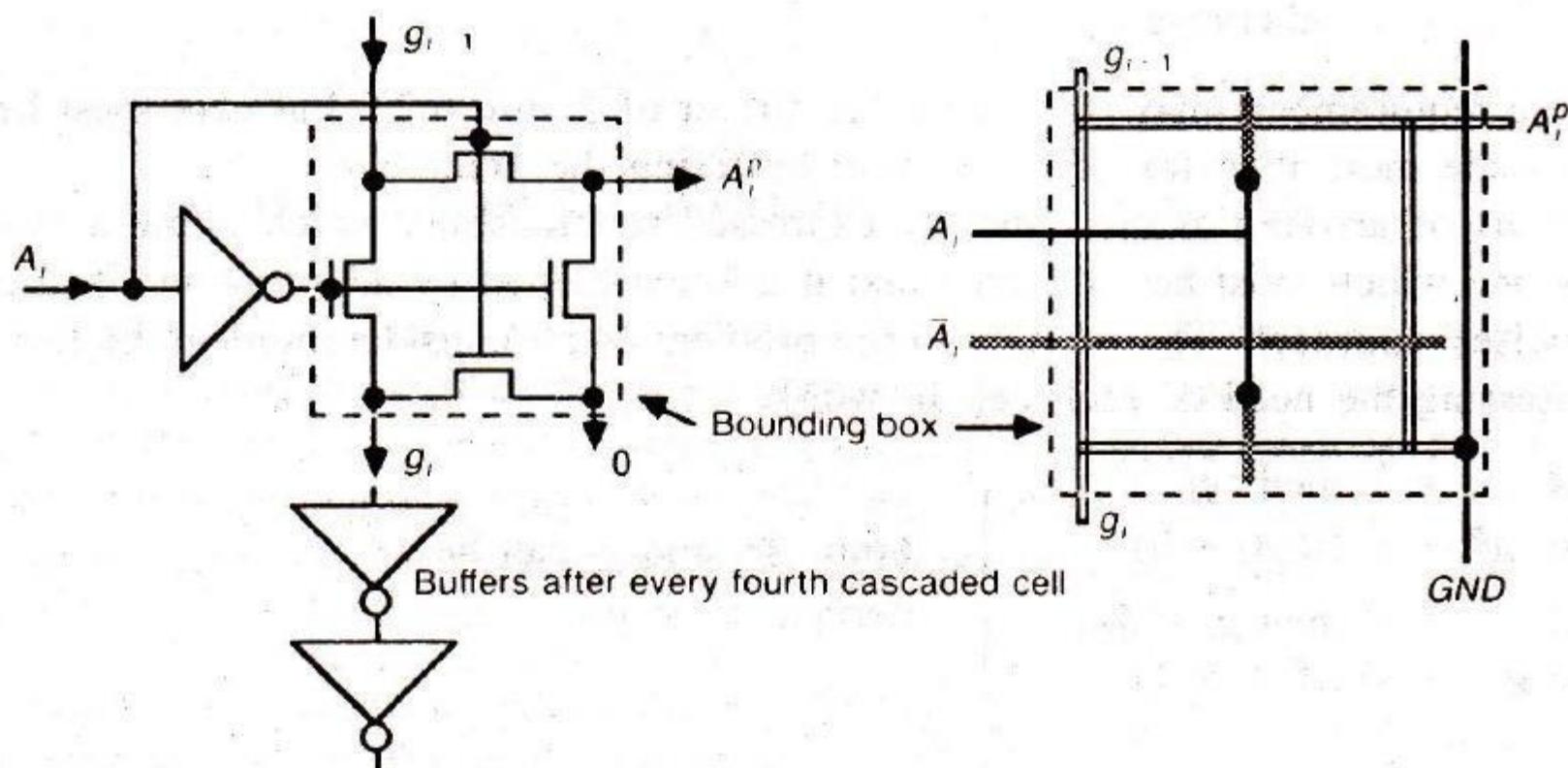
X = Don't care



Bus arbitration Logic



- x 1. g = grant line.
- 2. If grant line is 1, none of the lines above it wants priority.
- 3. If $A_i = 0$, pass grant.



(a) Circuit

(b) Stick diagram

$$A_i^P = \begin{cases} g_{i+1} & \text{if } A_i = 1 \\ \text{or } 0 & \text{otherwise} \end{cases} \quad \left. \begin{array}{l} \text{If } A_i = 1 \text{ then } A_i^P = g_{i+1}, \\ \text{else } A_i^P = 0 \text{ (if } A_i = 0\text{)} \end{array} \right\}$$

$$g_i = \begin{cases} 0 & \text{if } A_i = 1 \\ \text{or } g_{i+1} & \text{otherwise} \end{cases} \quad \left. \begin{array}{l} \text{If } A_i = 0 \text{ then } g_i = g_{i+1}, \\ \text{else } g_i = 0 \text{ (if } A_i = 1\text{)} \end{array} \right\}$$

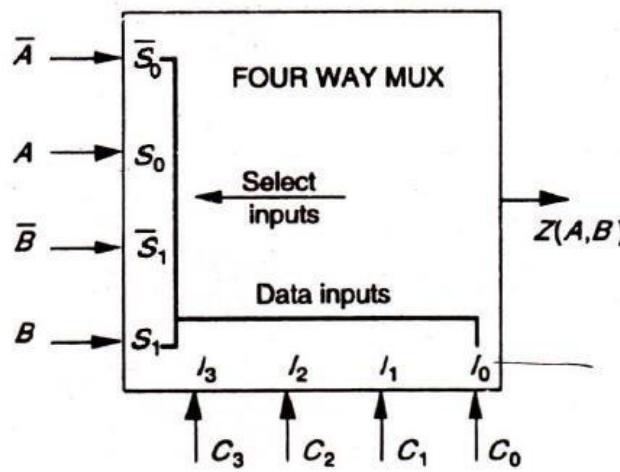
both A_i^P and g_i can be derived from g_{i+1}

$$A_i^p = A_i \cdot g_{i+1}; \bar{A}_i^p = \bar{A}_i + \bar{g}_{i+1}$$

$$g_i = \bar{A}_i g_{i+1}; \bar{g}_i = A_i + \bar{g}_{i+1}$$

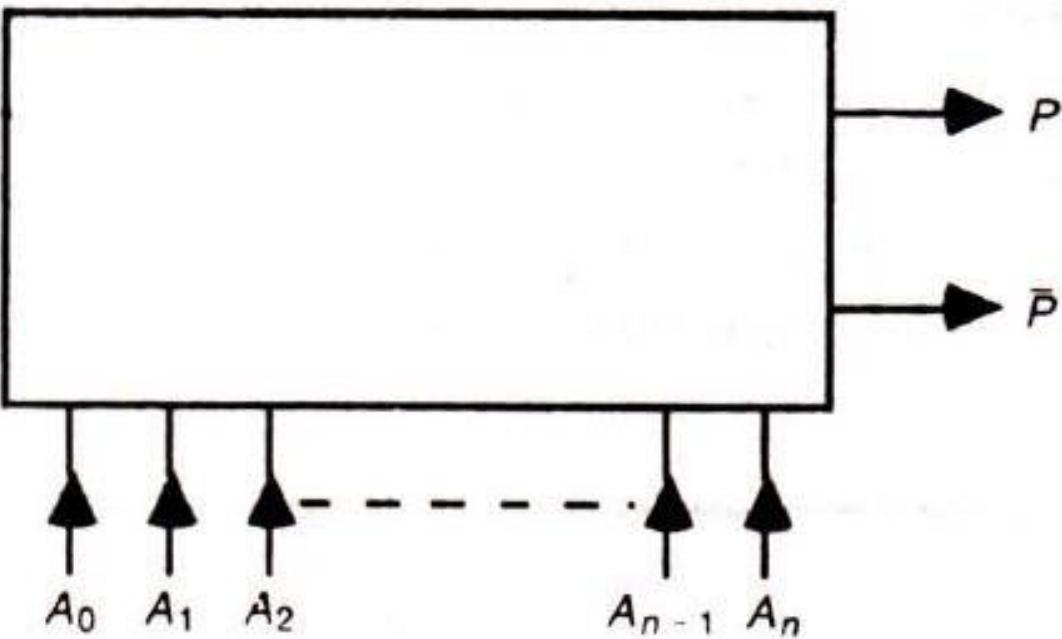
MUX

1,2,3,6,7,8,9,11,12,14,15,17,19,20,21,23,,24,27,26,28,29,30,31,32,33,35,37,38,40,41,43,45,
46,51,53,57,58,60,



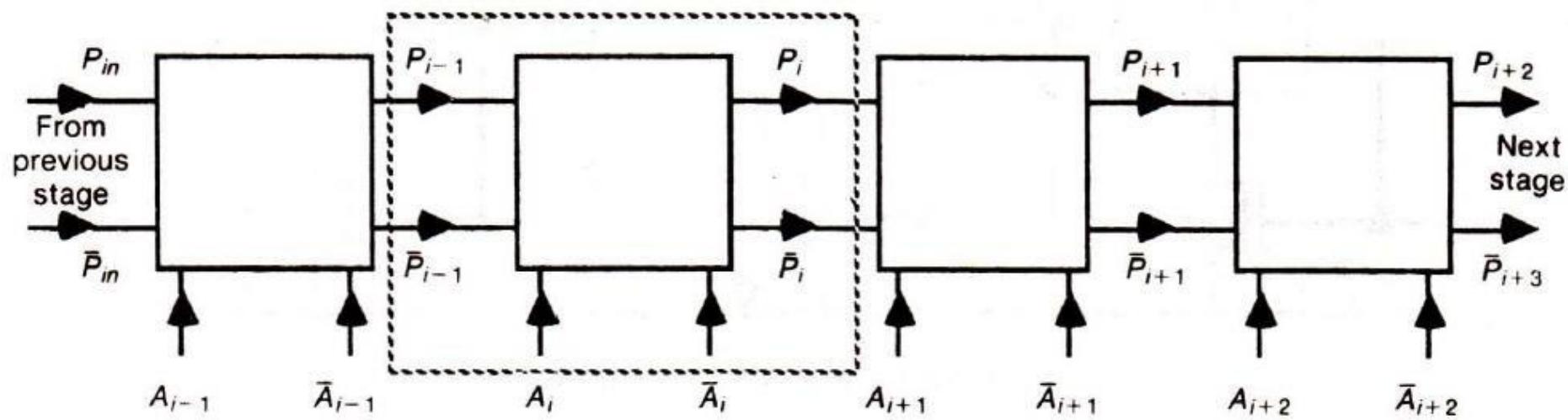
INPUT PROGRAMMING				FUNCTION $Z(A,B)$	
C_3	C_2	C_1	C_0	Z	
0	0	0	0	0	$Z = 0$
0	0	0	1	$\bar{A} \cdot \bar{B}; \bar{A} + \bar{B}$	$Not\ B$
0	0	1	0	$A \cdot \bar{B}$	
0	0	1	1	\bar{B}	$Not\ B$
0	1	0	0	$\bar{A} \cdot B$	
0	1	0	1	\bar{A}	$Not\ A$
0	1	1	0	$A \cdot \bar{B} + \bar{A} \cdot B$	$Exclusive-Or$
0	1	1	1	$\bar{A} + \bar{B}; \bar{A} \cdot B$	$Nand$
1	0	0	0	$A \cdot B$	And
1	0	0	1	$\bar{A} \cdot \bar{B} + A \cdot B$	$Comparator$
1	0	1	0	A	$O/P = A$
1	0	1	1	$A + \bar{B}$	
1	1	0	0	B	$O/P = B$
1	1	0	1	$\bar{A} + B$	Or
1	1	1	0	$A + B$	
1	1	1	1	1	$Z = 1$

Parity Generator



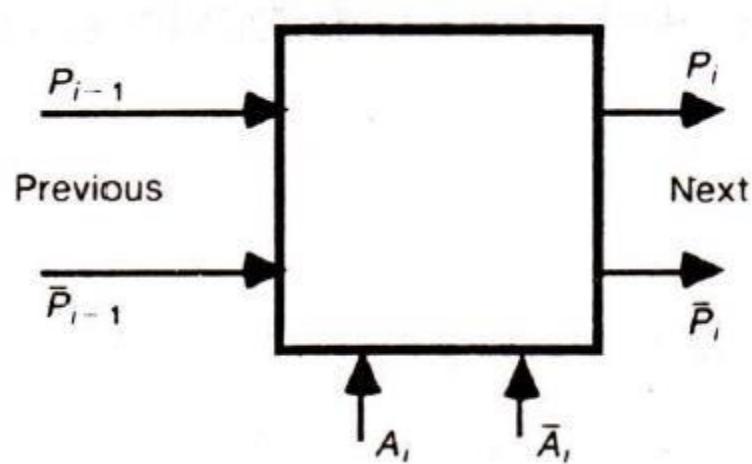
Note: $P = \begin{cases} 1 & \text{Even number of } 1\text{s at input} \\ 0 & \text{Odd number of } 1\text{s at input} \end{cases}$

FIGURE 6.15 Parity generator basic block diagram.



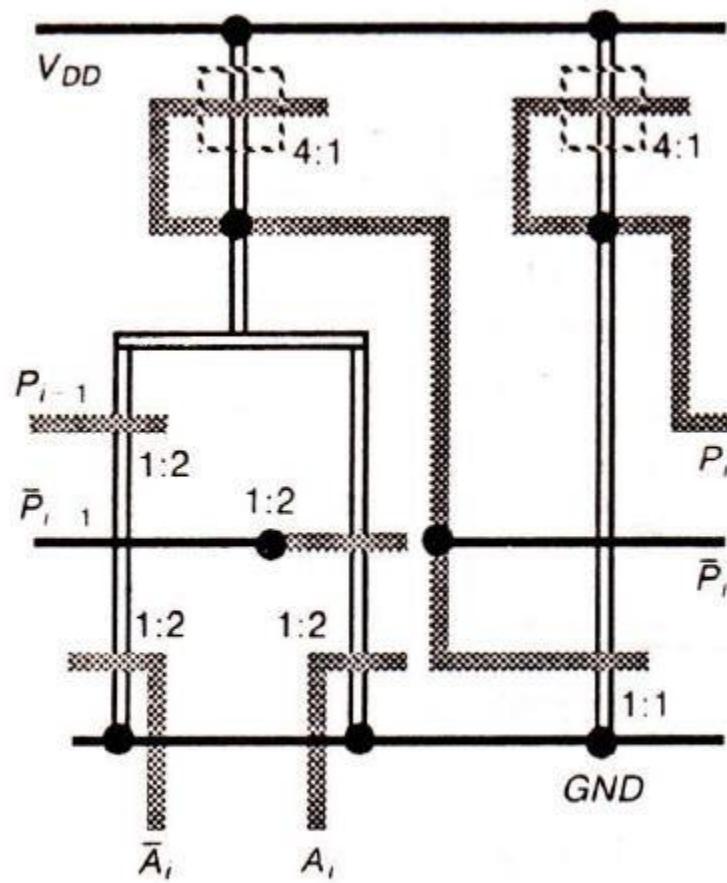
Note: Parity requirements are set at the left-most cell where $P_{in} = 1$ sets even and $P_{in} = 0$ sets odd parity.

FIGURE 6.16 Parity generator—structured design approach.

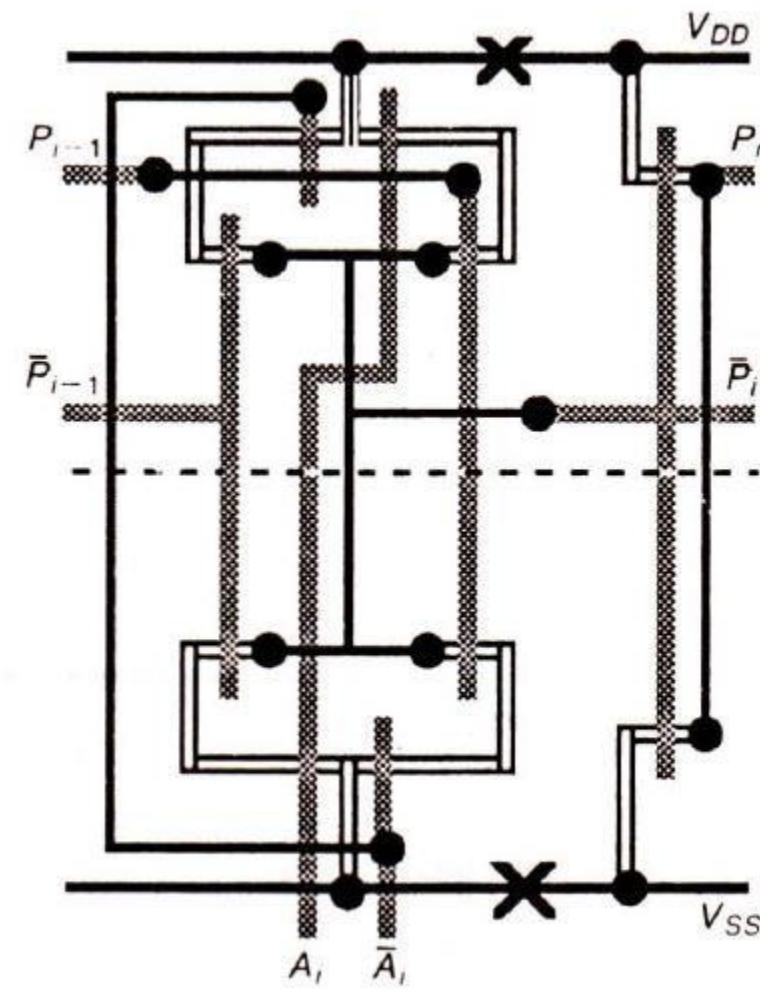


$A_i = 1$ parity is changed, $P_i = \bar{P}_{i-1}$
 $A_i = 0$ parity is unchanged, $P_i = P_{i-1}$

$$P_i = \bar{P}_{i-1} \cdot A_i + P_{i-1} \cdot \bar{A}_i$$



(a) nMOS

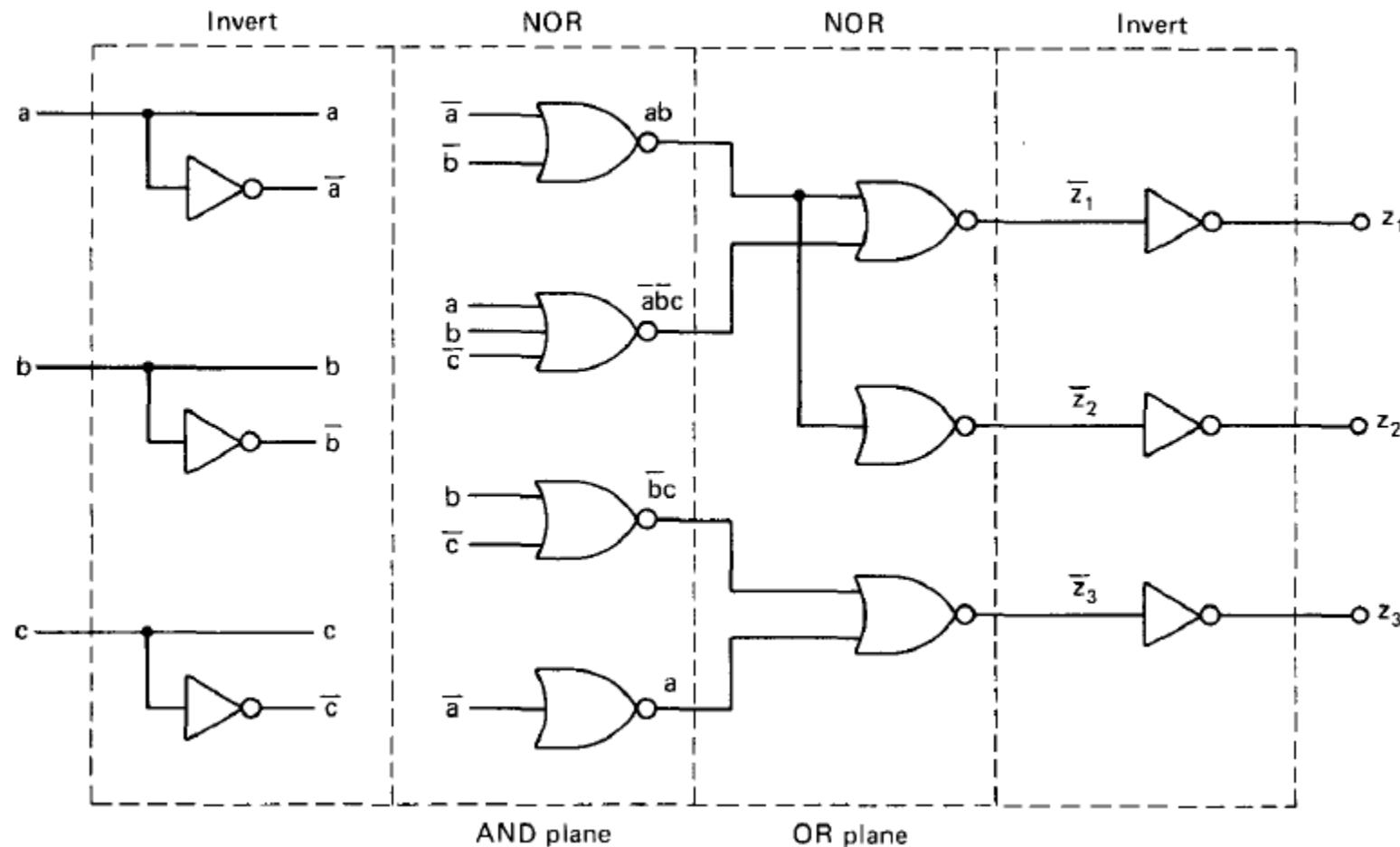


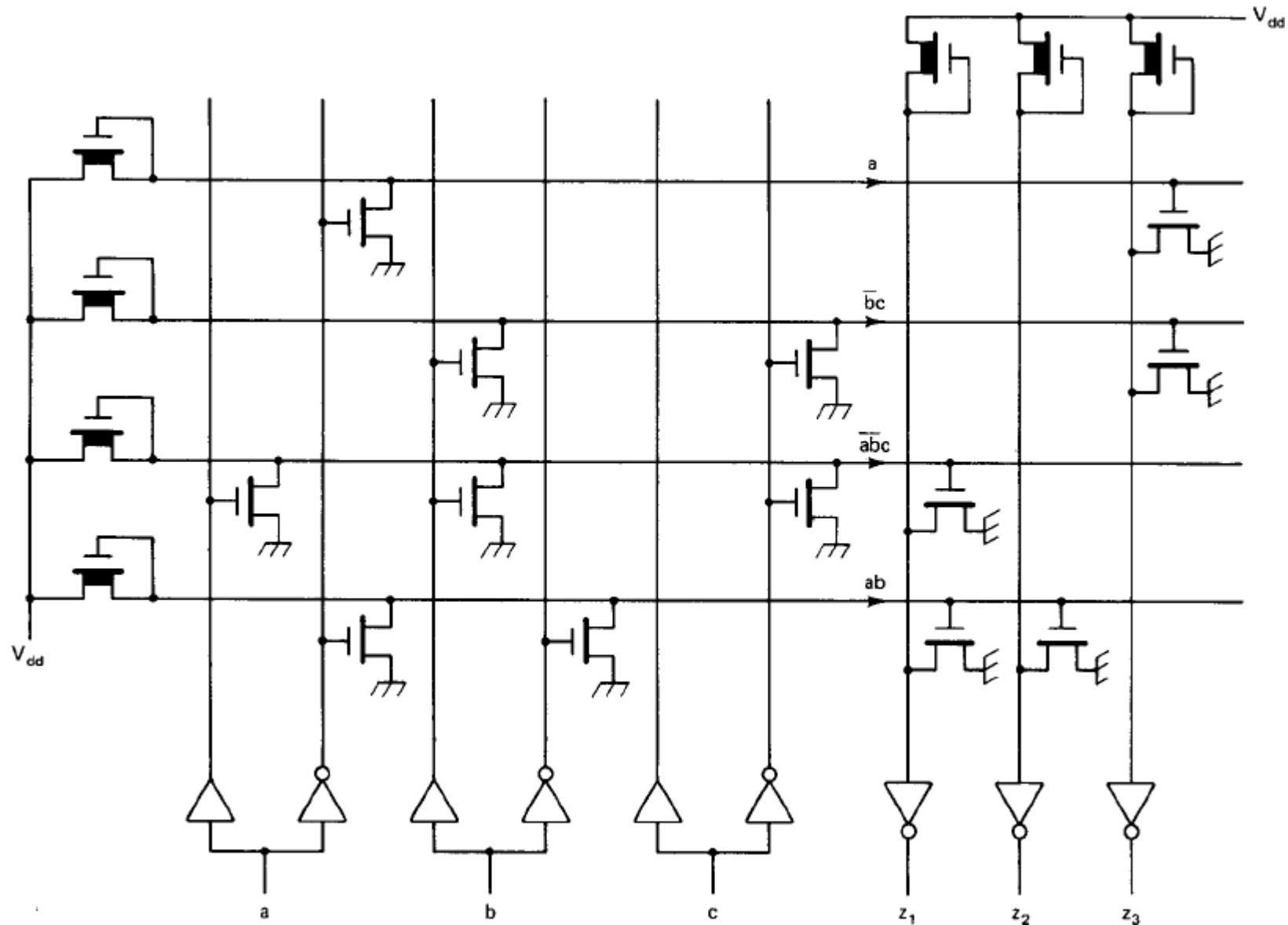
(b) CMOS

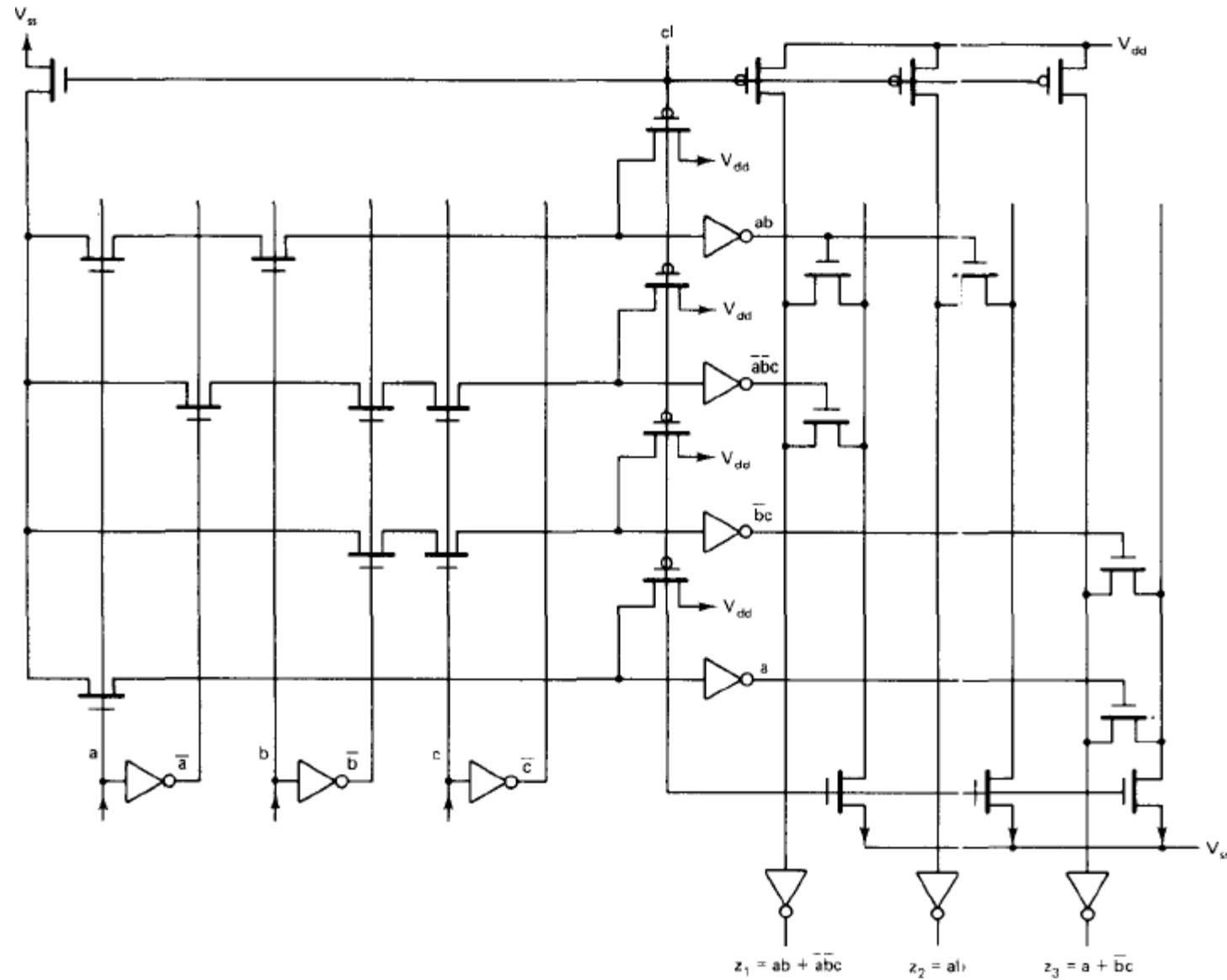
$$z_1 = ab + \bar{a}\bar{b}c$$

$$z_2 = ab$$

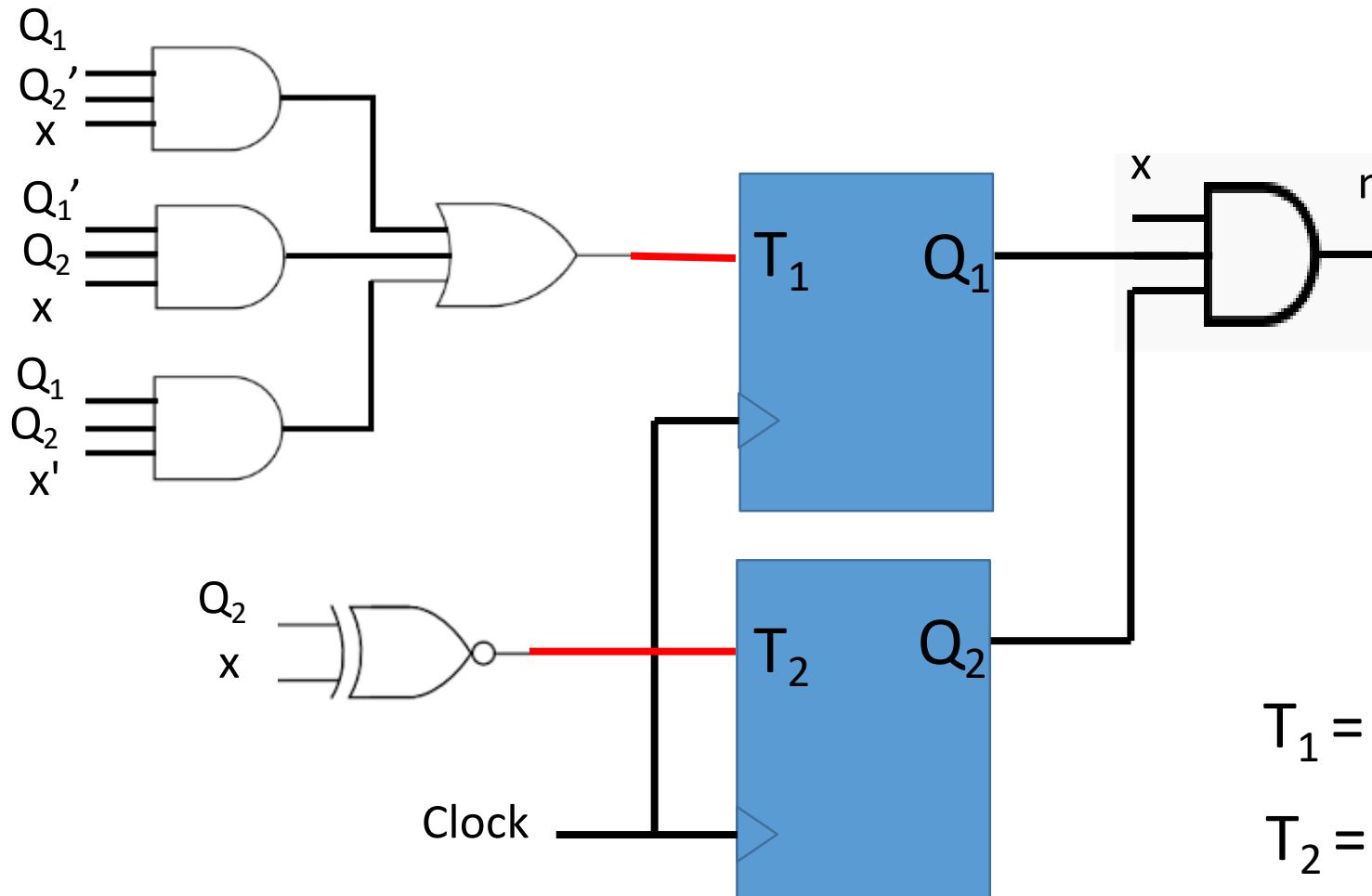
$$z_3 = a + \bar{b}c$$







Design of 0101 overlapping sequence detector using T-flip flop and basic gates.



$$T_1 = Q_1 Q_2' x + Q_1' Q_2 x + Q_1 Q_2 x'$$

$$T_2 = Q_2' x' + Q_2 x$$

$$n = Q_1 Q_2 x$$

Assignment-2

1. Explain the working of FAMOS. Implement the given words using Flash ROM

$$\begin{bmatrix} w0 \\ w1 \\ w2 \\ w3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

2. Find the optimal number of NMOS inverters to be cascaded so as to drive load capacitance of 0.54 pF off-chip capacitive load such that the total delay is minimized. Given that $1\Box C_g = 0.01$ pF. Give the cascaded structure with L:W ratios shown. Find the overall delay.

3. Explain the steps involved in photolithography. Perform the lithography using positive and negative photoresist with mask shown in Fig. Show the cross section view along 1-1' line

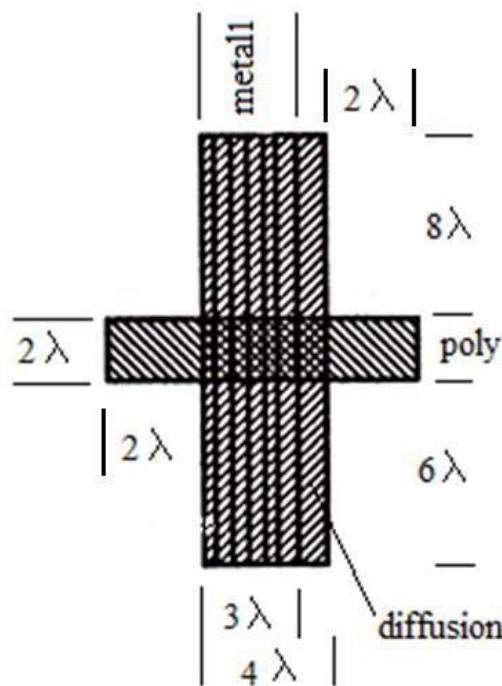


4. Given that $\lambda = 2.5 \mu\text{m}$

(i) A 3λ wide metal path crosses a 2λ wide polysilicon path at right angles. The layers are separated by a $0.5 \mu\text{m}$ thick layer of silicon dioxide. Find the capacitances associated with metal layer.

(ii) The polysilicon layer in turn crosses a 4λ wide diffusion region at right angles to form a transistor. Find the capacitance associated with poly layer.

TABLE 4.2 Typical area capacitance values for MOS circuits



Capacitance	Value in $\text{pF} \times 10^{-4}/\mu\text{m}^2$ (Relative values in brackets)		
	5 μm	2 μm	1.2 μm
Gate to channel	4 (1.0)	8 (1.0)	16 (1.0)
Diffusion (active)	1 (0.25)	1.75 (0.22)	3.75 (0.23)
Polysilicon* to substrate	0.4 (0.1)	0.6 (0.075)	0.6 (0.038)
Metal 1 to substrate	0.3 (0.075)	0.33 (0.04)	0.33 (0.02)
Metal 2 to substrate	0.2 (0.05)	0.17 (0.02)	0.17 (0.01)
Metal 2 to metal 1	0.4 (0.1)	0.5 (0.06)	0.5 (0.03)
Metal 2 to polysilicon	0.3 (0.075)	0.3 (0.038)	0.3 (0.018)

Notes: Relative value = specified value/gate to channel value for that technology.

*Poly. 1 and Poly. 2 are similar (also silicides where used).

- the structures shown when realized in 5 μm technology are capable of driving loads of 2 pF with 5 nsec rise-time.