JOINT INSTITUTE

交大密西根学院

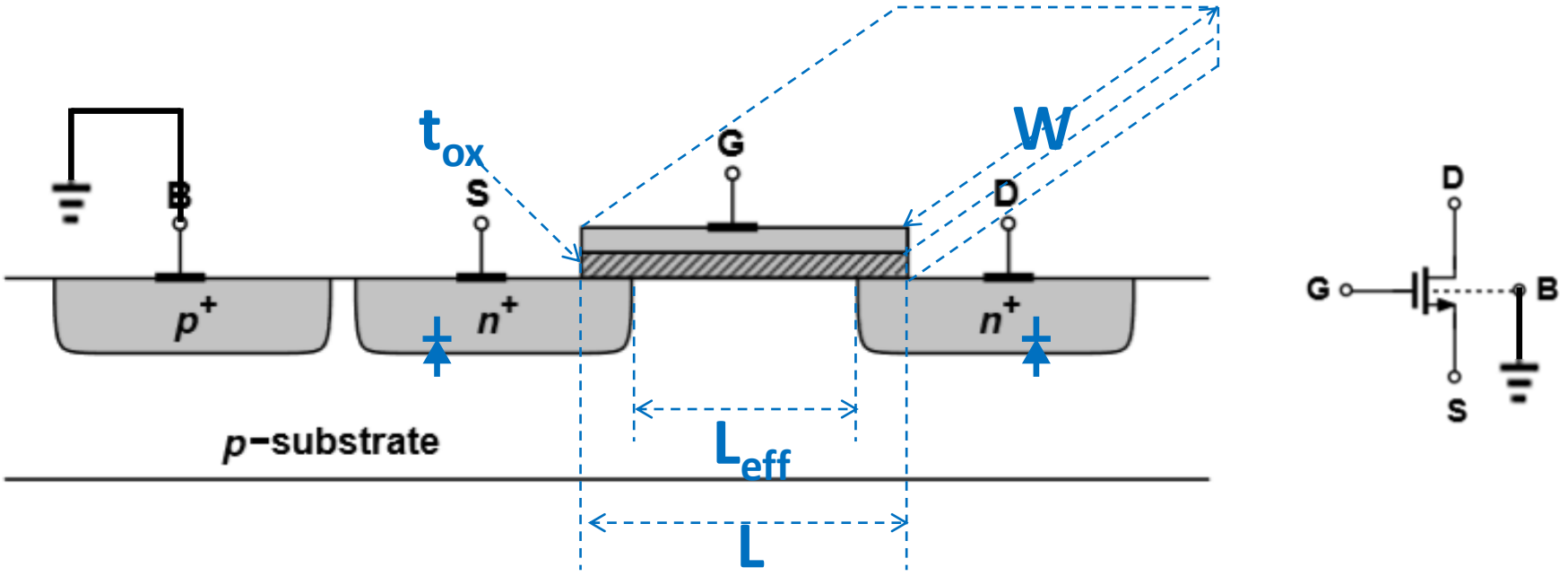
FET

Ve311 Electronic Circuits (Summer 2019)

Dr. Chang-Ching Tu

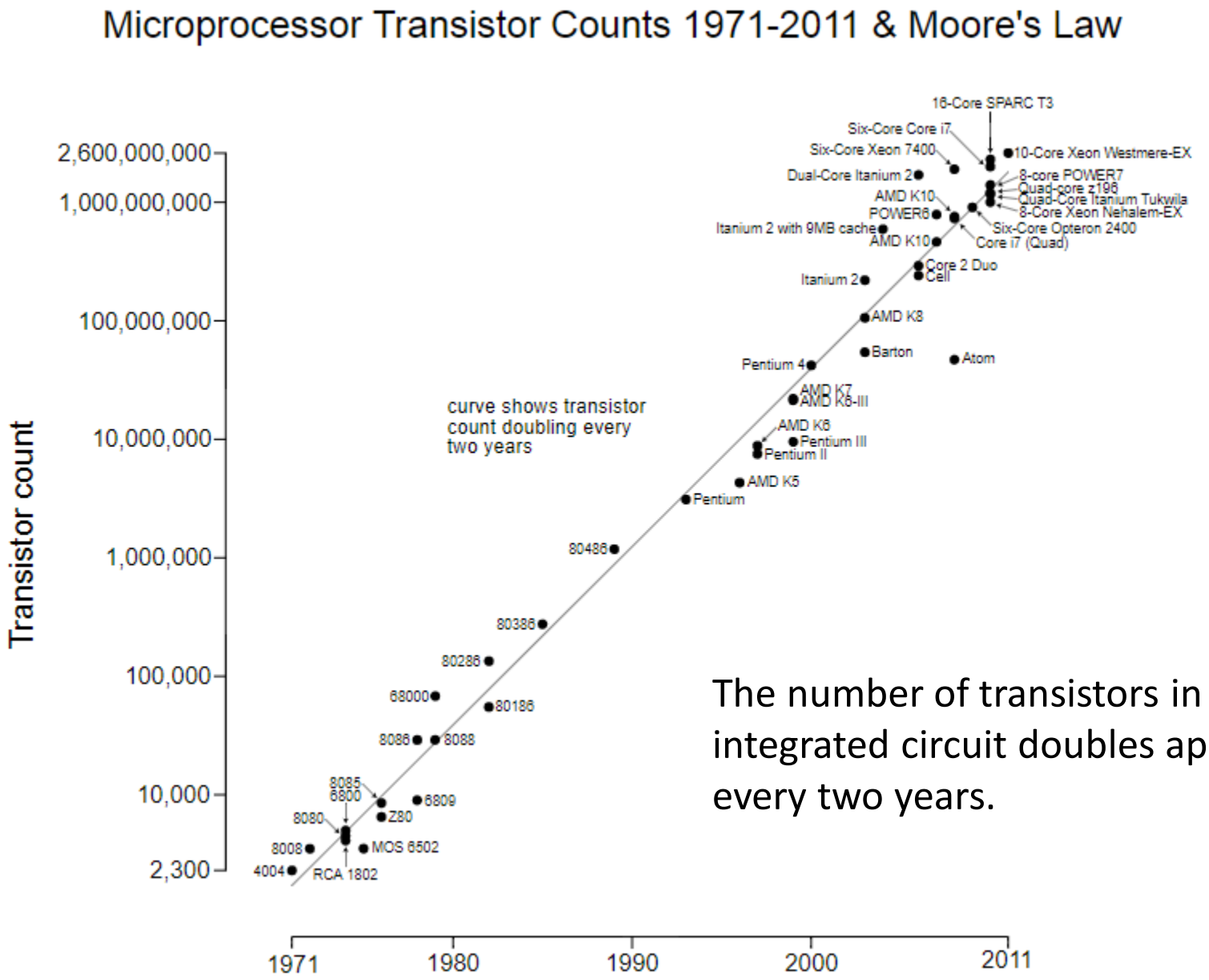
Last Updated on **June 17, 2019**

NMOS FET



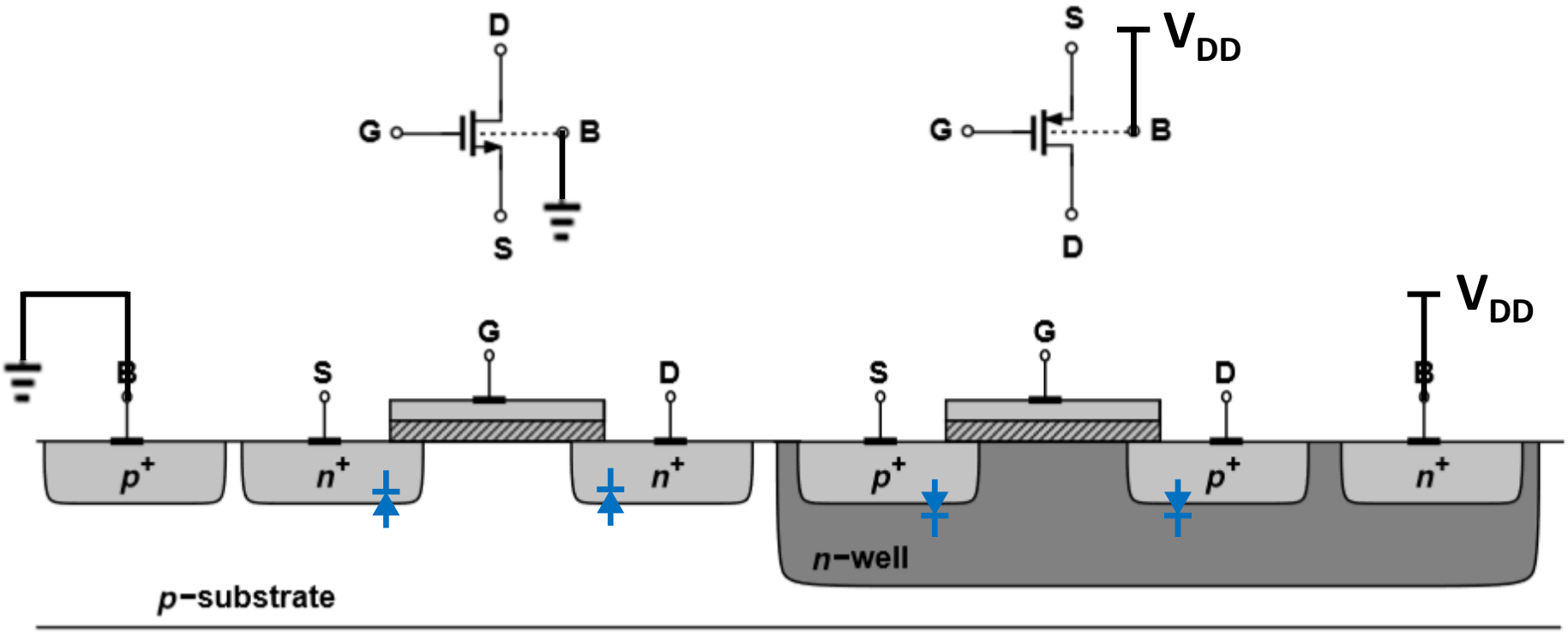
- MOS = Metal-Oxide-Semiconductor
- FET = Field effect Transistor
- CMOS Technology keeps on reducing t_{ox} and L_{eff} (Moore's Law), for: (1) lower cost, (2) faster speed, and (3) less power consumption.
- Substrate (Body) of NMOS is generally connected to ground.
- See Chapter 17 for the introduction of CMOS fabrication technology.

Moore's Law



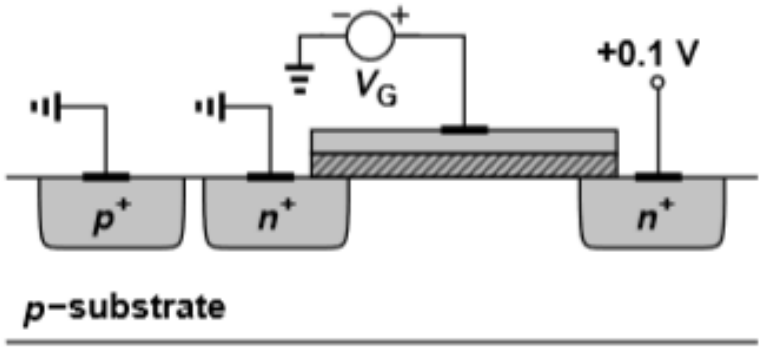
The number of transistors in a dense integrated circuit doubles approximately every two years.

CMOS

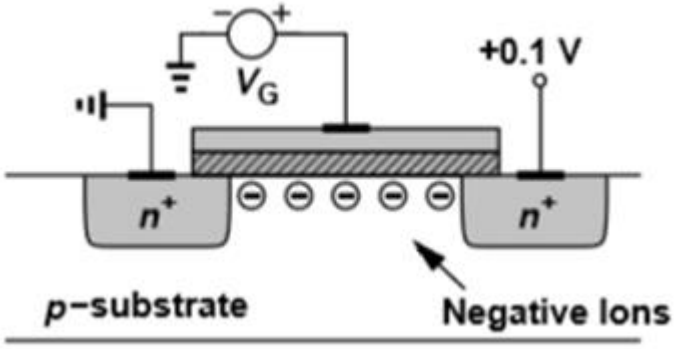


- CMOS = Complementary MOS
- Substrate (Body) of NMOS is generally connected to ground.
- N-well (Body) of PMOS is generally connected to V_{DD}. Sometimes, it can also be connected to source for eliminating body effect.

Threshold Voltage (V_{TH}) for NMOS

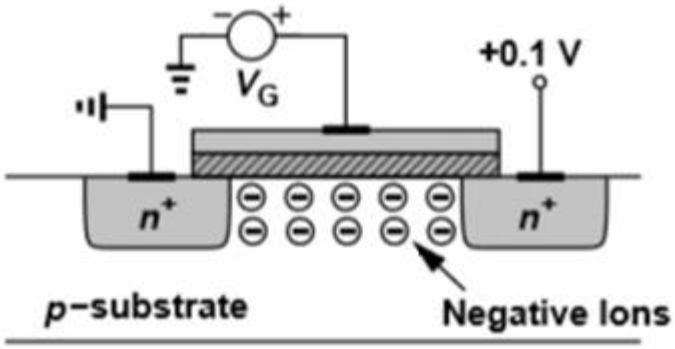


- V_G = 0 V
- No current flow

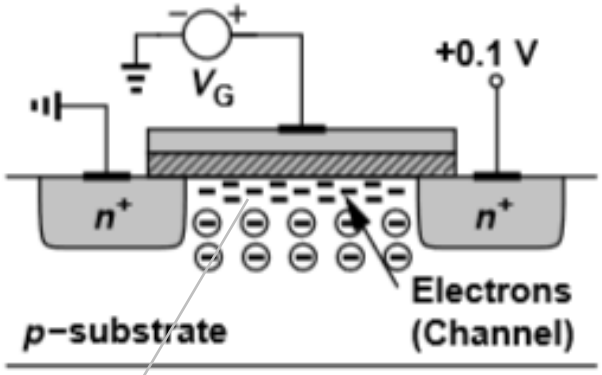


- As V_G increases from zero, holes in *p*-substrate are repelled, leaving negative ions (ionized boron dopants) behind to form a **depletion region**.
- Positive charges are mirrored at the gate.
- No charge carriers (electrons or holes) in the channel, so no current flow.

Threshold Voltage (V_{TH}) for NMOS



- Increasing V_G further increases the width of the **depletion region** and the potential at the oxide-silicon interface.



Interface as n-type as the substrate is p-type

- When V_G reaches sufficiently positive value, a channel of electrons (**inversion layer**) is formed beneath the gate oxide.
- Electrons flow from the source to the interface and eventually to the drain.
- The value of V_G at which the inversion layer occurs is the threshold voltage (V_{TH}).
- If V_G rises further, the charge in the depletion region remains relatively constant, while the charge in the inversion layer increases rapidly.

Threshold Voltage (V_{TH}) for NMOS

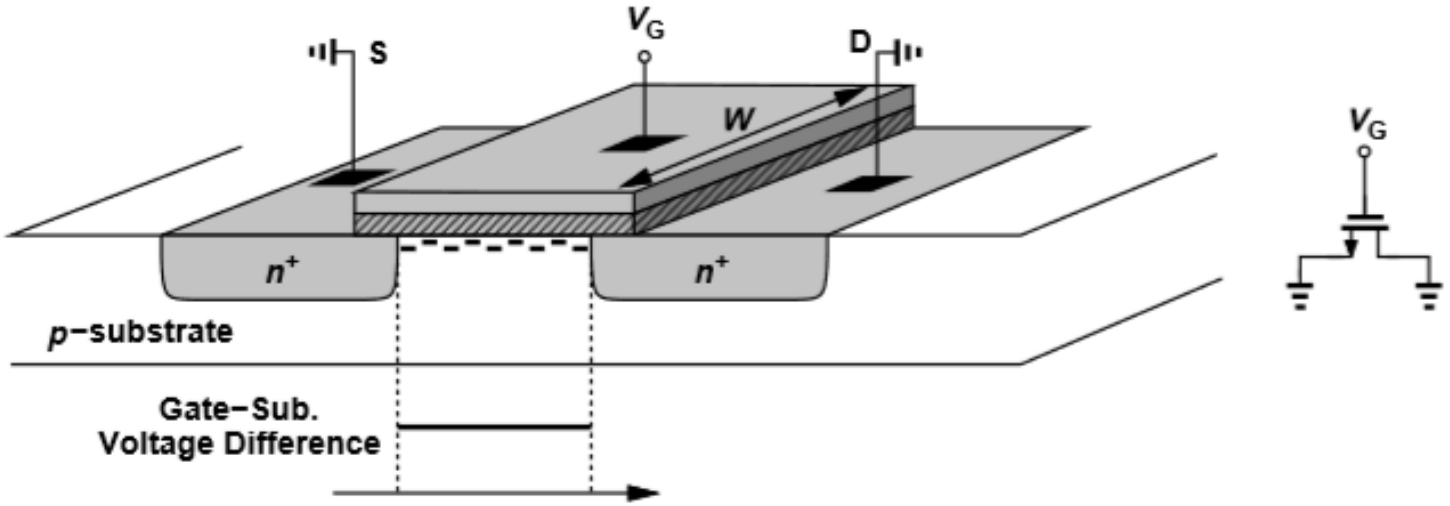
$$V_{TH} = \Phi_{MS} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}}$$

$$\Phi_F = \frac{kT}{q} \ln \frac{N_{sub}}{n_i}$$

$$Q_{dep} = \sqrt{4q\epsilon_{si}\Phi_F N_{sub}}$$

- Φ_{MS}: The difference between the work functions of the polysilicon gate and the silicon substrate (Φ_M = 4.15 V for n+ poly-Si; Φ_S = 4.15 + 0.56 + Φ_F V for the p- silicon substrate)
- **k**: Boltzmann's constant (1.38×10^{-23} J·K^{-1})
- **q**: Electron charge (1.60×10^{-19} coulomb)
- N_{sub}: The doping density of the silicon substrate (cm^{-3})
- n_i: The density of electrons or holes in undoped silicon
- Q_{dep}: The charge in the depletion region per unit area (coulomb·cm^{-2})
- C_{ox}: The gate oxide capacitance per unit area = $\epsilon_{silicon\ oxide}$ / t_{ox} (F·μm^{-2})
- ϵ_{si}: The electric permittivity of silicon ($11.7 \times 8.85 \times 10^{-12}$ F·m^{-1})

I-V Characteristics for NMOS (Triode)

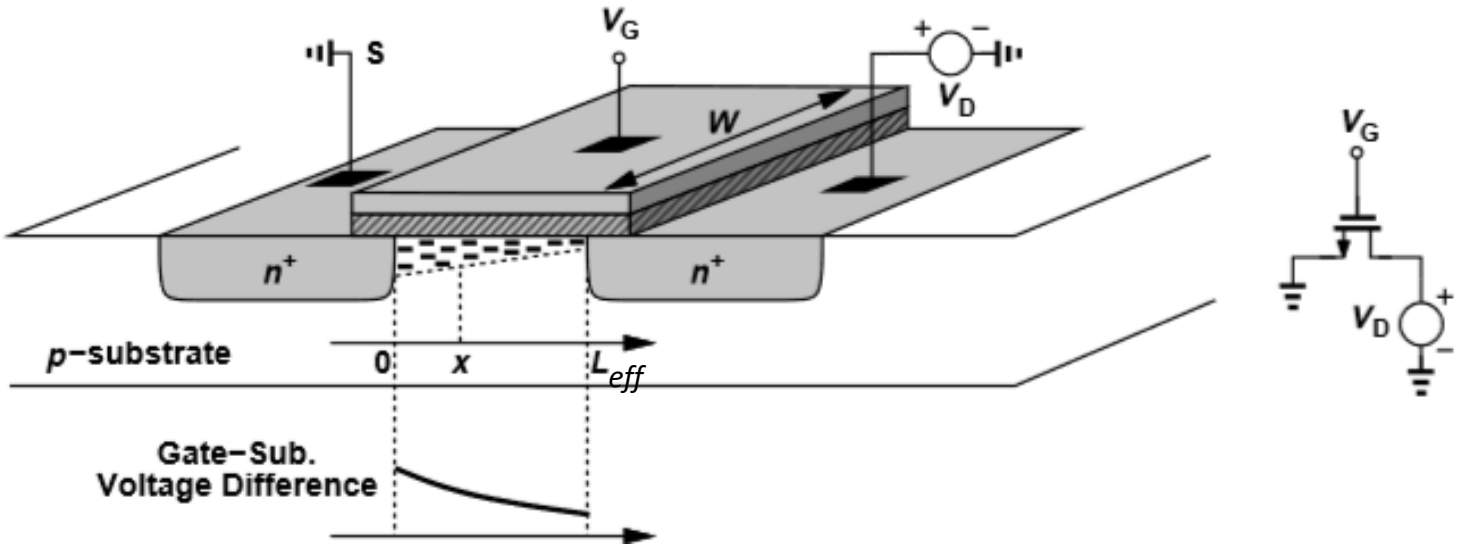


For $V_{GS} \geq V_{TH}$

$Q = -WL_{eff}C_{ox}(V_{GS} - V_{TH})$ (unit: coulomb)

$Q_d = -WC_{ox}(V_{GS} - V_{TH})$ (unit: coulomb·m^{-1})

I-V Characteristics for NMOS (Triode)



$$I_D = Q_d \times v = Q_d \times (\mu_n \mathcal{E}) = -WC_{ox}[V_{GS} - V_{TH} - V(x)] \times (\mu_n \mathcal{E}) \qquad \mathcal{E} = -dV(x)/dx$$

$$= WC_{ox}[V_{GS} - V_{TH} - V(x)] \times \mu_n \times \frac{dV(x)}{dx}$$

$$\int_{x=0}^{x=L_{eff}} I_D \cdot dx = \int_{V(0)=0}^{V(L)=V_{DS}} \mu_n C_{ox} W[V_{GS} - V_{TH} - V(x)] \cdot dV(x)$$

I_D: constant along channel

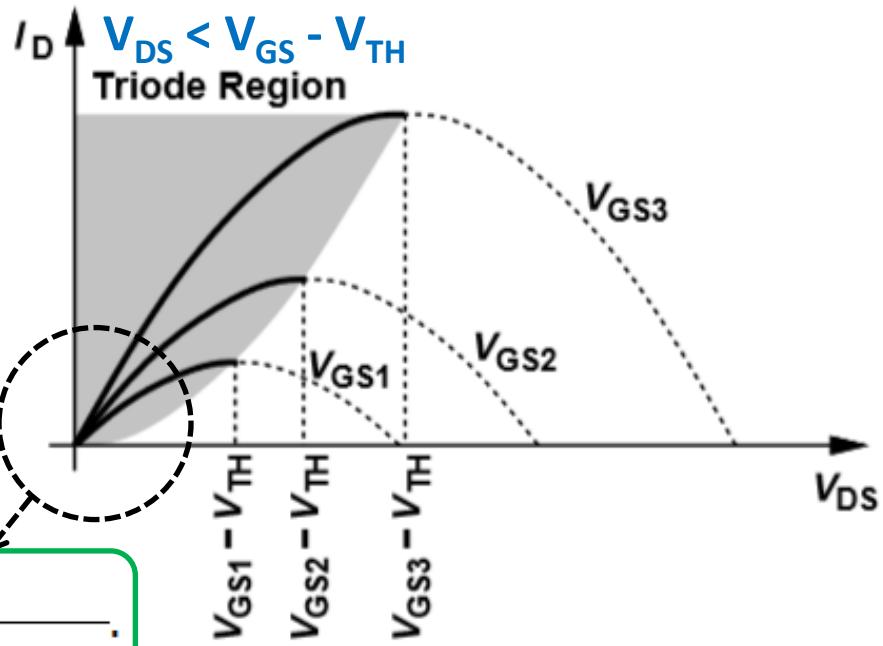
$$\mathbf{I_D = \mu_n C_{ox} \frac{W}{L_{eff}} \left[(V_{GS} - V_{TH})V_{DS} - \frac{1}{2}V_{DS}^2\right]}$$

I-V Characteristics for NMOS (Triode)

$$I_D = \mu_n C_{ox} \frac{W}{L_{eff}} [(V_{GS} - V_{TH})V_{DS} - \frac{1}{2} V_{DS}^2]$$

$$I_{D,max} = \frac{1}{2} \mu_n C_{ox} \frac{W}{L_{eff}} (V_{GS} - V_{TH})^2$$

$V_{DS} = V_{GS} - V_{TH}$



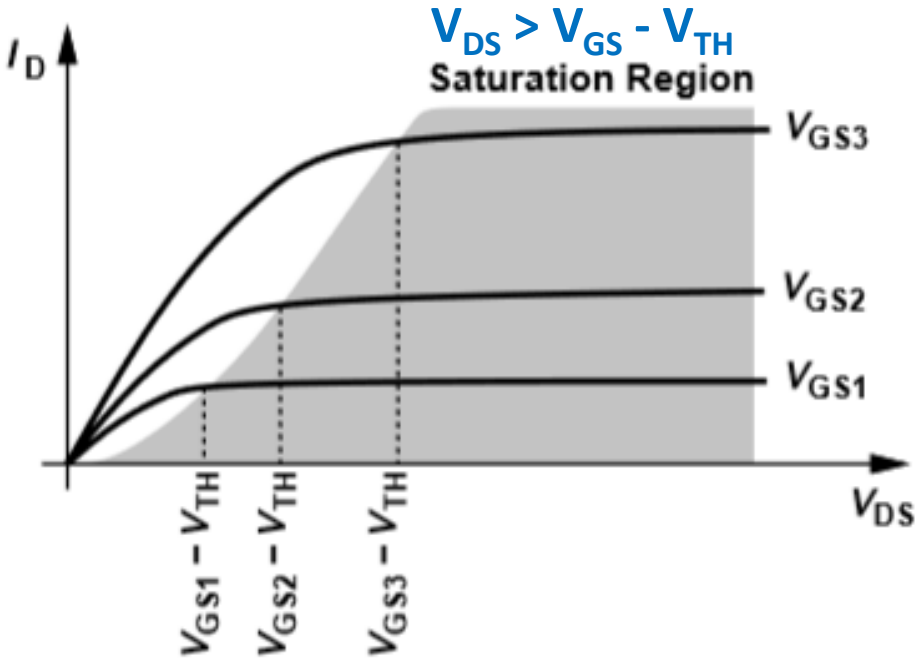
$V_{DS} << 2(V_{GS} - V_{TH})$

Deep triode region

$$R_{on} = \frac{1}{\mu_n C_{ox} \left(\frac{W}{L_{eff}}\right) (V_{GS} - V_{TH})}.$$

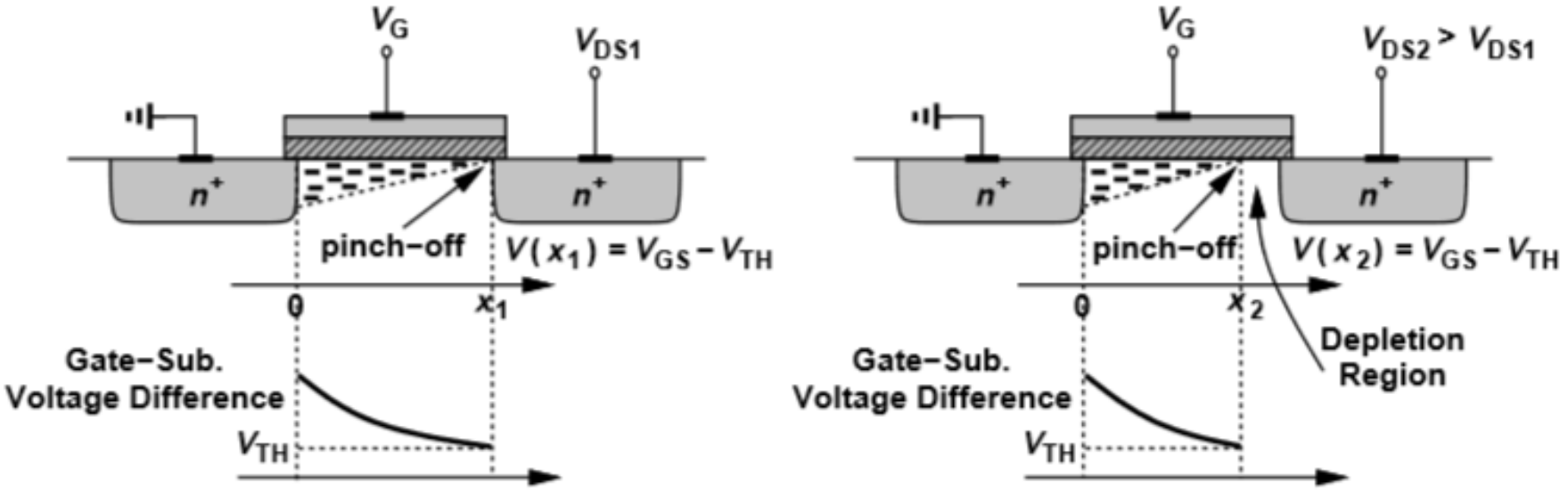
- For digital circuit, MOSFET, as a switch, usually operates in deep triode region.
- This is why reducing t_{ox} and L_{eff} can improve speed.

I-V Characteristics for NMOS (Saturation)



- In reality, for $V_{DS} > V_{GS} – V_{TH}$, I_D becomes relatively constant.
- $V_{DS} = V_{GS} – V_{TH}$ is the minimum value for the NMOS to operate in saturation region.

I-V Characteristics for NMOS (Saturation)



$$\int_{x=0}^{x=L'} I_D \cdot dx = \int_{V(0)=0}^{V(L')=V_{GS}-V_{TH}} \mu_n C_{ox} W[V_{GS} - V_{TH} - V(x)] \cdot dV(x)$$

$$\mathbf{I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2}$$

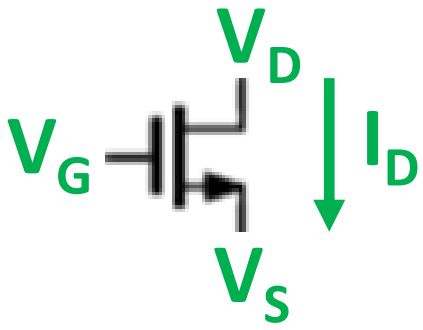
I_D: constant along channel

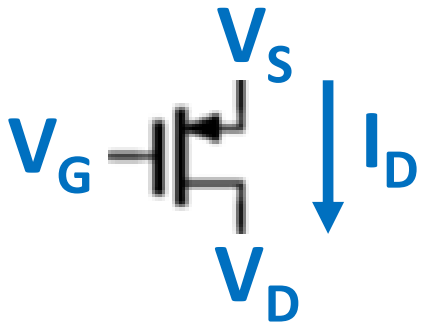
L': the point at which Q_d drops to zero

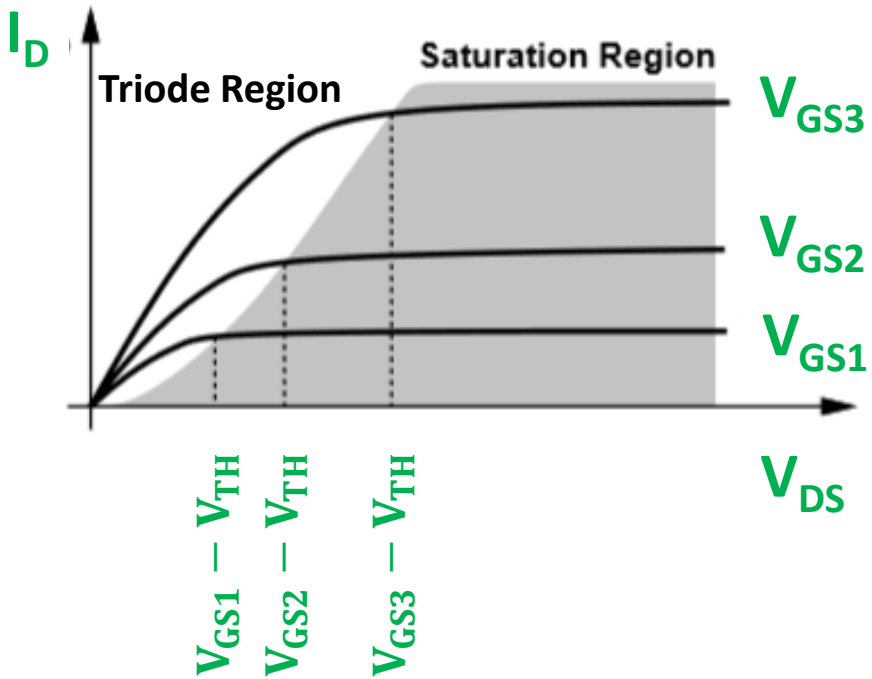
$V_{GS} - V_{TH}$: the overdrive voltage

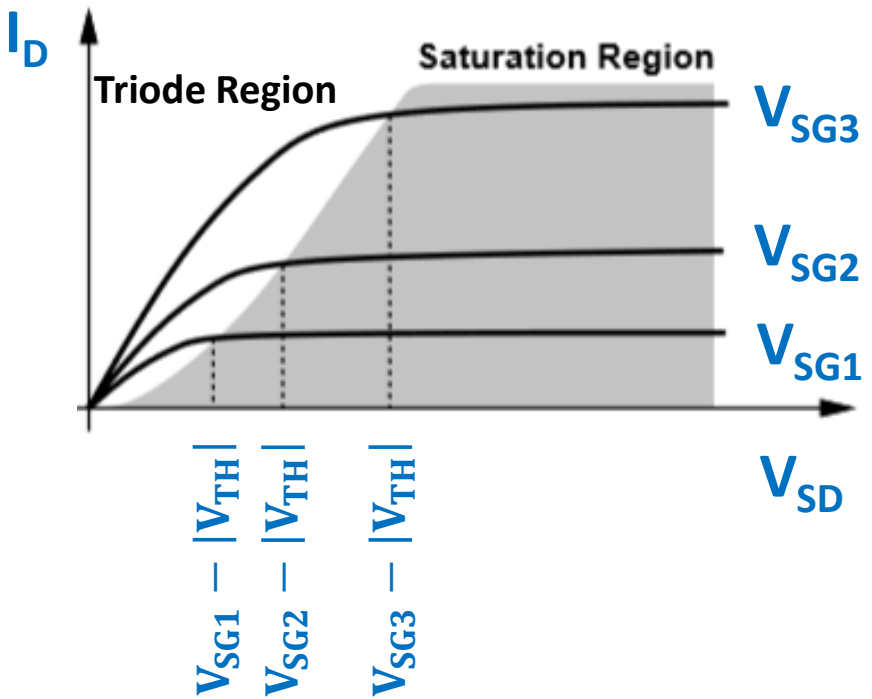
- Electron velocity ($v = I_D / Q_d$) becomes tremendously high at the pinch off point ($Q_d \rightarrow 0$), such that electrons shoot through the depletion region and arrive at the drain terminal.

NMOS vs PMOS


V_D
V_G
I_D
V_S

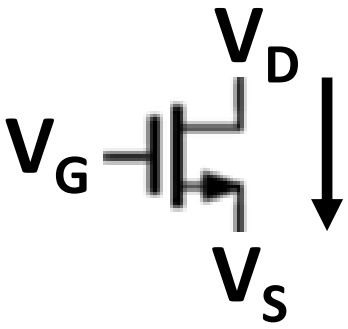

V_S
V_G
I_D
V_D


I_D
Triode Region
Saturation Region
V_{GS3}
V_{GS2}
V_{GS1}
V_{DS}
$V_{GS1} - V_{TH}$
$V_{GS2} - V_{TH}$
$V_{GS3} - V_{TH}$


I_D
Triode Region
Saturation Region
V_{SG3}
V_{SG2}
V_{SG1}
V_{SD}
$V_{SG1} - |V_{TH}|$
$V_{SG2} - |V_{TH}|$
$V_{SG3} - |V_{TH}|$

Transconductance

- For the NMOS operating in the saturation region ($V_{DS} \geq V_{GS} - V_{TH}$):



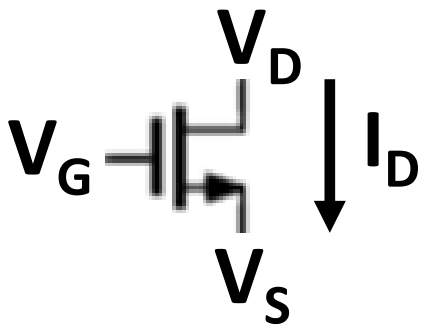
$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2$$

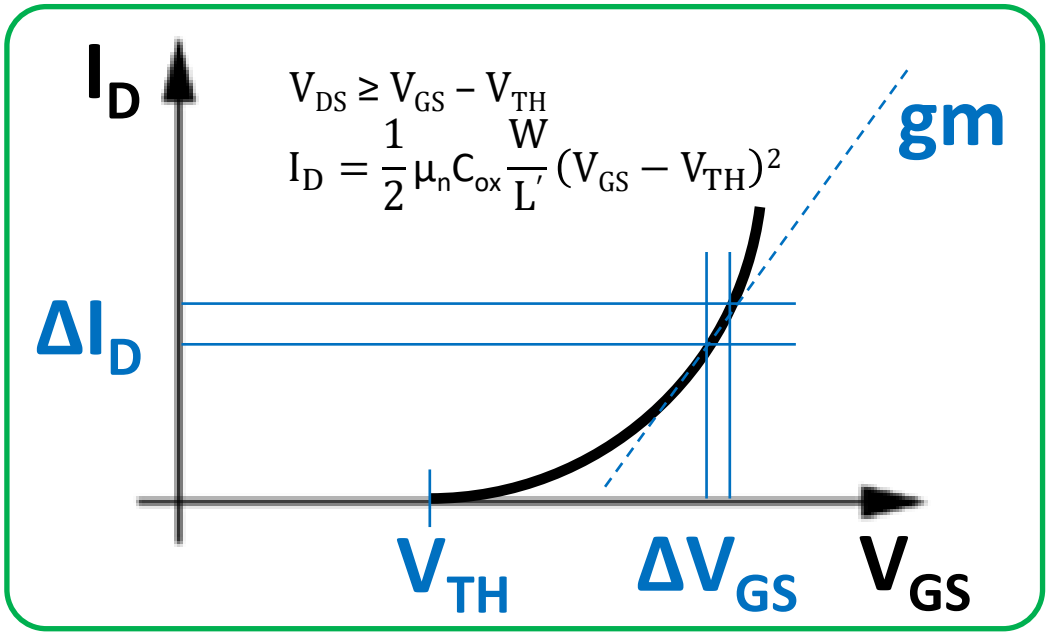
- ΔV_{GS} results in $\Delta I_D = gm \times \Delta V_{GS}$.

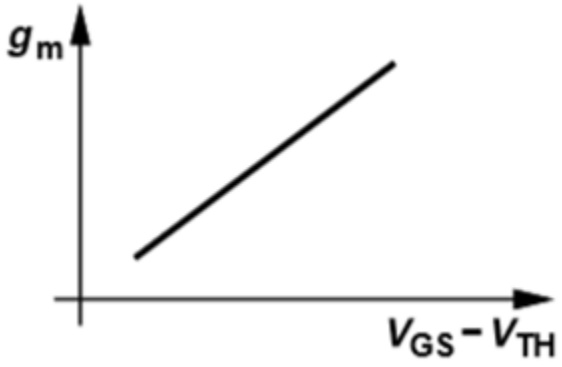
$$gm = \frac{\partial I_D}{\partial V_{GS}} = \mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})$$

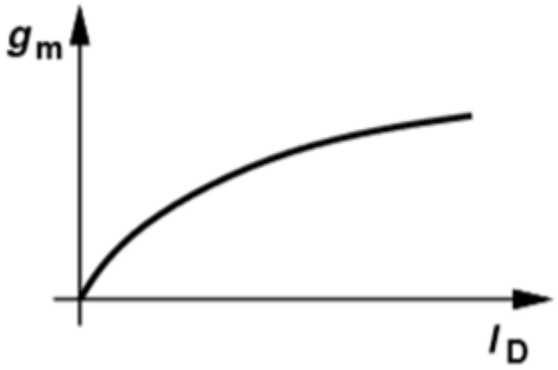
$$= \sqrt{2\mu_n C_{ox}\frac{W}{L'} I_D} = \frac{2I_D}{V_{GS} - V_{TH}}$$

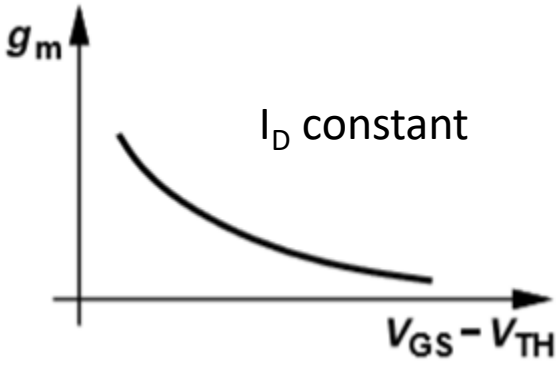
Transconductance





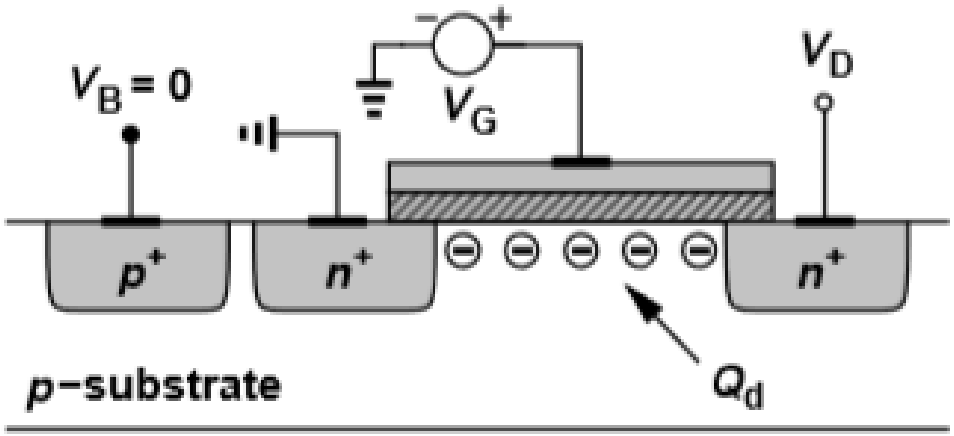




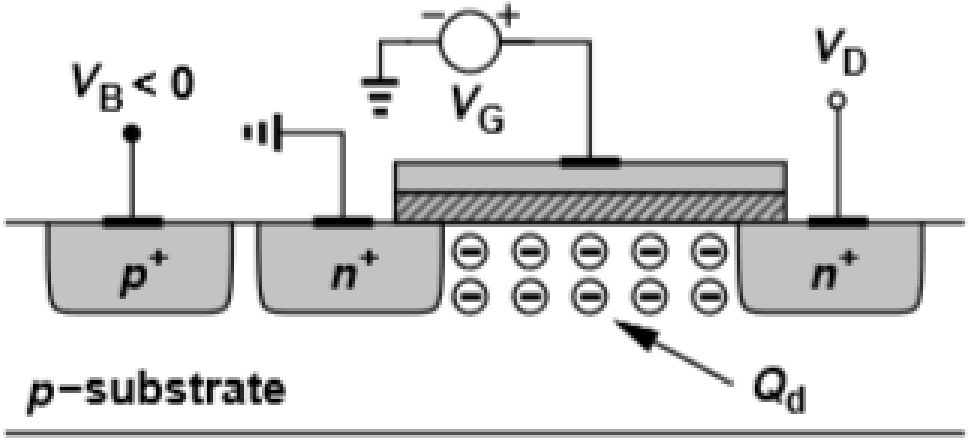


- For a given NMOS, gm changes according to the DC biasing condition.
- If a small signal is applied to a NMOS with defined biasing values, we assume the signal amplitude is small enough that the variation in gm is negligible.

Body Effect



- $V_D = V_S = 0$ V
- $V_G < V_{TH}$
- $V_B = 0$ V
- The depletion region is formed.



- $V_D = V_S = 0$ V
- $V_G < V_{TH}$
- $V_B < 0$ V
- More holes are attracted to the substrate connection, leaving a wider depletion region behind.
- Q_{dep} increases, thus V_{TH} increases.

$$V_{TH} = \Phi_{MS} + 2\Phi_F + \frac{Q_{dep}}{C_{ox}}$$

Body Effect

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

$$\Phi_F = \frac{kT}{q}\ln\frac{N_{sub}}{n_i} \qquad \gamma = \frac{\sqrt{2q\epsilon_{Si}N_{sub}}}{C_{ox}}$$

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2$$

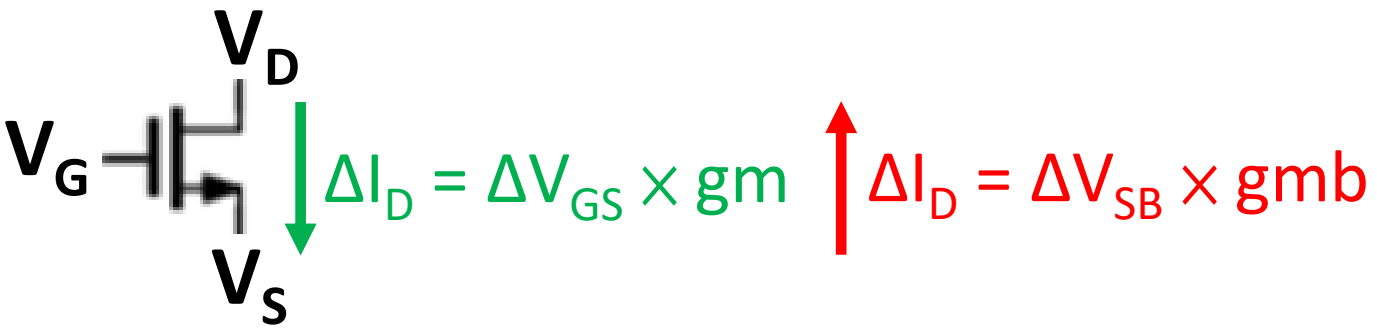
Body Effect

$$\mathbf{gmb} = \frac{\partial I_D}{\partial V_{SB}} = \frac{\partial I_D}{\partial V_{TH}} \cdot \frac{\partial V_{TH}}{\partial V_{SB}}$$

$$= -\mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) \cdot \frac{\partial V_{TH}}{\partial V_{SB}}$$

$$= -\mu_n C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) \cdot \frac{\gamma}{2} \frac{1}{\sqrt{|2\Phi_F + V_{SB}|}}$$

$$= -\mathbf{gm} \cdot \boldsymbol{\eta}$$



- V_{GS} increases, I_D increases.
- V_{SB} increases, V_{TH} increases and thus I_D decreases.

Channel-Length Modulation

$$I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L'}(V_{GS} - V_{TH})^2$$

$$L' = L - \Delta L$$

$$\frac{1}{L'} = \frac{1}{L - \Delta L} = \frac{1}{L}\cdot\frac{1}{1 - \frac{\Delta L}{L}} \approx \frac{1}{L}\cdot\left(1 + \frac{\Delta L}{L}\right)$$

Note: $L = L_{eff}$ here

$$\boldsymbol{I_D = \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2\left(1 + \frac{\Delta L}{L}\right)}$$

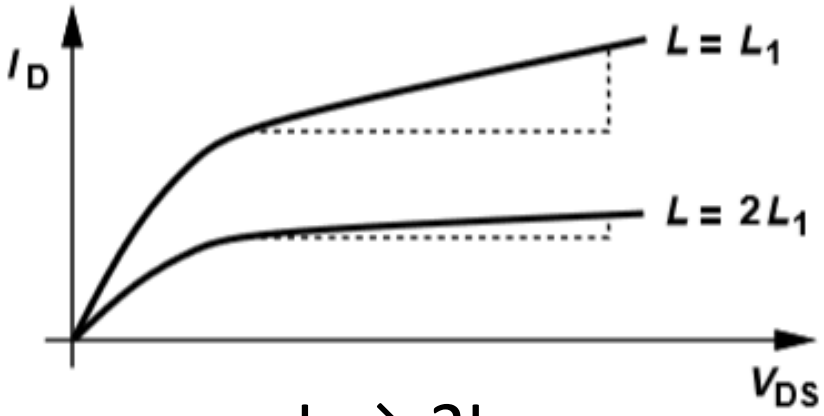
$$\boldsymbol{= \frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2(1 + \lambda V_{DS})}$$

Channel-Length Modulation

$$r_o = \frac{\partial V_{DS}}{\partial I_D} = 1 \Big/ \frac{\partial I_D}{\partial V_{DS}}$$

$$= \frac{1}{\frac{1}{2}\mu_n C_{ox}\frac{W}{L}(V_{GS} - V_{TH})^2 \cdot \lambda}$$

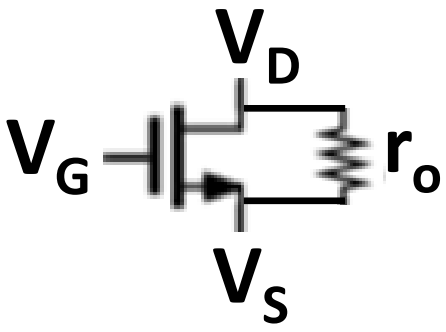
$$\approx \frac{1}{I_D \cdot \lambda}$$



$L \rightarrow 2L$

$r_o \rightarrow 4r_o$

- Longer L leads to larger r_o, closer to the ideal current source.



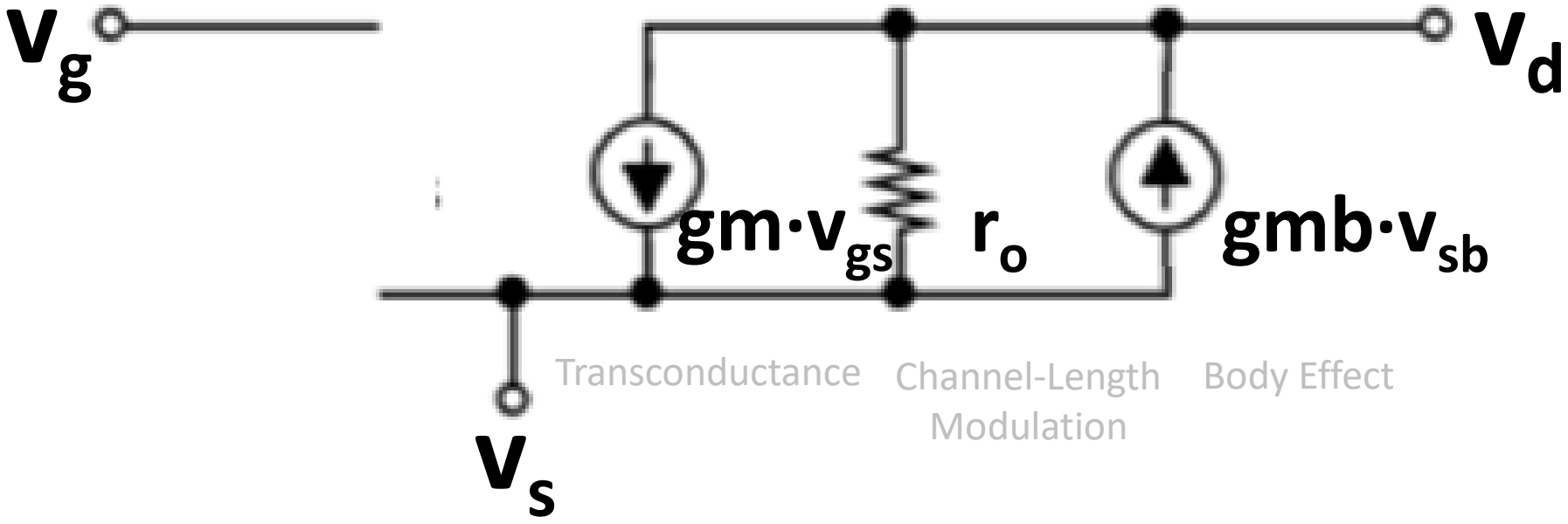
Small-Signal Model for NMOS

$V_d = V_D + v_d$

$V_g = V_G + v_g$

$I_D \downarrow$ $i_d = gm \cdot v_{gs}$ $\uparrow$ $i_d = gmb \cdot v_{sb}$

$V_s = V_S + v_s$



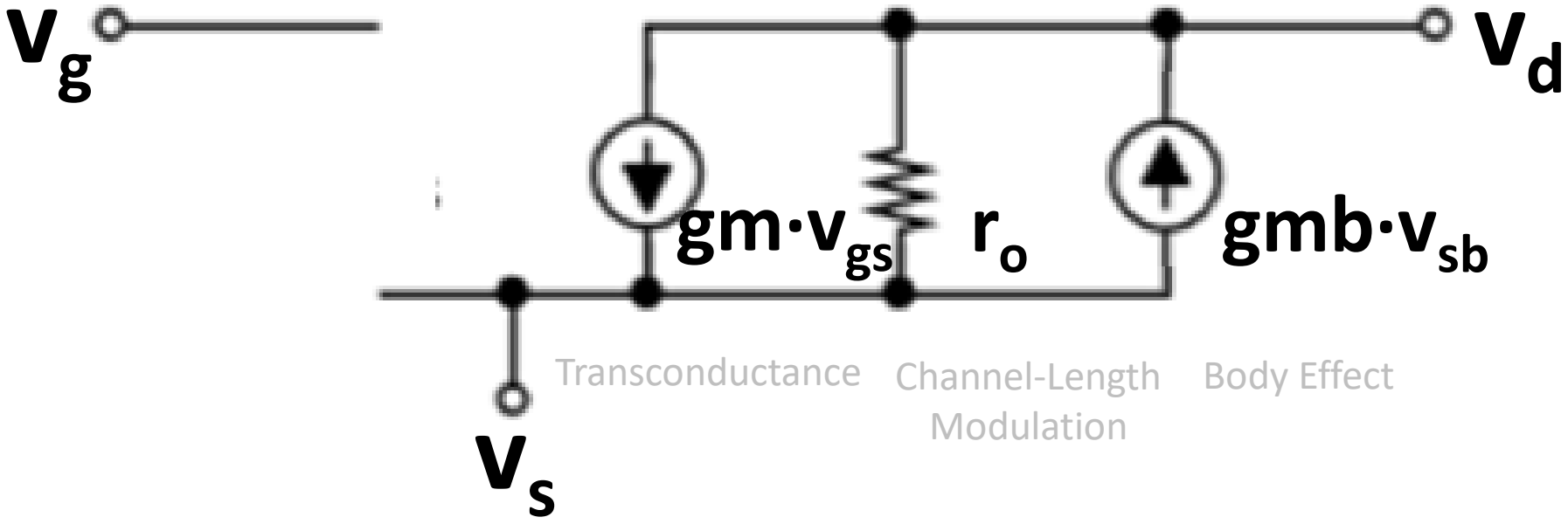
Small-Signal Model for PMOS

$V_s = V_S + v_s$

$V_g = V_G + v_g$

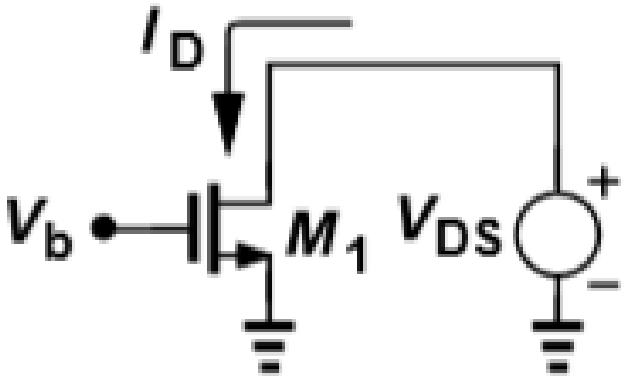
$I_D \uparrow i_d = gm \cdot v_{gs} \downarrow i_d = gmb \cdot v_{sb}$

$V_d = V_D + v_d$

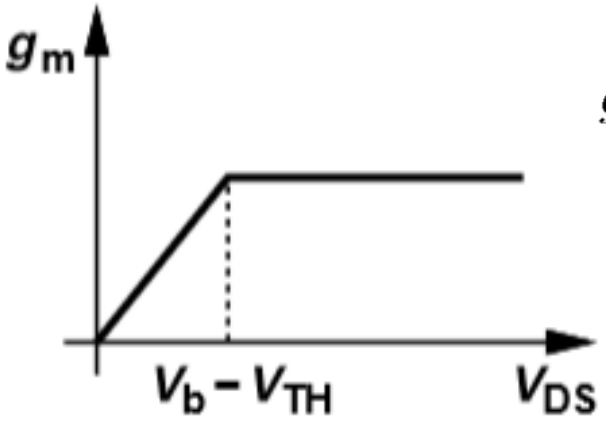


Example

Plot gm of M_1 as a function of V_{DS}. Assume $\lambda = \gamma = 0$.



Solution:



For the triode region:

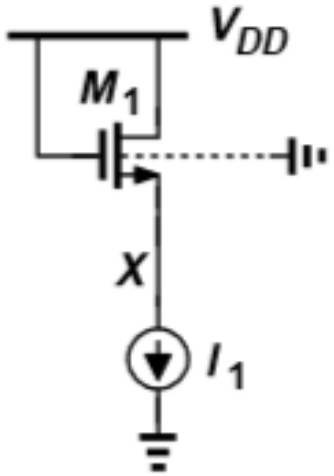
$$g_m = \frac{\partial}{\partial V_{GS}} \left\{ \frac{1}{2} \mu_n C_{ox} \frac{W}{L} \left[2(V_{GS} - V_{TH}) V_{DS} - V_{DS}^2 \right] \right\}$$

$$= \mu_n C_{ox} \frac{W}{L} V_{DS}.$$

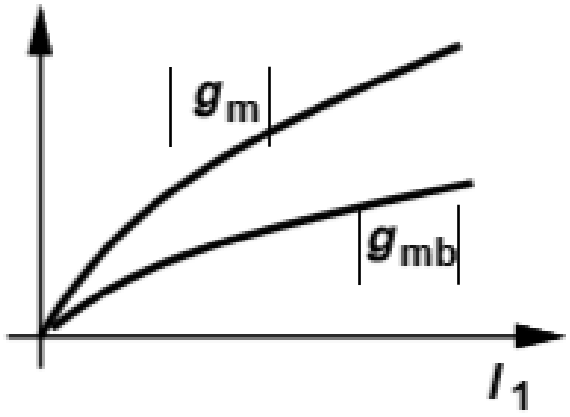
- This is why we want MOSFET to operate in the saturation region for the highest amplification.

Example

Sketch gm and gmb of M_1 as a function of the bias current I_1.



Solution:

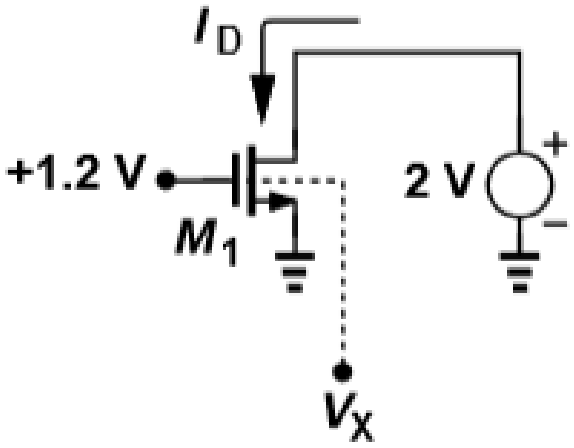


$$gm = \frac{\partial I_D}{\partial V_{GS}} = \mu_n C_{ox} \frac{W}{L'} (V_{GS} - V_{TH}) = \sqrt{2\mu_n C_{ox} \frac{W}{L'} I_D}$$

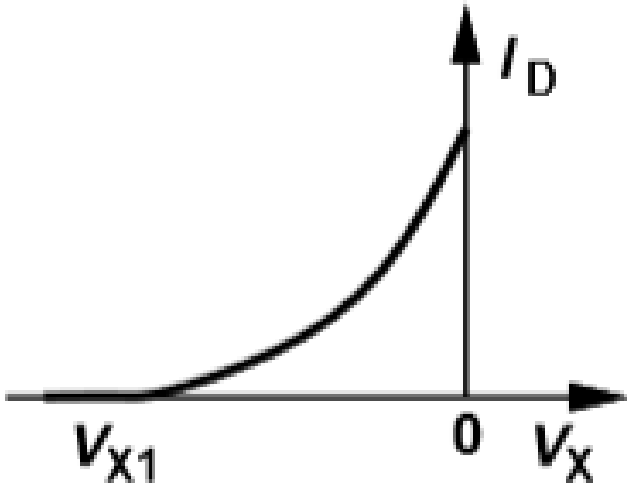
$$gmb = -gm \cdot \frac{\gamma}{2} \frac{1}{\sqrt{|2\Phi_F + V_{SB}|}}$$

Example

Sketch I_D as a function of V_X increasing from $-\infty$ to 0. Assume V_{TH} = 0.6 V, γ = 0.4 $V^{1/2}$ and $2\Phi_F$ = 0.7 V.



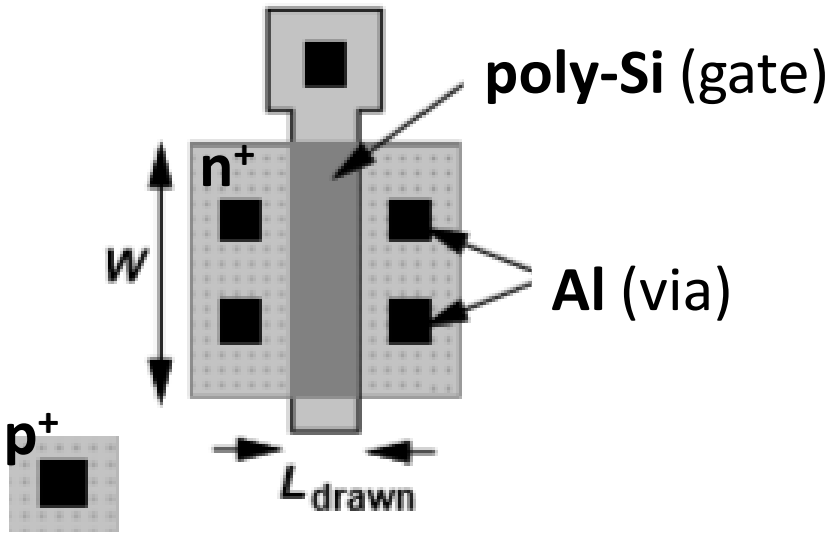
Solution:



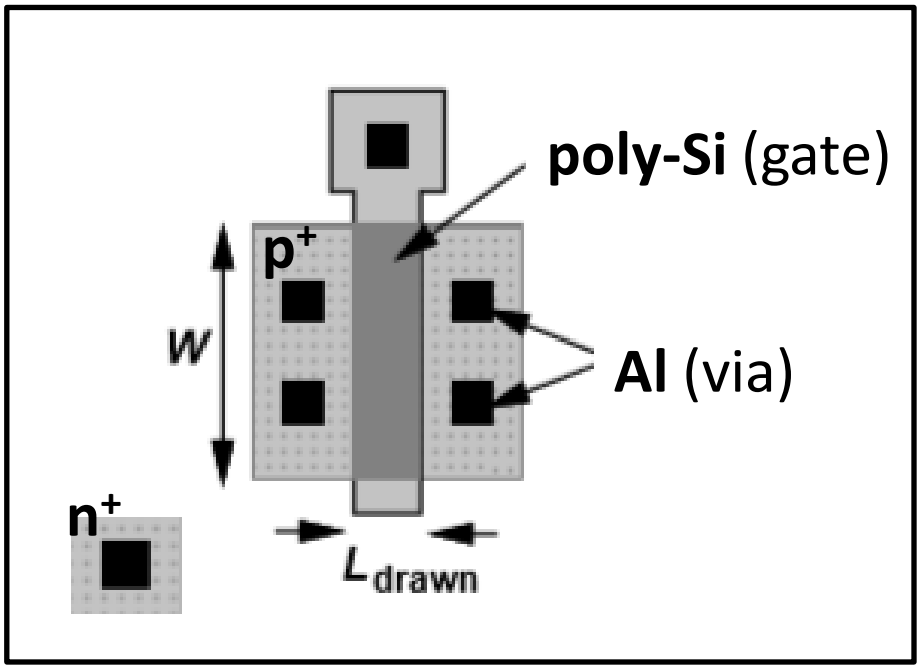
$$1.2 = 0.6 + 0.4(\sqrt{0.7 - V_X} - \sqrt{0.7}),\ V_X = -4.76\ \text{V}.$$

Layout

NMOS



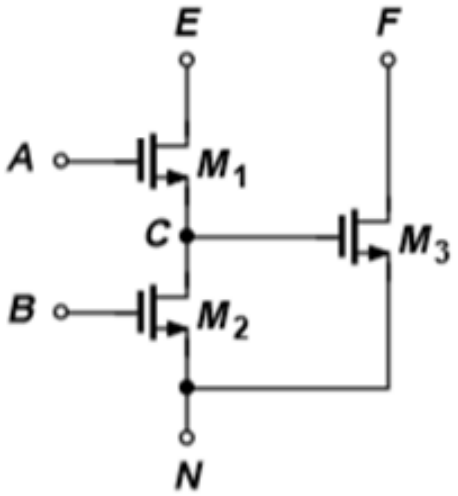
PMOS



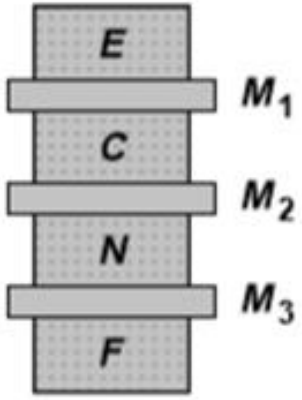
- W/L is chosen to determine gm. Minimum L is dictated by the process.
- Design rules: (1) Poly-Si extends beyond the channel area by some amount. (2) Enough n^+, p^+ or poly-Si area surrounding each via. (3) Enough distance between two vias. (4) Many others.

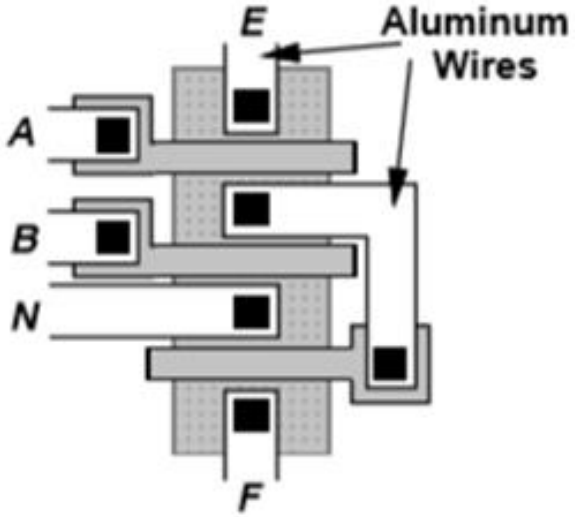
Example

What is the layout for the schematic below which saves the most area and introduces the least parasitic capacitance?



Solution:





- Share S/D

SPICE Model

NMOS Model

LEVEL = 1	VTO = 0.7	GAMMA = 0.45	PHI = 0.9
NSUB = 9e+14	LD = 0.08e−6	UO = 350	LAMBDA = 0.1
TOX = 9e−9	PB = 0.9	CJ = 0.56e−3	CJSW = 0.35e−11
MJ = 0.45	MJSW = 0.2	CGDO = 0.4e−9	JS = 1.0e−8

PMOS Model

LEVEL = 1	VTO = −0.8	GAMMA = 0.4	PHI = 0.8
NSUB = 5e+14	LD = 0.09e−6	UO = 100	LAMBDA = 0.2
TOX = 9e−9	PB = 0.9	CJ = 0.94e−3	CJSW = 0.32e−11
MJ = 0.5	MJSW = 0.3	CGDO = 0.3e−9	JS = 0.5e−8

- Simulators such as SPICE and Cadence need accurate models for each device.
- Above is the simplest MOS SPICE model, known as “Level 1,” and provide typical values for each parameter corresponding to 0.5-µm technology.

NMOS vs PMOS in Performance

- PMOS devices are quite inferior to NMOS in most CMOS technology.
- Lower mobility of holes ($\mu_p C_{ox} \approx 0.5\mu_n C_{ox}$) yield lower current drive and conductance.
- NMOS exhibit higher output resistance, providing more ideal current sources and higher voltage gain.
- It is preferable to use NMOS rather than PMOS wherever possible.