University of Southern California

Viterbi School of Engineering

EE577A

VLSI System Design

Introduction

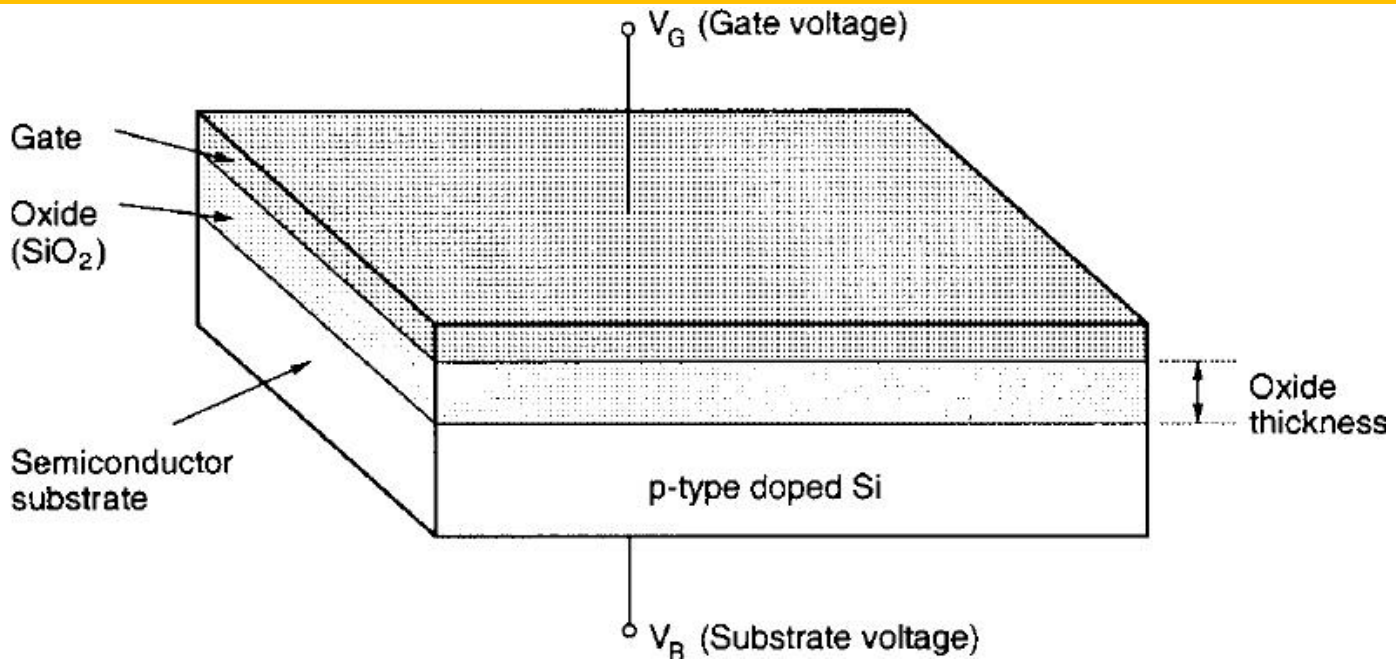
References: Professor Massoud Pedram's lecture slides, books listed in the syllabus, and online resources

Shahin Nazarian

Spring 2013

EE577A Objectives

Two-Terminal MOS Structure



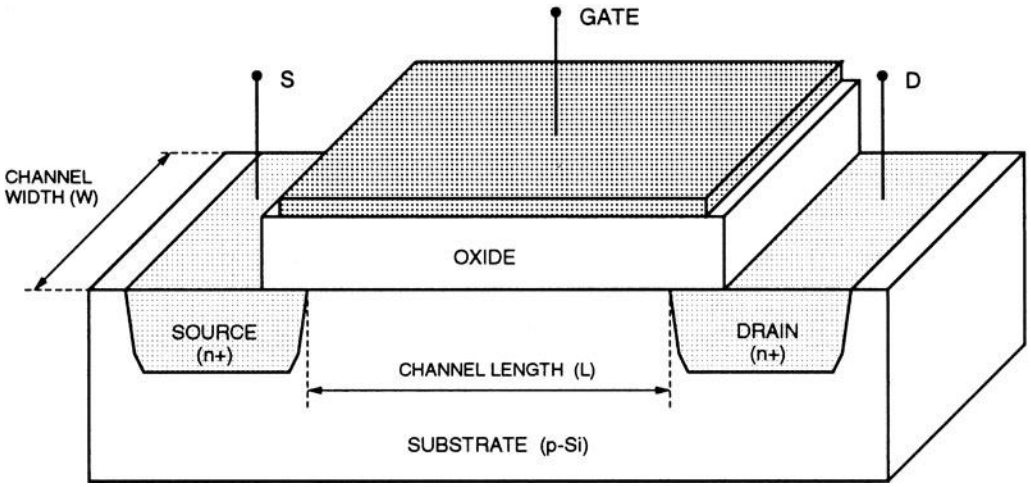
Under thermal equilibrium (Mass Action Law)

$$np = n_i^2 \; (n_i \approx 1.45 \times 10^{10}\, cm^{-3})$$

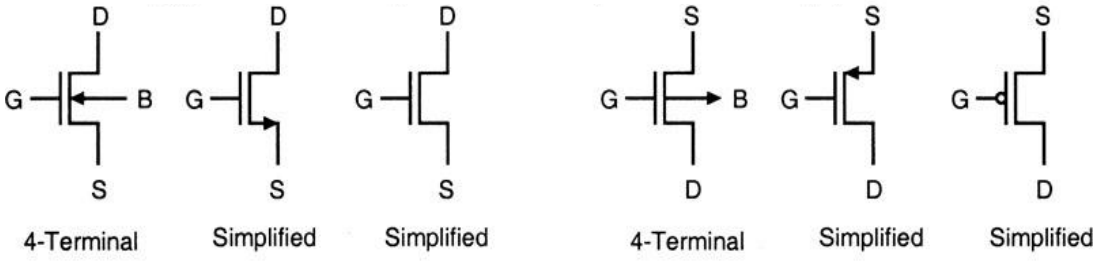
If substrate is uniformly doped @ *N_A (Typically 10^{14} to 10^{16})*

$$p_{p0} = N_A \;\; \rightarrow \;\; n_{p0} \approx \frac{n_i^2}{N_A}$$

Structure of MOS Transistor

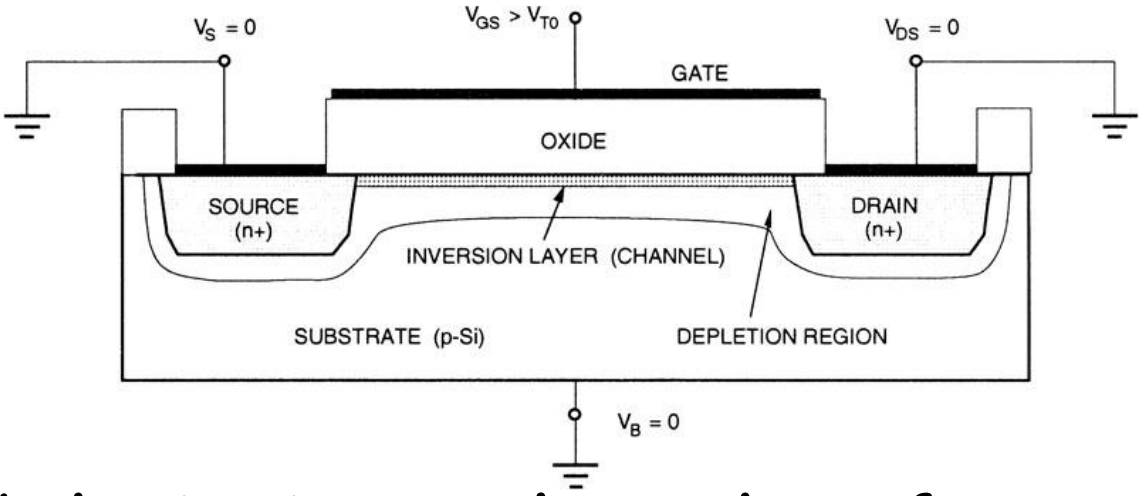


- The inverted surface layer is essential for current conduction. The two n+ regions are added as the current conducting terminals
- Conventionally all terminal voltages are defined wrt V_S



n-channel MOSFET p-channel MOSFET

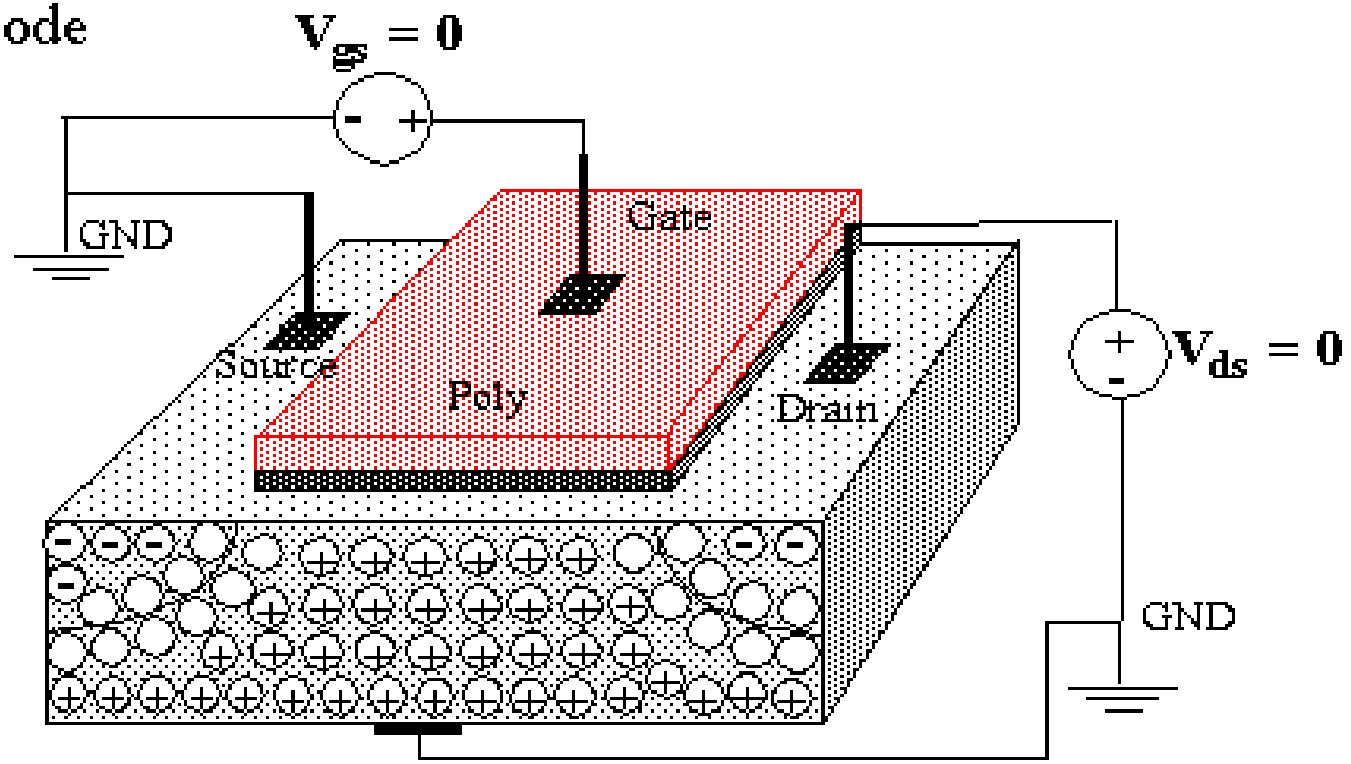
Threshold Voltage



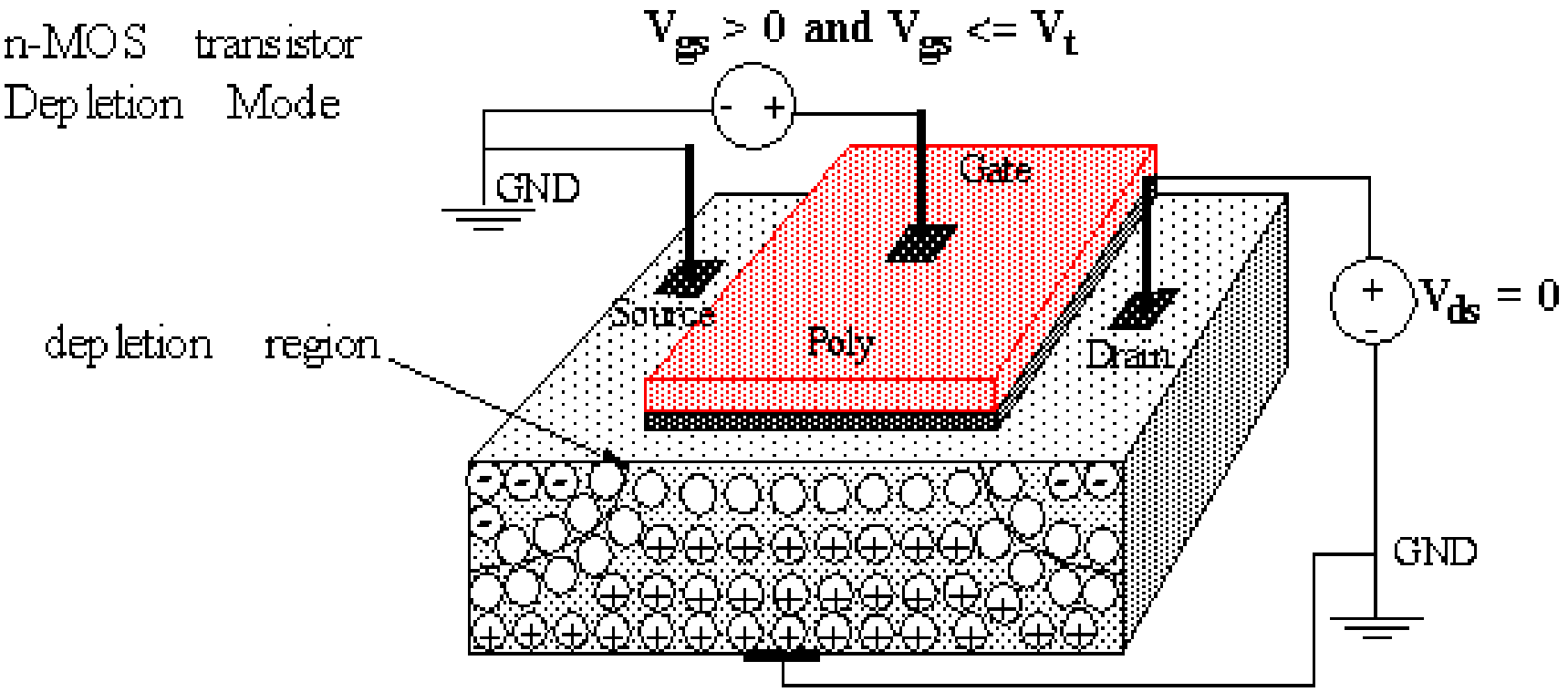
- Channel behavior is exactly as that of a metal-gate-oxide: increase V_G and surfaces is inverted as soon as surface potential reaches a critical threshold voltage V_{T0}
- Note that increasing V_{GS} beyond V_{T0} will not increase the depletion region depth
- V_{DS} would then allow current flow

nMOS – Accumulation

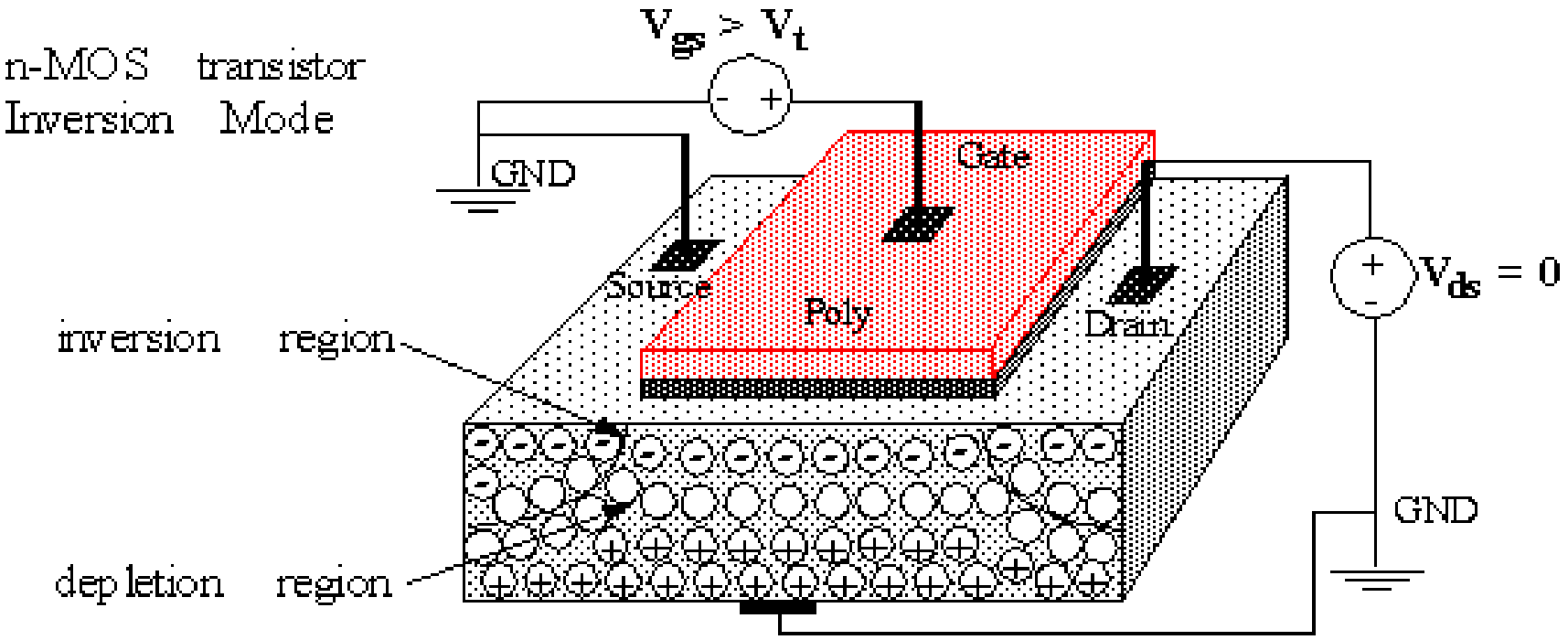
n-MOS transistor
Accumulation Mode



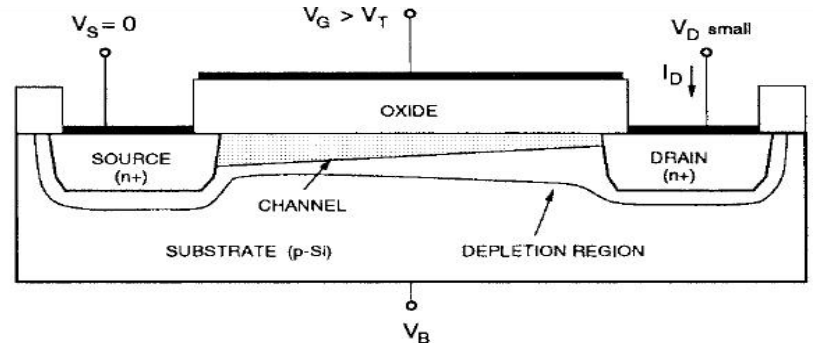
nMOS - Depletion



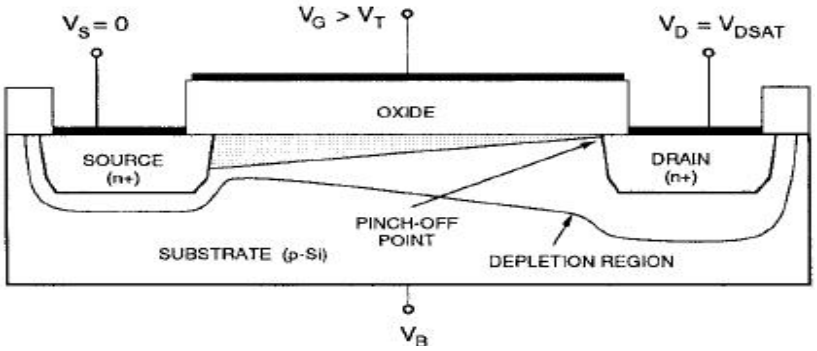
nMOS – Inversion



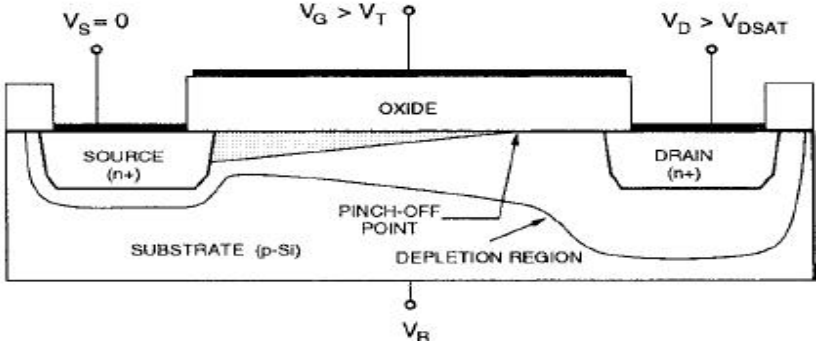
nMOS Linear and Saturation Operations



Operating in the linear region



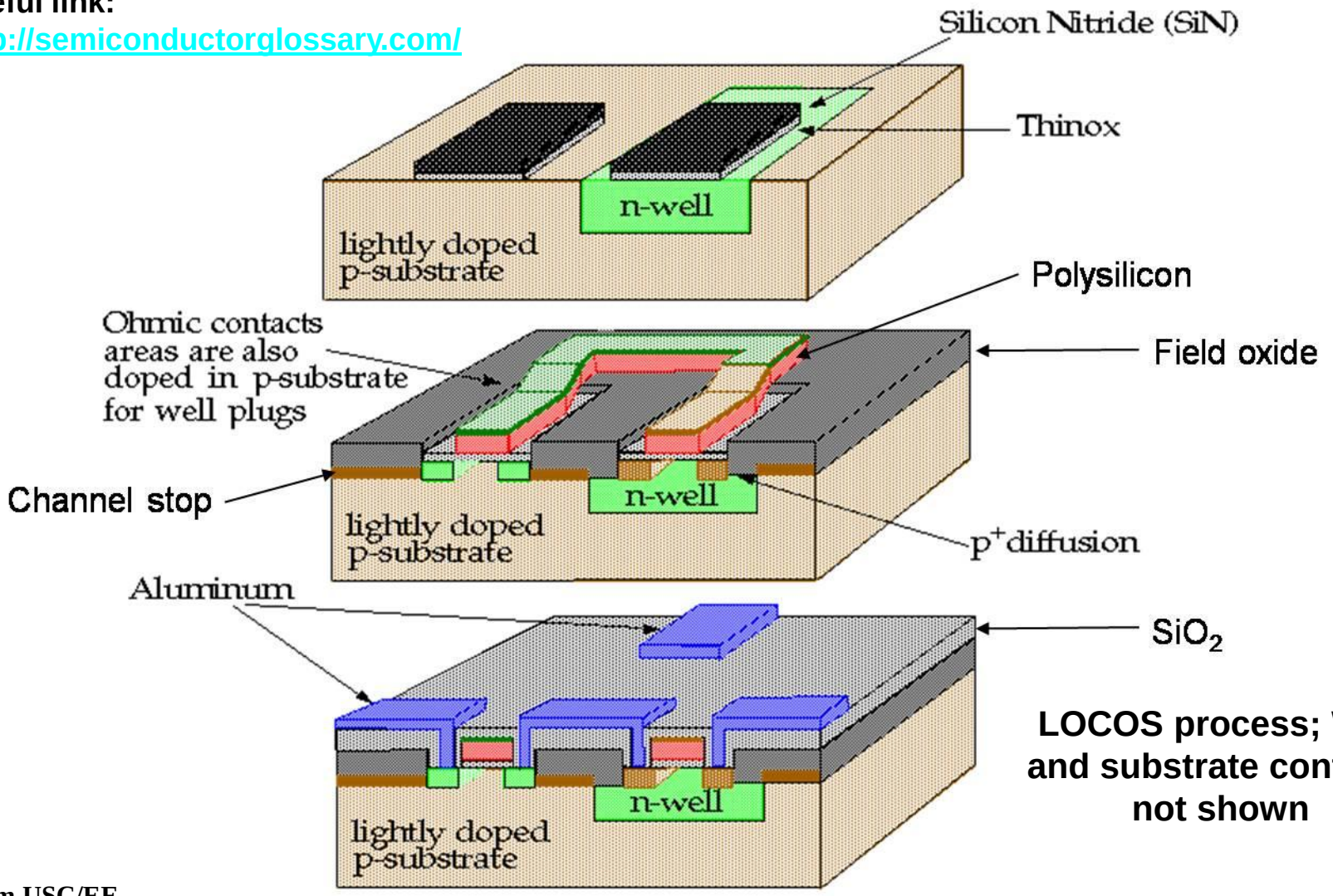
Operating at the edge of saturation



Operating beyond saturation

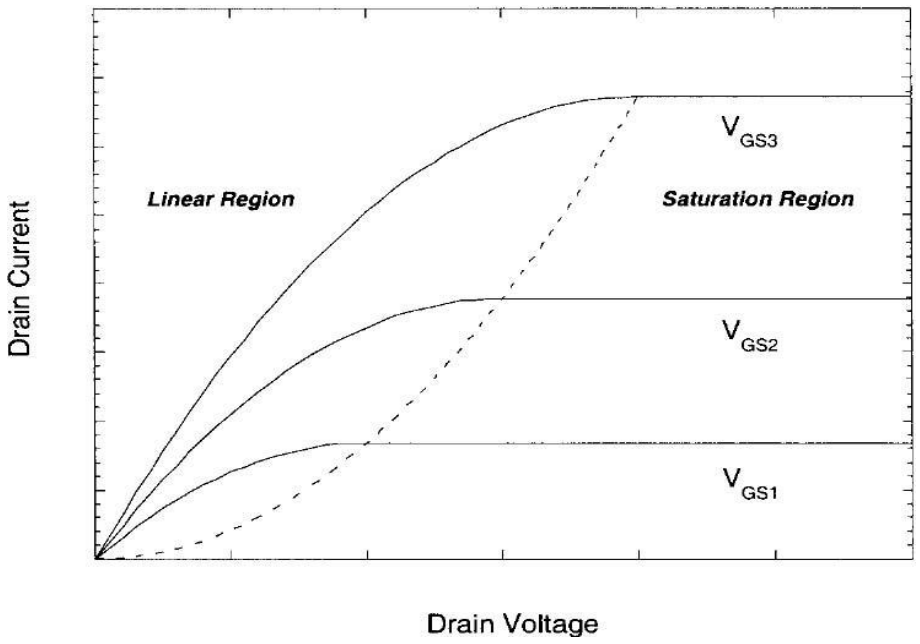
CMOS Processing Technology Steps

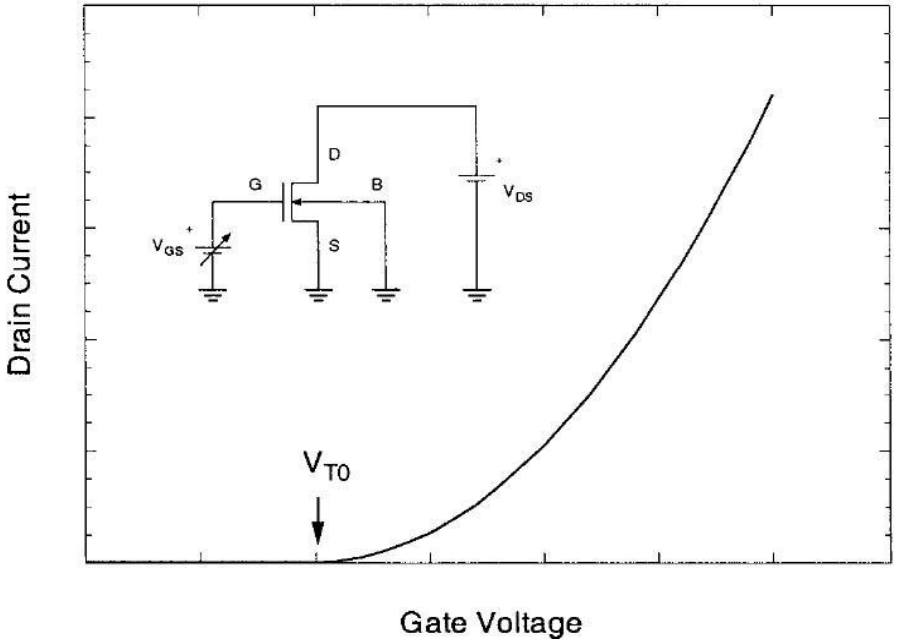
Useful link:
http://semiconductorglossary.com/



LOCOS process; Well and substrate contacts not shown

NMOS I_D-V_{DS} and I_D-V_{GS} Curves





$$I_D(lin) = \frac{\mu_n C_{ox}}{2}\frac{W}{L}\left(2(V_{GS} - V_{T0})V_{DS} - V_{DS}^2\right)$$

$$I_D(sat) = \frac{\mu_n C_{ox}}{2}\frac{W}{L}\left(V_{GS} - V_{T0}\right)^2$$

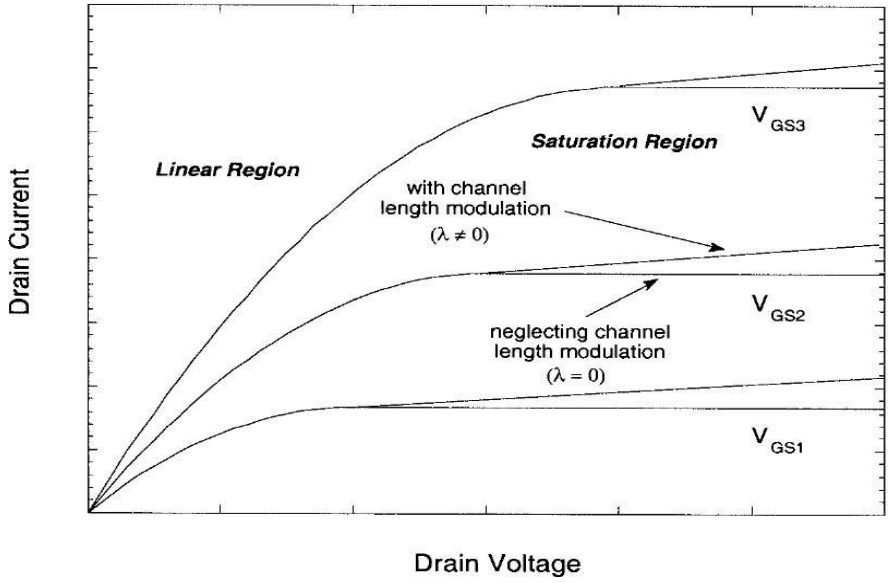
Channel Length Modulation

- The channel is pinched off under saturation condition; The pinch-off point moves from the drain end towards source as V_{DS} increases:

$$I_{Dsat} \propto \frac{1}{L-\Delta L} \cong \frac{1}{L}\left(1+\frac{\Delta L}{L}\right) \quad \text{where} \quad \frac{\Delta L}{L} = \lambda V_{DS}$$

- We can then write the saturation current as:

$$I_D(sat) = \frac{\mu_n C_{ox}}{2}\frac{W}{L}\left(V_{GS}-V_{T0}\right)^2\left(1+\lambda V_{DS}\right)$$



Substrate Bias Effect

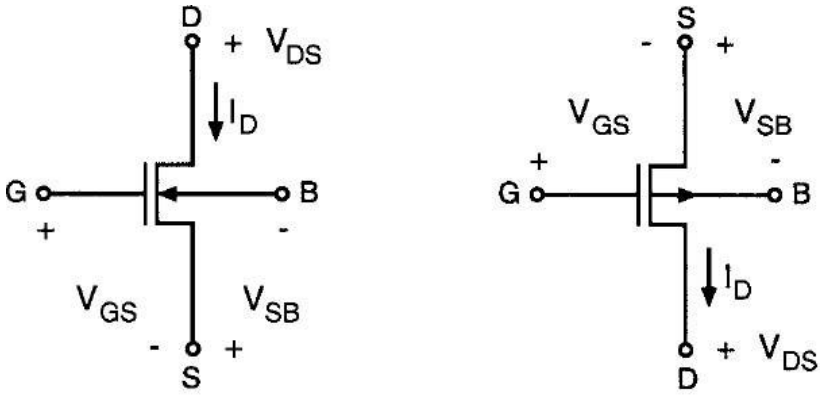
- Applying a substrate voltage changes the threshold voltage

$$V_T(V_{SB}) = V_{T0} + \gamma\left(\sqrt{|2\phi_F| + V_{SB}} - \sqrt{|2\phi_F|}\right)$$

- We can simply replace the threshold voltage terms in linear-mode and saturation-mode current equations with $V_T(V_{SB})$

$$I_D(lin) = \frac{\mu_n C_{ox}}{2}\frac{W}{L}\left(2\left(V_{GS} - V_T(V_{SB})\right)V_{DS} - V_{DS}^2\right)$$

$$I_D(sat) = \frac{\mu_n C_{ox}}{2}\frac{W}{L}\left(V_{GS} - V_T(V_{SB})\right)^2\left(1 + \lambda V_{DS}\right)$$



MOSFET Current-Voltage Equations (a.k.a. Shichman-Hodges Equations)

- **nMOS transistor, with** $k_n = k_n' \frac{W}{L} = \mu_n C_{ox} \frac{W}{L}$

$$I_D(cutoff) = 0 \quad V_{GS} < V_T$$

$$I_D(lin) = \frac{k_n}{2}\left(2\left(V_{GS} - V_T(V_{SB})\right)V_{DS} - V_{DS}^2\right) \quad V_{GS} \geq V_T, V_{DS} < V_{GS} - V_T$$

$$I_D(sat) = \frac{k_n}{2}\left(V_{GS} - V_T(V_{SB})\right)^2\left(1 + \lambda V_{DS}\right) \quad V_{GS} \geq V_T, V_{DS} \geq V_{GS} - V_T$$

- **pMOS transistor, with** $k_p = k_p' \frac{W}{L} = \mu_p C_{ox} \frac{W}{L}$

$$I_D(cutoff) = 0 \quad V_{GS} > V_T$$

$$I_D(lin) = \frac{k_p}{2}\left(2\left(V_{GS} - V_T(V_{SB})\right)V_{DS} - V_{DS}^2\right) \quad V_{GS} \leq V_T, V_{DS} > V_{GS} - V_T$$

$$I_D(sat) = \frac{k_p}{2}\left(V_{GS} - V_T(V_{SB})\right)^2\left(1 + \lambda V_{DS}\right) \quad V_{GS} \leq V_T, V_{DS} \leq V_{GS} - V_T$$

Example

- For an n-channel MOS transistor with μ_n=600 cm²/V.s, C_{ox}=7x10⁻⁸ F/cm², W=20 µm, L=2 µm and V_{T0}=1.0 V, plot the relationship between I_D and the terminal voltages

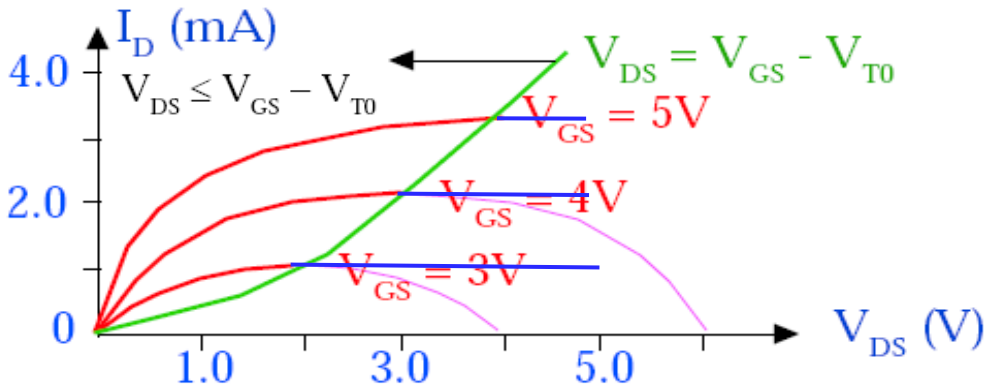
$$k = 600 cm^2/V.s \times 7\times10^{-8} F/cm^2 \times \frac{20\mu m}{2\mu m} = 0.42 \quad mA/V^2$$

For Linear operation:

$$V_{GS} \geq V_{T0}$$

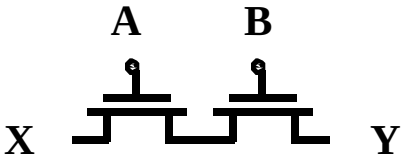
$$V_{GD} = V_{GS} - V_{DS} \geq V_{T0} \quad \Rightarrow V_{DS} \leq V_{GS} - V_{T0}$$

$$I_D = 0.21 \quad mA/V^2 \left[2(V_{GS} - 1)V_{DS} - V_{DS}^2 \right]$$

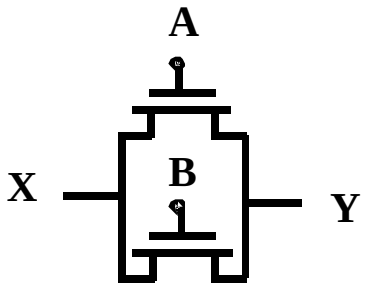


NMOS Transistors in Series/Parallel Connection

- Transistors can be thought as a switch controlled by its gate signal
- NMOS switch closes when switch control input is high



Y = X if A and B

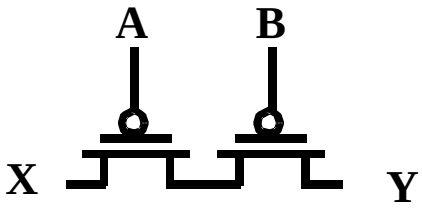


Y = X if A OR B

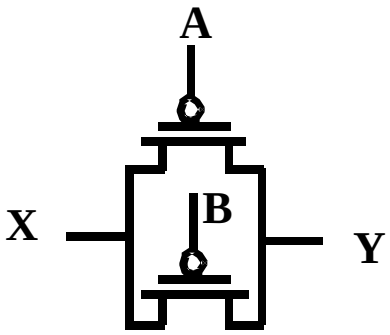
NMOS Transistors pass a "strong" 0 but a "weak" 1

PMOS Transistors in Series/Parallel Connection

- PMOS switch closes when switch control input is low



$Y = X$ if $\overline{A}$ AND $\overline{B} = \overline{A + B}$

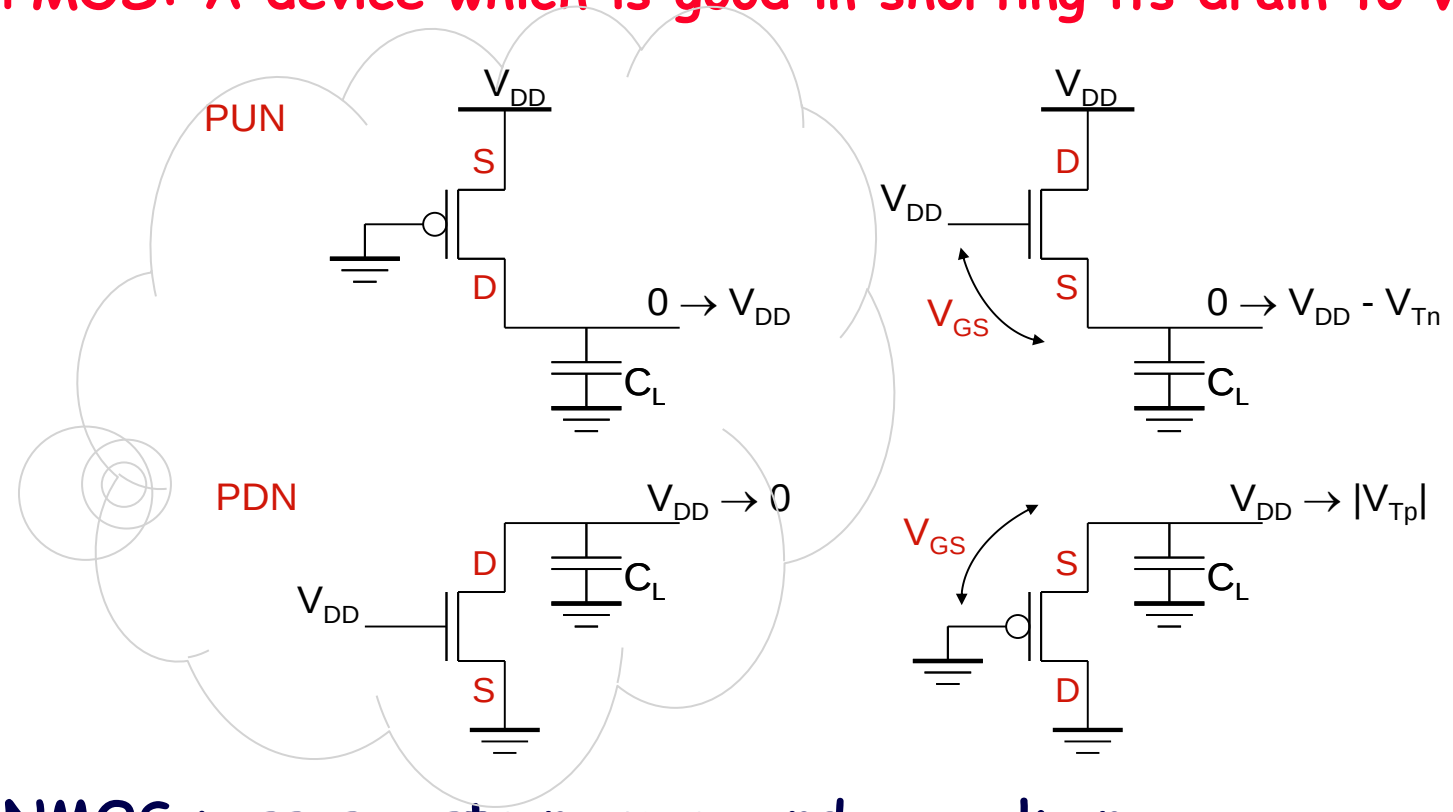


$Y = X$ if $\overline{A}$ OR $\overline{B} = \overline{AB}$

PMOS Transistors pass a "strong" 1 but a "weak" 0

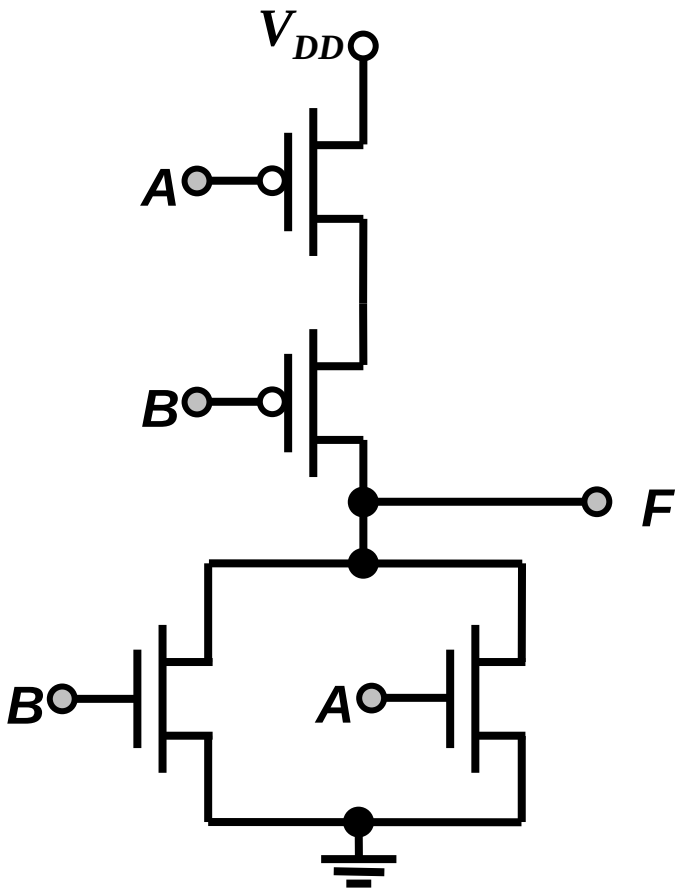
Threshold Drops

- NMOS: A device which is good in shorting its drain to the ground
- PMOS: A device which is good in shorting its drain to V_{DD}



- NMOS passes a strong zero and a weak one
- PMOS passes a strong one and a weak zero
- Note: NMOS gate = 0, PMOS gate = 1 are in high Impedance

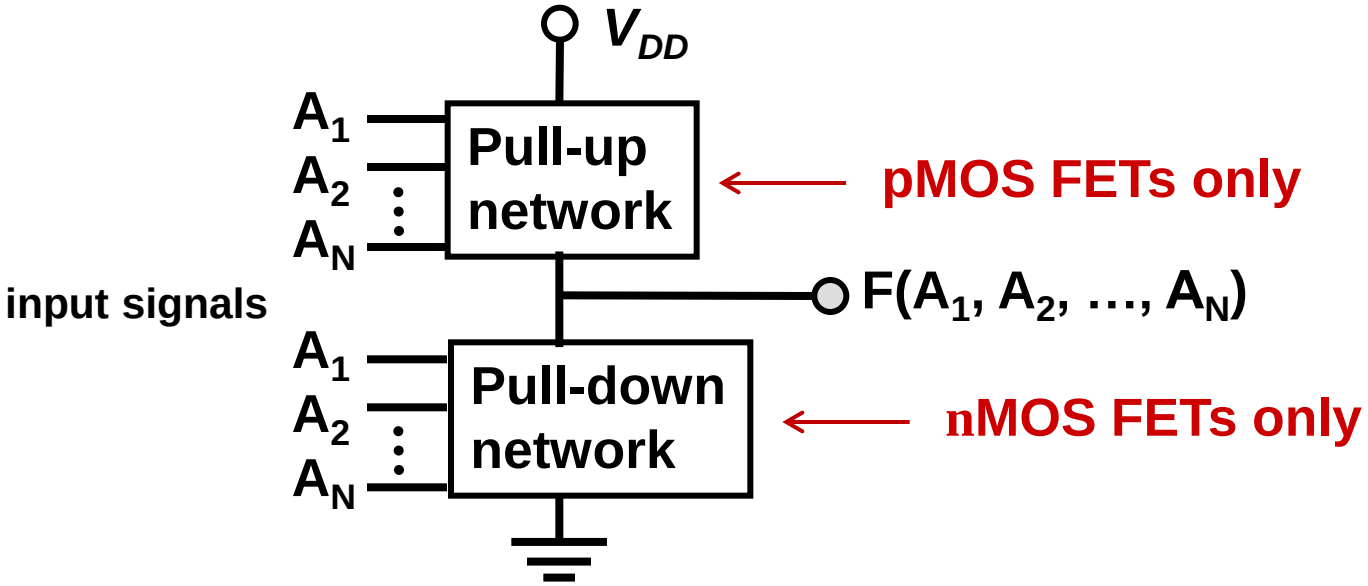
CMOS NOR2 Gate



A	B	F
0	0	1
0	1	0
1	0	0
1	1	0

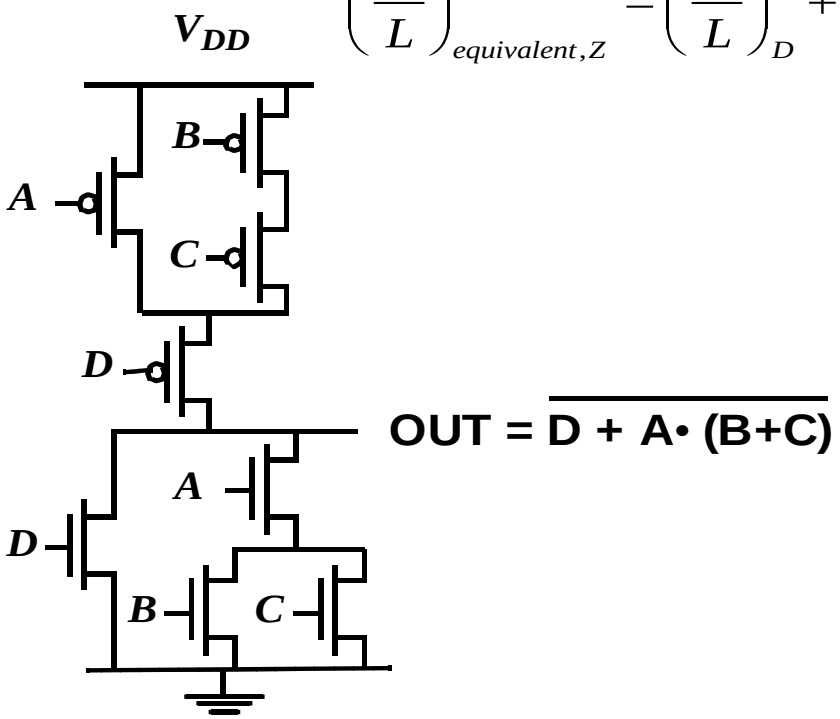
Complex (Static) CMOS Gates

- In CMOS logic gates, nMOS FETs are used to connect the output to GND, whereas pMOS FETs are used to connect the output to V_{DD}
  - An nMOS FET functions as a pull-down device when it is turned on (gate voltage = V_{DD})
  - A pMOS FET functions as a pull-up device when it is turned on (gate voltage = GND)

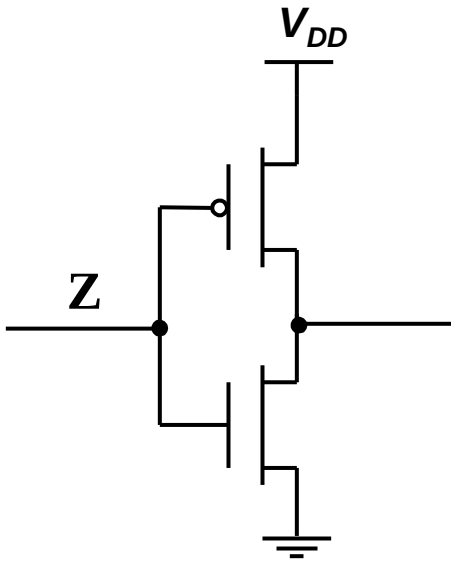


Reduction of Complex Gates to Inverters

$$\left(\frac{W}{L}\right)_{equivalent,Z} = \left(\frac{W}{L}\right)_D + \frac{1}{\frac{1}{\left(\frac{W}{L}\right)_A} + \frac{1}{\left(\frac{W}{L}\right)_B + \left(\frac{W}{L}\right)_C}}$$



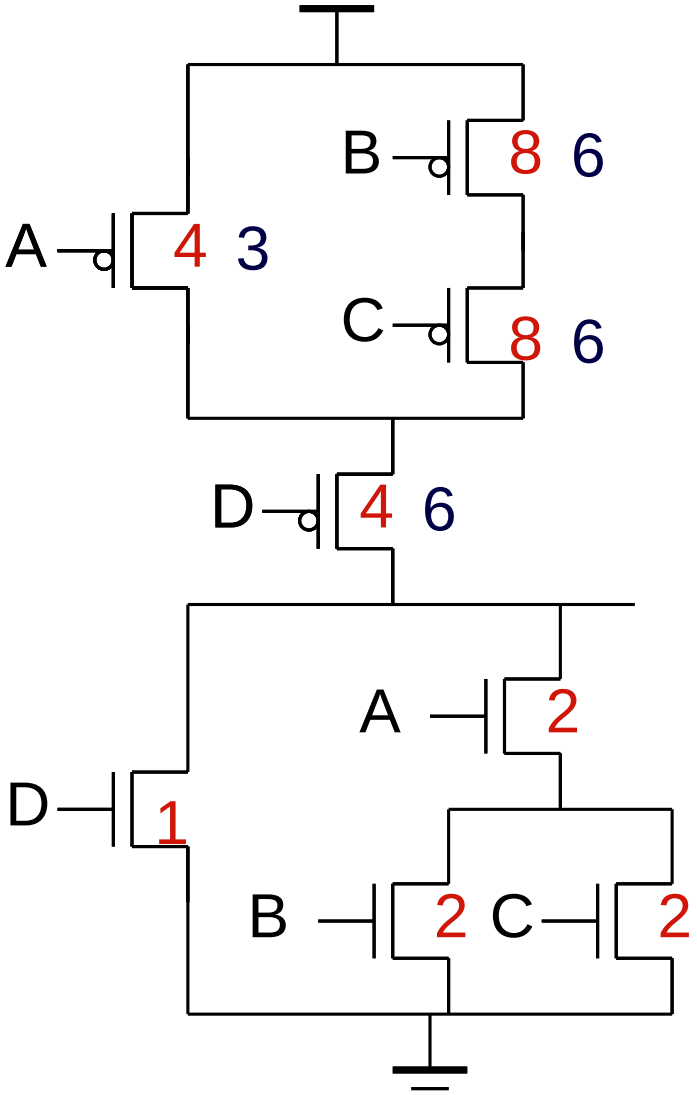
$$OUT = \overline{D + A \cdot (B+C)}$$



- Note that reduction for worst case t_{pHL} is:

$$\left(\frac{W}{L}\right)_{worstcasedelay,Z} = MIN\left(\left(\frac{W}{L}\right)_D, \frac{1}{\frac{1}{\left(\frac{W}{L}\right)_A} + \frac{1}{\left(\frac{W}{L}\right)_B}}, \frac{1}{\frac{1}{\left(\frac{W}{L}\right)_A} + \frac{1}{\left(\frac{W}{L}\right)_C}}\right)$$

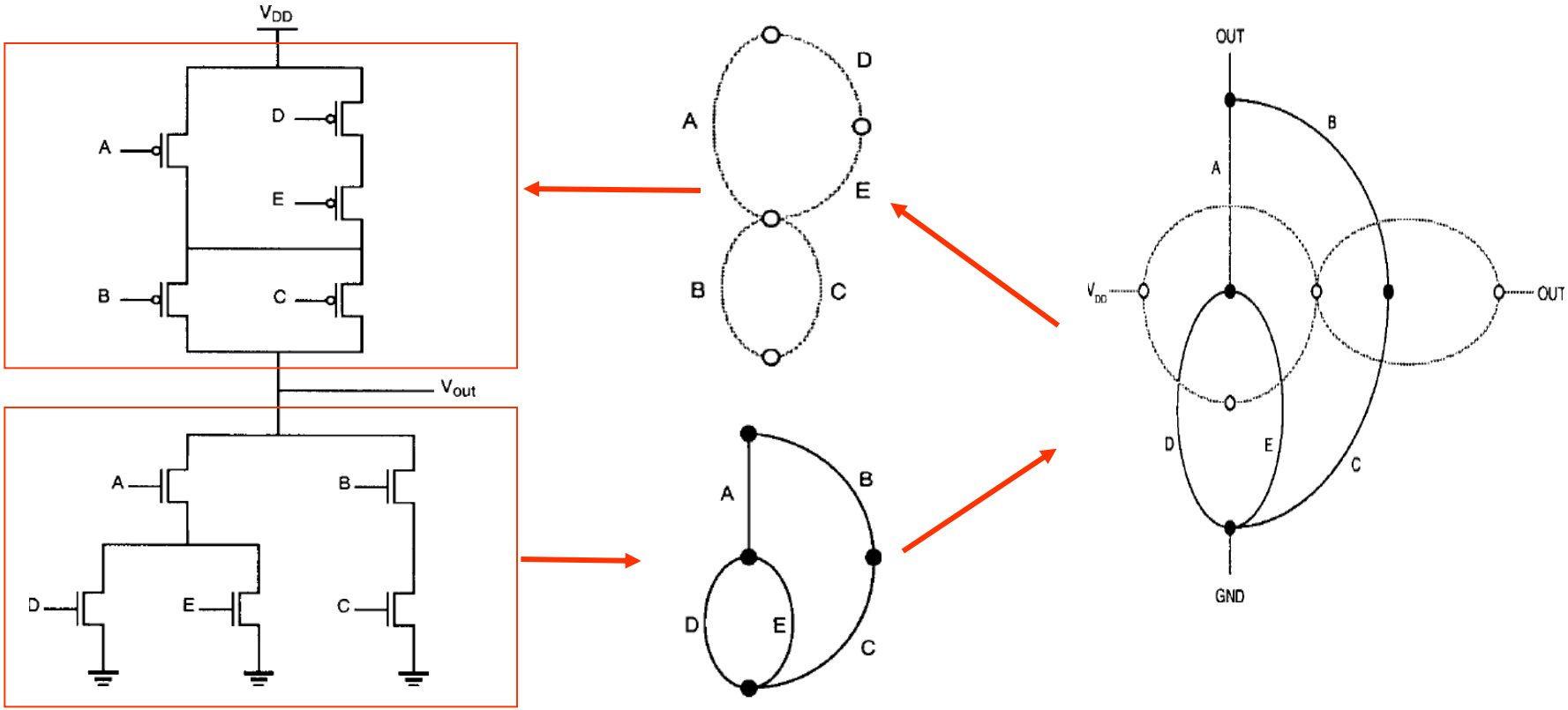
Sizing – Complex CMOS Gates



$$OUT = \overline{D + A \bullet (B + C)}$$

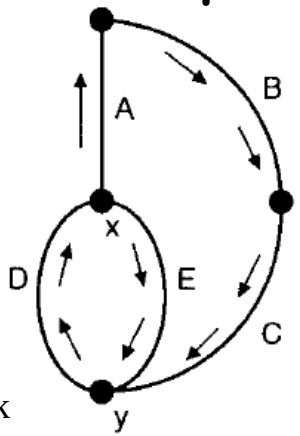
Layout Design – Stick Diagram

- Construction of the dual pull-up graph from the pull-down graph



Stick Diagram (Cont.)

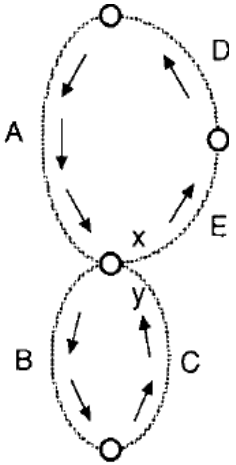
- The number of diffusion breaks can be minimized by changing the ordering of the polysilicon columns
- A simple method for finding the optimum gate ordering is the Euler-path approach
  - Find a common Euler path for both pull-down and pull-up graphs
  - The polysilicon columns can be arranged according to the sequence (in Euler-path)
- Diffusion will be unbroken if identically labeled Euler paths can be found for the p and n trees



nMOS network

Common Euler Path

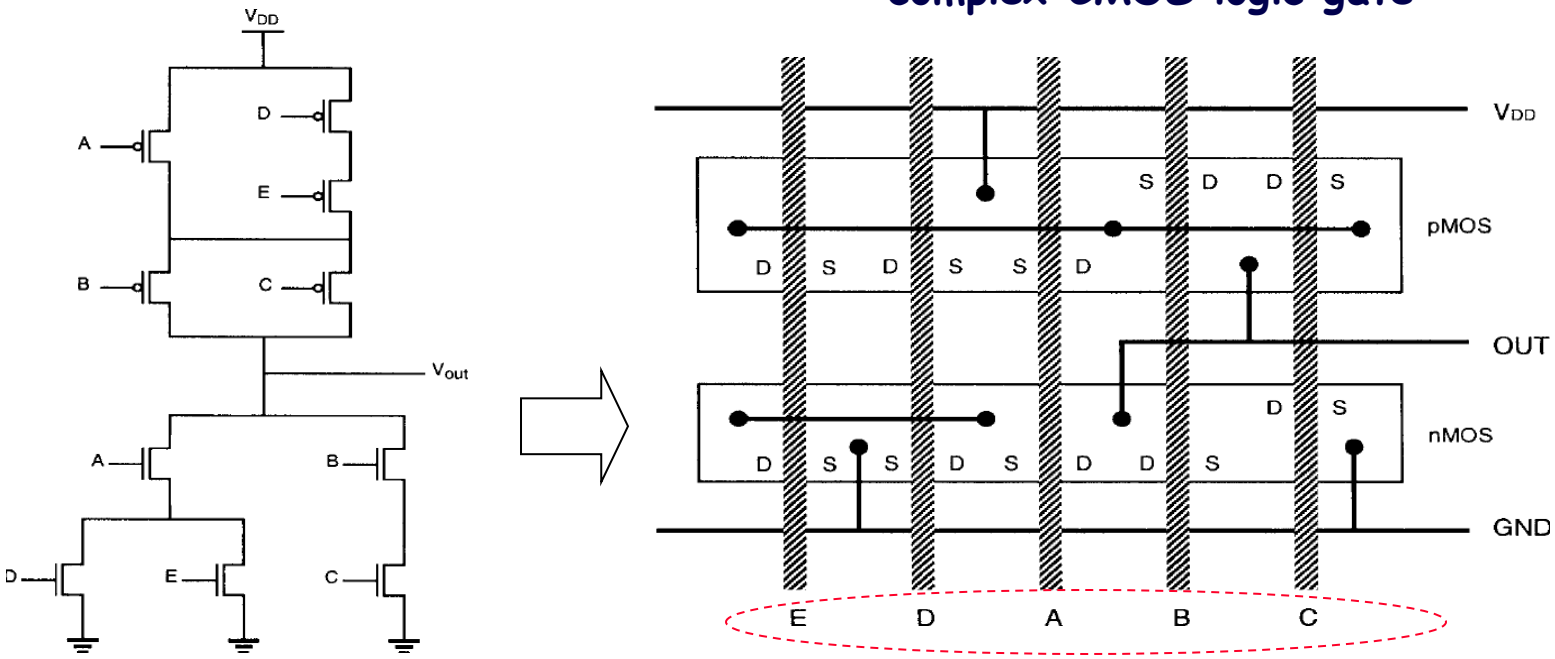
E – D – A – B – C



pMOS network

Layout of Complex Gates (Cont.)

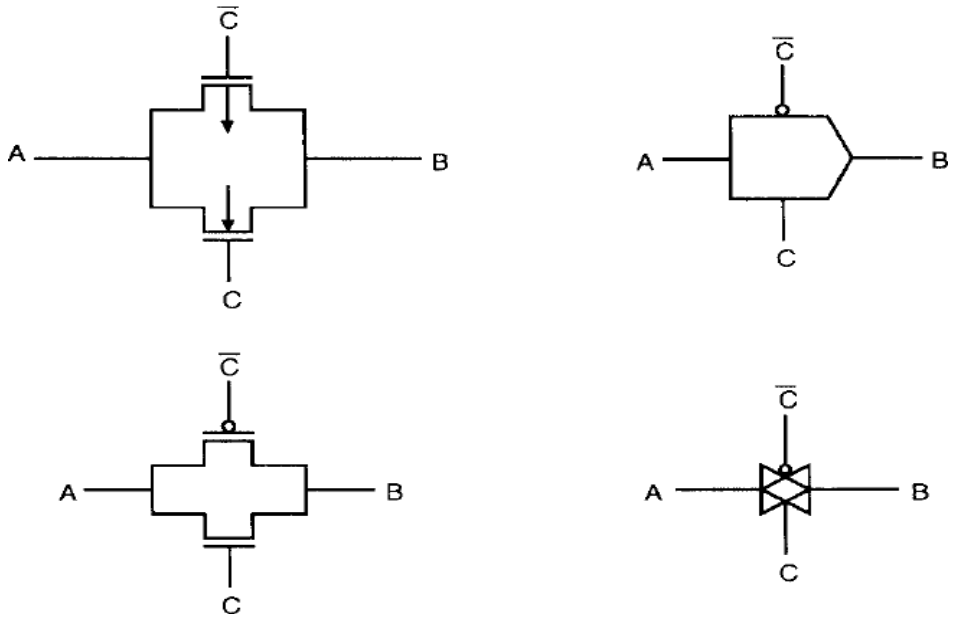
- Optimize stick-diagram layout of the complex CMOS logic gate



- The advantages of this new layout are more compact layout area, simple routing of signals, and consequently, lower parasitic capacitance

CMOS Transmission Gates (Pass Gates)

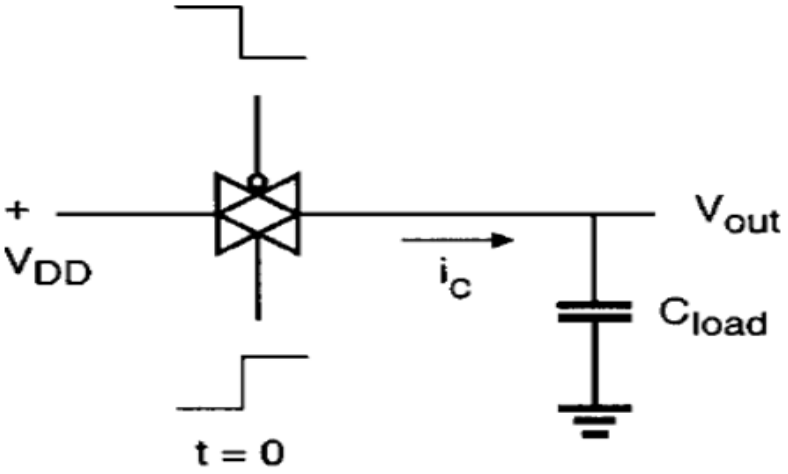
- The CMOS transmission gate (TG) consists of one NMOS and one PMOS transistor, connected in parallel
- The CMOS TG operates as a bidirectional switch between nodes A and B, which is controlled by signal C

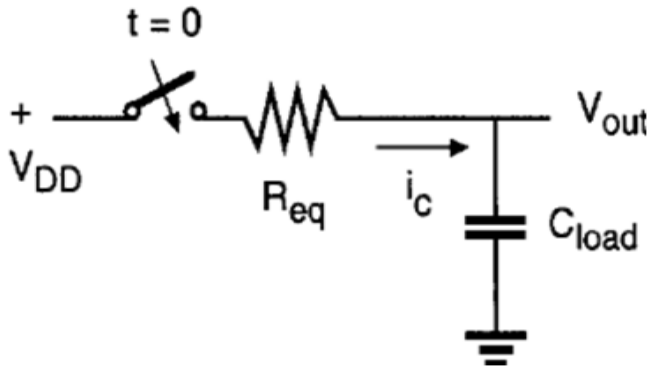


Four different representations of the CMOS transmission gate

Model of a CMOS TG

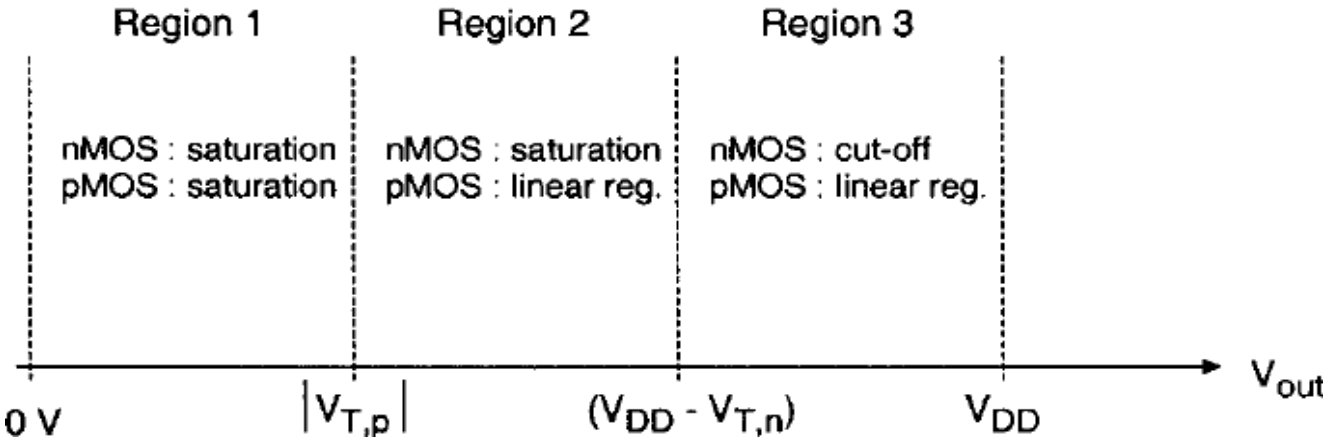
- A CMOS pass gate which is turned on by a logic-high control signal can be replaced by its **simple equivalent resistance** for dynamic analysis





Pass Gate Resistance Calculation

- Unlike the nMOS transistor, the pMOS remains ON regardless of V_{out}
- If nMOS is ON, it would be in saturation region
- We can identify three operating regions for the CMOS transmission gate, depending on the output voltage level



- The total current flowing through the transmission gate is the sum of the nMOS drain and the pMOS drain current:

$$I_D = I_{DS,n} + I_{SD,p}$$

Pass Gate Resistance Calculation (Cont.)

- At this point, we devise an equivalent resistance for each transistor in this structure as follows

$$R_{eq,n} = \frac{V_{DD} - V_{out}}{I_{DS,n}}$$

$$R_{eq,p} = \frac{V_{DD} - V_{out}}{I_{SD,p}}$$

- The total equivalent resistance of the CMOS TG will be the parallel equivalent of these two resistances

- We will next calculate the equivalent resistance values for the three operating regions of the transmission gate

Pass Gate Resistance Calculation (Cont.)

- In region 1, $V_{out} < |V_{T,p}|$
  - pMOS and nMOS transistors both in saturation

$$R_{eq,n} = \frac{2(V_{DD} - V_{out})}{k_n (V_{DD} - V_{out} - V_{T,n})^2} \qquad R_{eq,p} = \frac{2(V_{DD} - V_{out})}{k_p (V_{DD} - |V_{T,p}|)^2}$$

  - V_{SB} of NMOS transistor is equal to the output voltage, while that for pMOS transistor is equal to zero, therefore body bias for the nMOS transistor must be taken into account

- In region 2, $|V_{T,p}| < V_{out} < V_{DD} - V_{T,n}$
  - pMOS in linear and nMOS in saturation

$$R_{eq,n} = \frac{2(V_{DD} - V_{out})}{k_n (V_{DD} - V_{out} - V_{T,n})^2} \qquad R_{eq,p} = \frac{2}{k_p [2(V_{DD} - |V_{T,p}|) - (V_{DD} - V_{out})]}$$

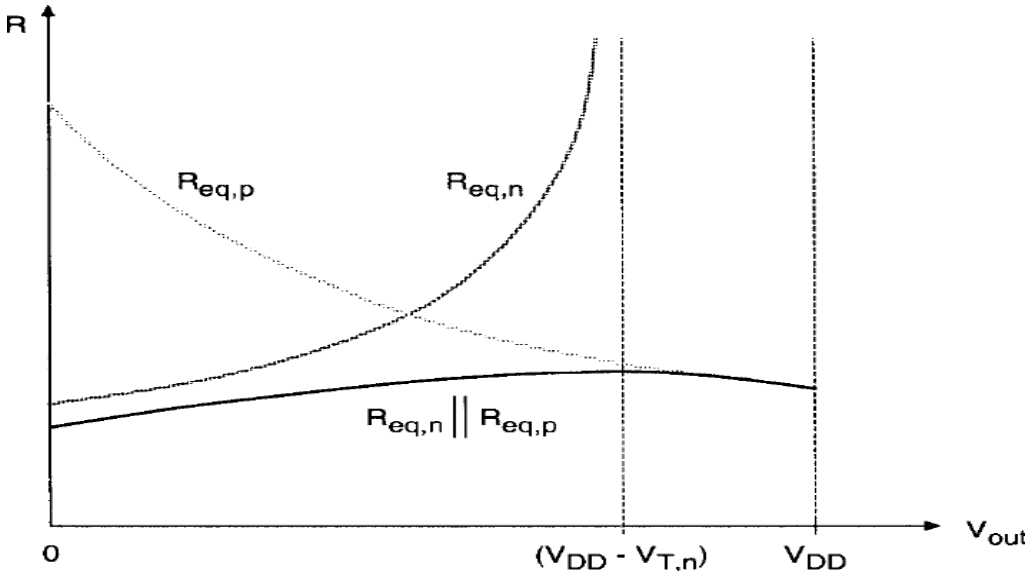
- In region 3, the output voltage is $V_{out} > V_{DD} - V_{T,n}$
  - pMOS in linear and nMOS in cutoff

$$R_{eq,n} = \infty \qquad R_{eq,p} = \frac{2}{k_p [2(V_{DD} - |V_{T,p}|) - (V_{DD} - V_{out})]}$$

Pass Gate Resistance Calculation (Cont.)

- While individual $R_{eq,n}$ and $R_{eq,p}$ are strongly dependent on V_{out}, the total R_{eq} remains relatively constant, i.e., its value is almost independent of V_{out} (a quite desirable properly)

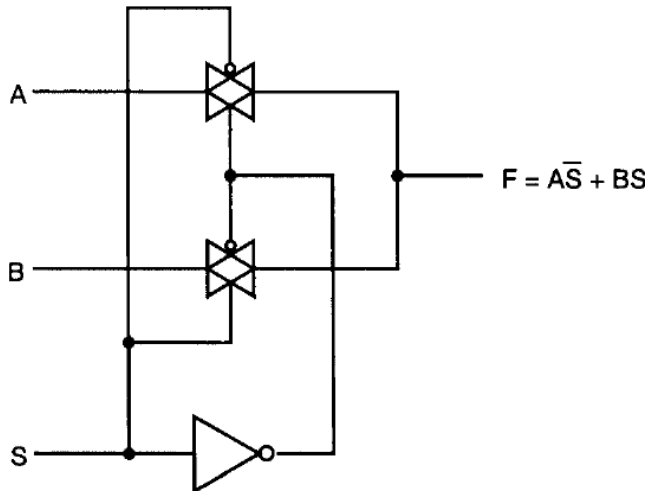
- A CMOS pass gate which is turned on by a logic-high control signal can be replaced by its simple equivalent resistance for dynamic analysis



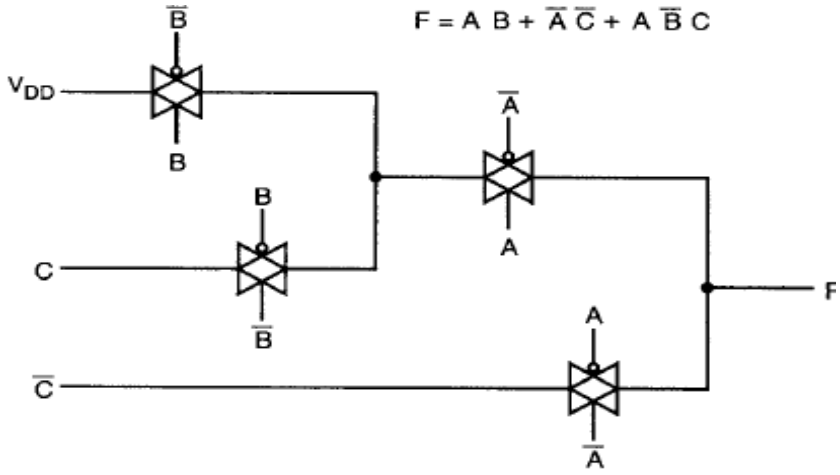
Equivalent resistance of the CMOS transmission gate

TG-based Cell Design

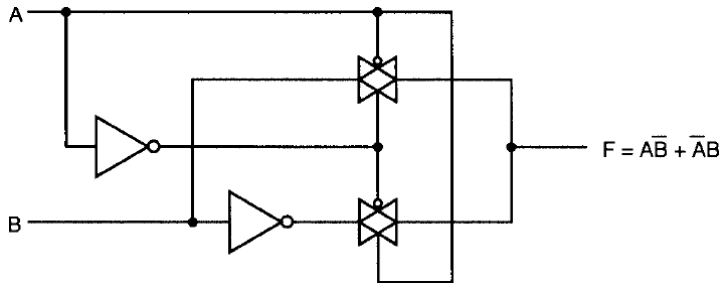
Two-input multiplexer



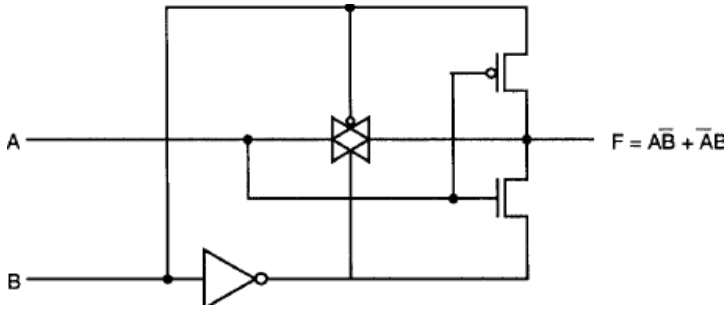
An arbitrary logic



- XOR



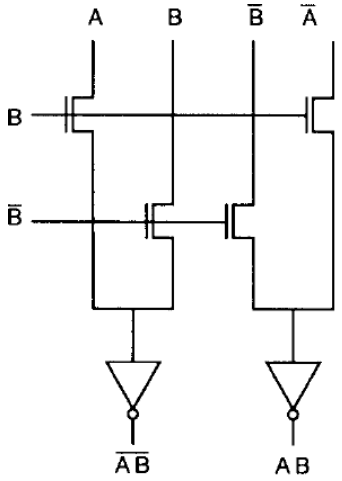
Eight-transistor CMOS TG implementation



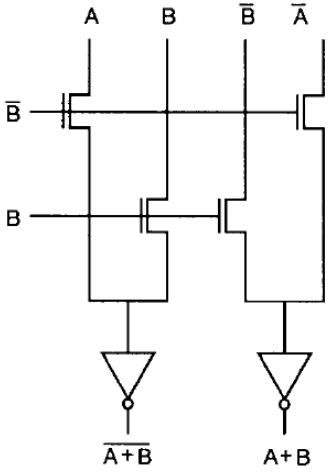
Six-transistor CMOS TG implementation

Complementary Pass-Transistor Logic

- The complexity of full CMOS pass-gate logic circuits can be reduced by adopting **Complementary Pass-transistor Logic** (**CPL**)
- The idea is to use a purely nMOS pass-transistor network for the logic operations, instead of a CMOS TG network
- Every input signal and its complement must be provided



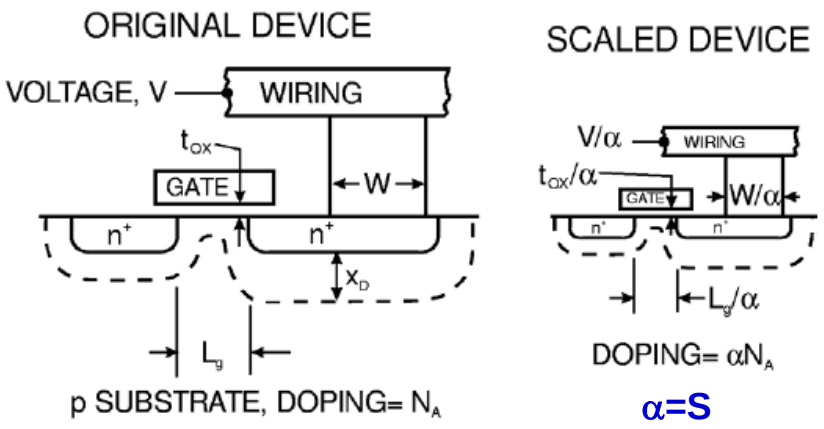
CPL NAND2 gate



CPL NOR2 gate

Full Scaling (Constant-Field Scaling)

Quantity	Before Scaling	After Scaling
Channel length	L	$L'=L/S$
Channel width	W	$W'=W/S$
Gate oxide thickness	t_{ox}	$t_{ox}'= t_{ox}/S$
Junction depth	x_j	$x_j'=x_j/S$
Power supply voltage	V_{DD}	$V_{DD}'=V_{DD}/S$
Threshold voltage	V_{T0}	$V_{T0}'= V_{T0}/S$
Doping densities	N_A & N_D	$N_A'=S.N_A$ & $N_D'= S.N_D$



Full Scaling (Cont.)

- I_D for linear and saturation modes:

$$I_D'(lin) = \frac{k_n'}{2}.\left[2.(V_{GS}' - V_T').V_{DS}' - V_{DS}'^2\right]$$

$$= \frac{S.k_n}{2}.\frac{1}{S^2}.\left[2.(V_{GS} - V_T).V_{DS} - V_{DS}^2\right] = \frac{I_D(lin)}{S}$$

$$I_D'(sat) = \frac{k_n'}{2}.(V_{GS}' - V_T')^2 = \frac{S.k_n}{2}.\frac{1}{S^2}.(V_{GS} - V_T)^2 = \frac{I_D(sat)}{S}$$

- For power dissipation of the transistor, we obtain:

$$P' = I_D'.V_{DS}' = \frac{1}{S^2}.I_D.V_{DS} = \frac{P}{S^2}$$

- Device area reduction by S^2, so *power density* per unit area remains unchanged for the scaled device
- $C_g = WLC_{ox}$ scales down by S, therefore faster charge-up charge-down for the scaled device
  - Reduced dimensions would result in reduction of various parasitic caps and resistances

Technology Trend

- http://www.itrs.net/Links/2011ITRS/Home2011.htm