Design and Implementation of VLSI Systems
Lecture04

MOSFET

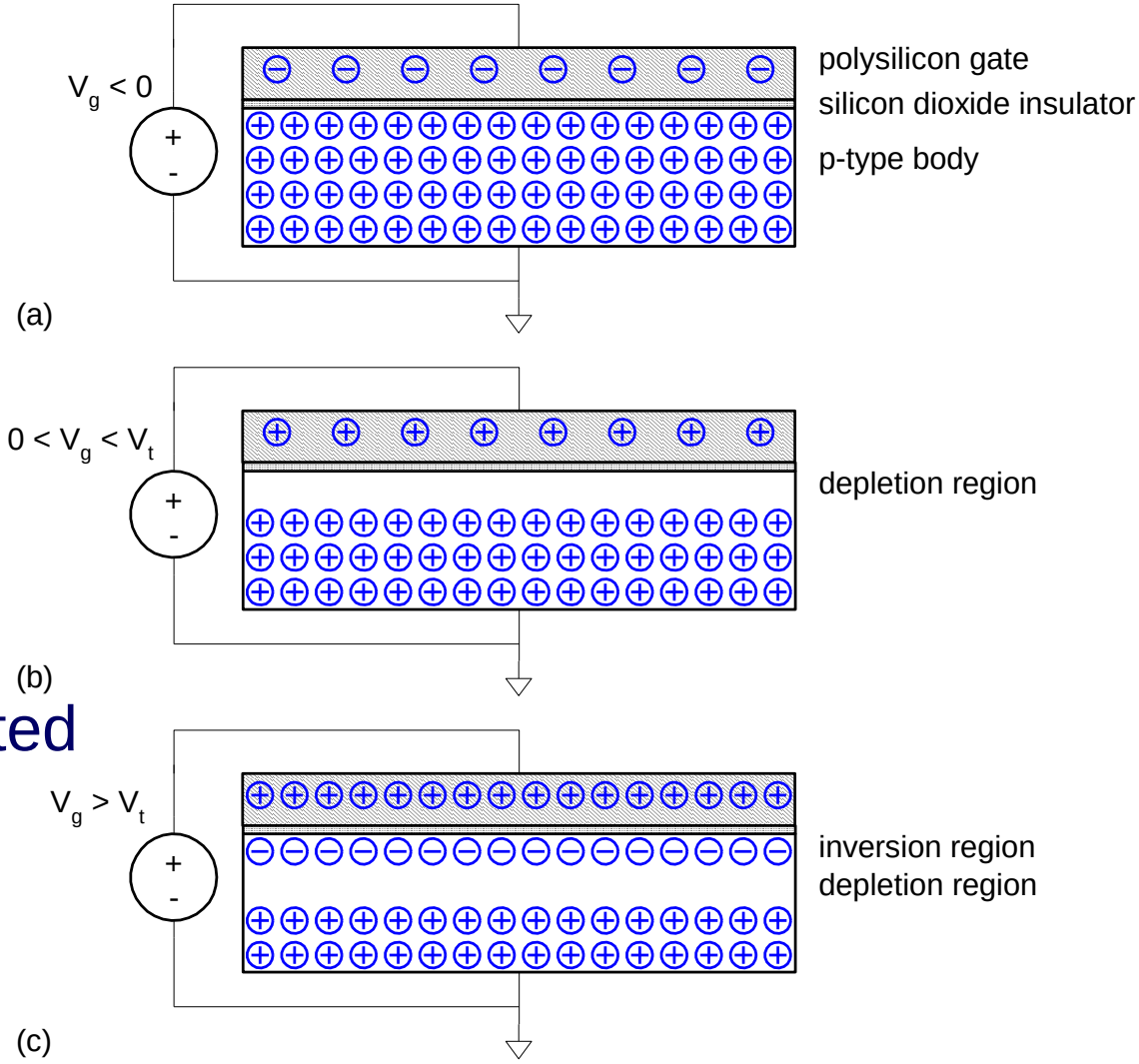
MOS transistor theory

- Schedule for 4 lectures
  - Ideal (Shockley) Model
  - Non-ideal model
  - Inverter DC characteristics
  - SPICE

gate-oxide-body sandwich = capacitor

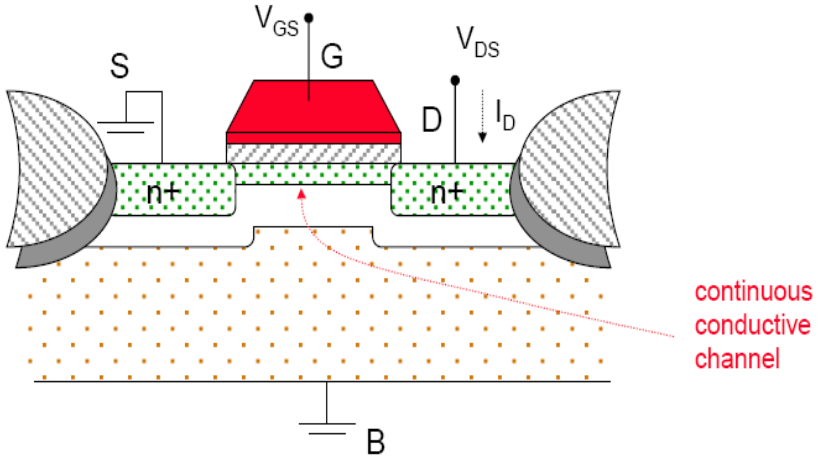
Operating modes

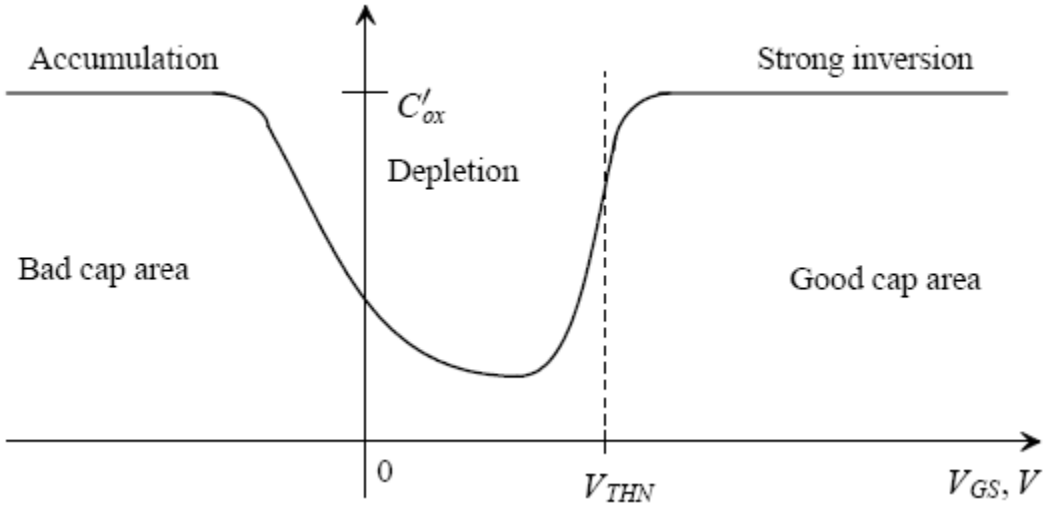
- Accumulation
- Depletion
- Inversion



- The charge accumulated is proportional to the *excess* gate-channel voltage (V_{gc}-V_t)

Gate capacitance as a function of Vgs





The MOS transistor has three regions of operation

- Cut off

  $V_{gs} < V_t$

- Linear (resistor):

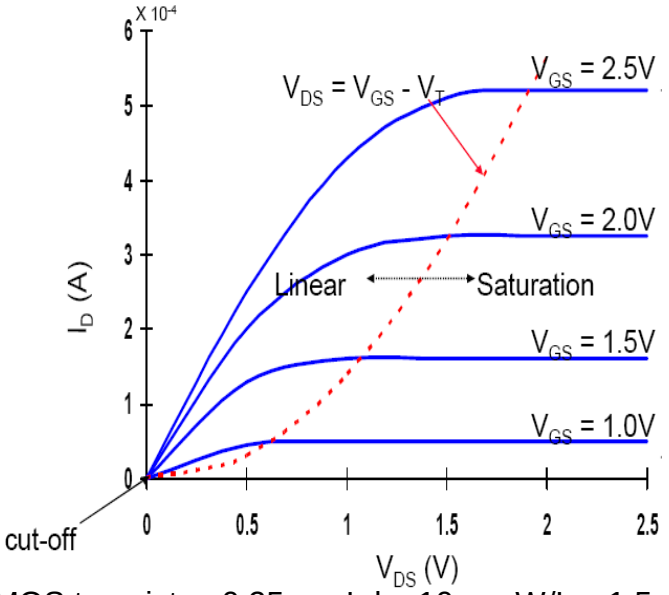
  $V_{gs} > V_t$ & $V_{ds} < V_{SAT}=V_{gs}-V_t$

  Current prop to V_{ds}

- Saturation:

  $V_{gs} > V_t$ and $V_{ds} \geq V_{SAT}=V_{gs}-V_t$

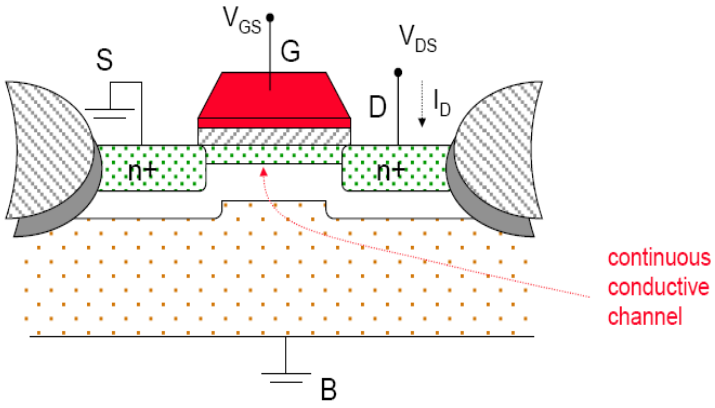
  Current is independent of V_{ds}

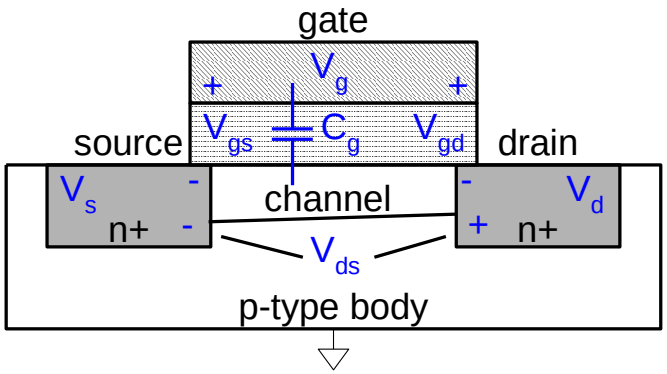


NMOS transistor, 0.25um, Ld = 10um, W/L = 1.5, VDD = 2.5V, VT = 0.4V

How to calculate the current value?

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel
- $Q_{channel} = CV$
- $C = \varepsilon_{ox}WL/t_{ox} = C_{ox}WL$ (where $C_{ox}=\varepsilon_{ox}/t_{ox}$)
- $V = V_{gc} – V_t = (V_{gs} – V_{ds}/2) – V_t$





Carrier velocity is a factor in determining the current

- Charge is carried by electrons
- Carrier velocity *v* proportional to lateral E-field between source and drain
- $v = \mu E$ μ called mobility
- $E = V_{ds}/L$
- Time for carrier to cross channel:

  $t = L / v$

I=Q/t

- Now we know
  - How much charge $Q_{channel}$ is in the channel
  - How much time *t* each carrier takes to cross

$$I_{ds} = \frac{Q_{\text{channel}}}{t}$$

$$= \mu C_{\text{ox}} \frac{W}{L}\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$

$$= \beta\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$

In linear mode ($V_{gs} > V_t$ & $V_{ds} < V_{gs}-V_t$)

$$I_{ds} = \frac{Q_{\text{channel}}}{t}$$

$$= \mu C_{\text{ox}} \frac{W}{L}\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$

$$= \beta\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$

Can be ignored for small V_{ds}

$$I_{ds} = \beta((V_{gs} - V_t)V_{ds} - V_{ds}^2/2)$$

$$I_{ds} = \beta(V_{gs} - V_t)V_{ds}$$

- *For a given V_{gs}, I_{ds} is proportional (linear) to V_{ds}*

In saturation mode ($V_{gs} > V_t$ and $V_{ds} \geq V_{gs}-V_t$)

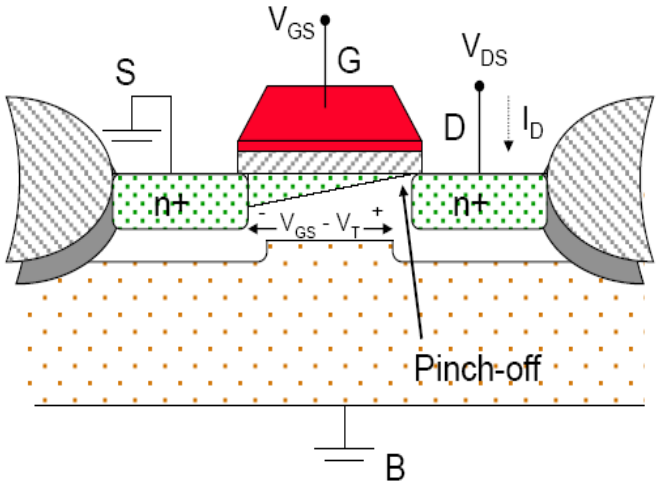
$$I_{ds} = \frac{Q_{\text{channel}}}{t}$$

$$= \mu C_{\text{ox}} \frac{W}{L}\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$

$$= \beta\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$

$$I_{ds} = \beta\left(V_{gs} - V_t - V_{dsat}/2\right)V_{dsat}$$

$$= \frac{\beta}{2}\left(V_{gs} - V_t\right)^2$$



pinched off

➢ *Now drain voltage no longer increases current*

Operation modes summary

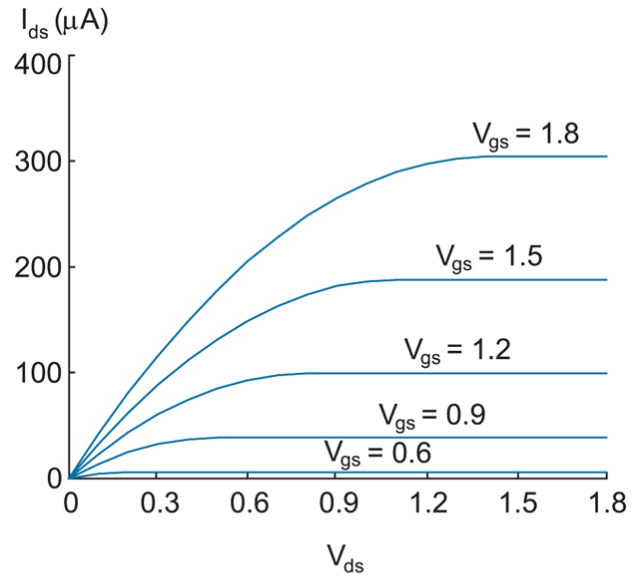


FIG 2.7 I-V characteristics of ideal nMOS transistor

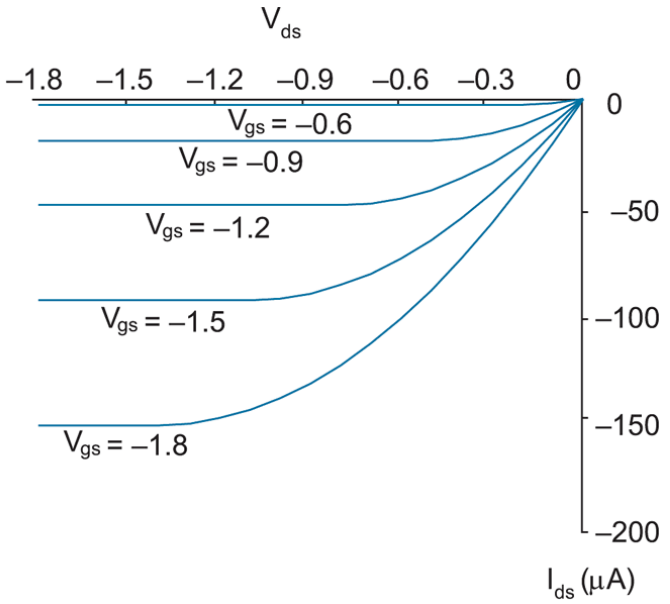


FIG 2.8 I-V characteristics of ideal pMOS transistor

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t \quad \text{cutoff} \\ \beta\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds} & V_{ds} < V_{dsat} \quad \text{linear} \\ \frac{\beta}{2}\left(V_{gs} - V_t\right)^2 & V_{ds} > V_{dsat} \quad \text{saturation} \end{cases}$$

CURRENT-VOLTAGE RELATIONS

$$\beta = \mu C_{ox} \frac{W}{L}$$ ß: transconductance parameter of transistor

W/L: width-to-length ratio

- As W increases, more carriers available to conduct current.
- As L increases, V_{ds} diminishes in effect (more voltage drop) → takes longer to push carriers across the transistor → reducing current flow.

Transistor capacitance

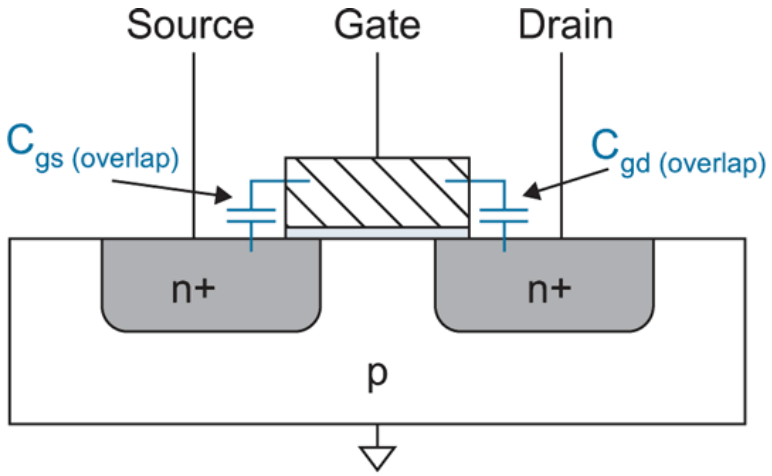


FIG 2.10 Overlap capacitance

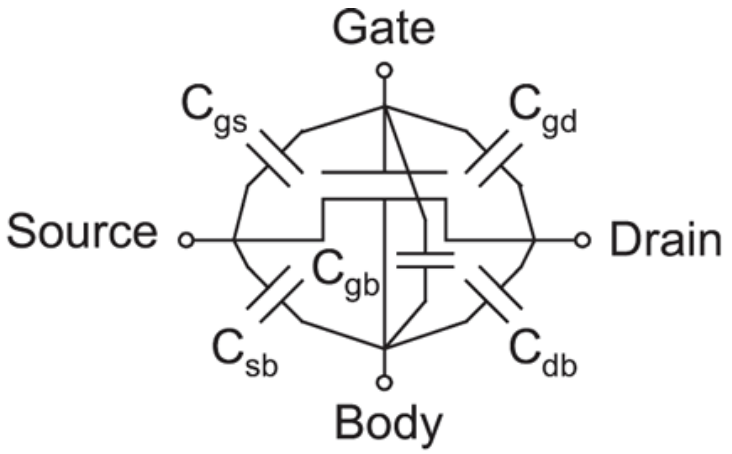
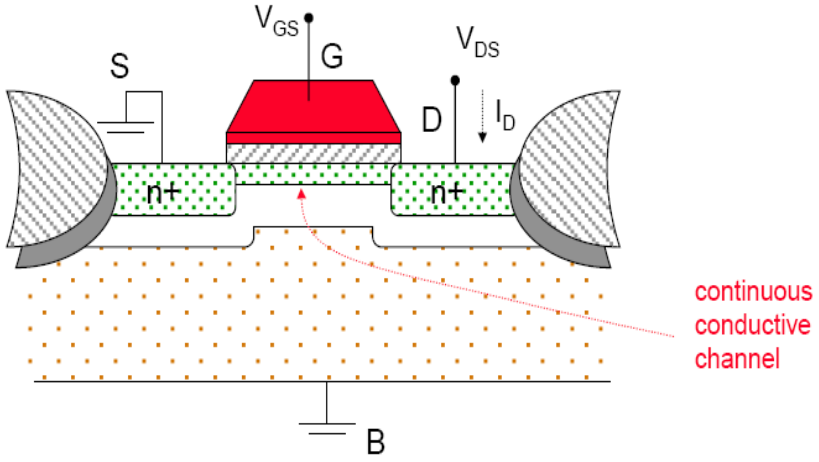
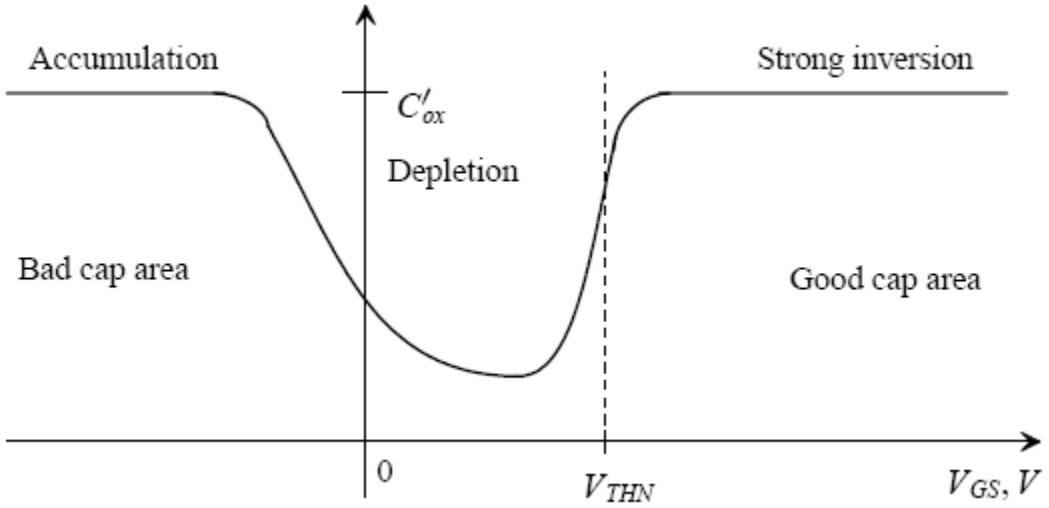


FIG 2.14 Capacitances of an MOS transistor

- Gate capacitance: to body + to drain + to source
- Diffusion capacitance: source-body and drain-body capacitances

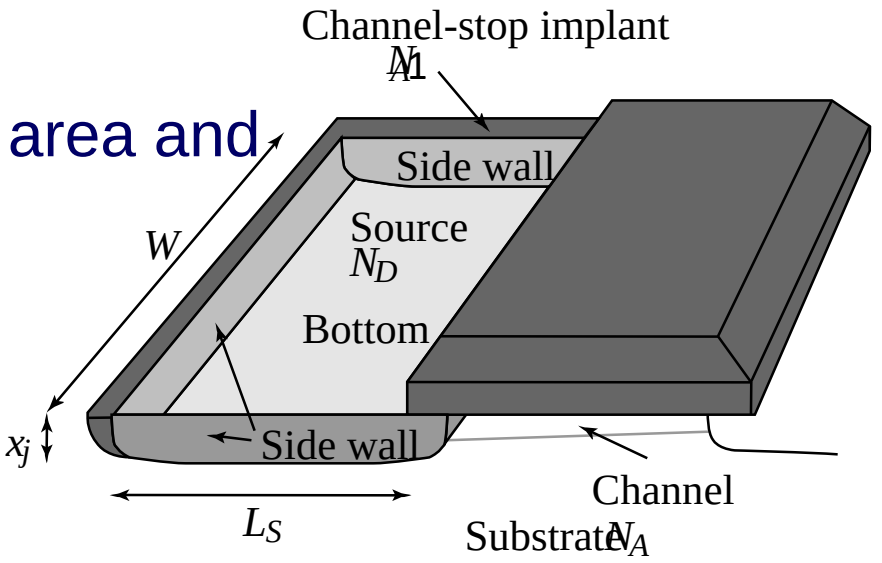
Gate capacitance as a function of Vgs





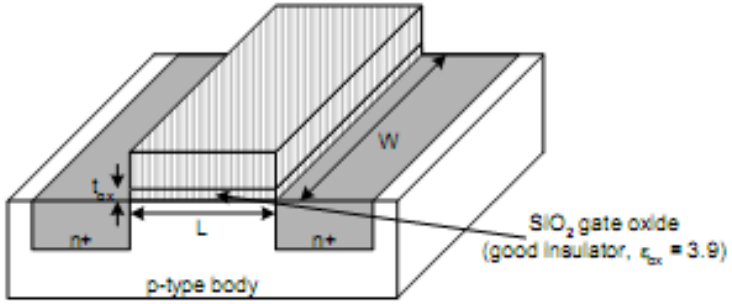
Source/Drain diffusion capacitance

- C_{sb}, C_{db}
- Undesirable, called *parasitic* capacitance
- Capacitance depends on area and perimeter
  - Use small diffusion nodes
  - Comparable to C_g
  - Varies with process



$$C_{diff} = C_{bottom} + C_{sw} = C_j \times AREA + C_{jsw} \times PERIMETER$$
$$= C_j L_S W + C_{jsw}(2L_S + W)$$

SUMMARY OF SHOCKLEY MODEL



$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t \quad \text{cutoff} \\ \beta\left(V_{gs} - V_t - \frac{V_{ds}}{2}\right)V_{ds} & V_{ds} < V_{dsat} \quad \text{linear} \\ \frac{\beta}{2}\left(V_{gs} - V_t\right)^2 & V_{ds} > V_{dsat} \quad \text{saturation} \end{cases}$$

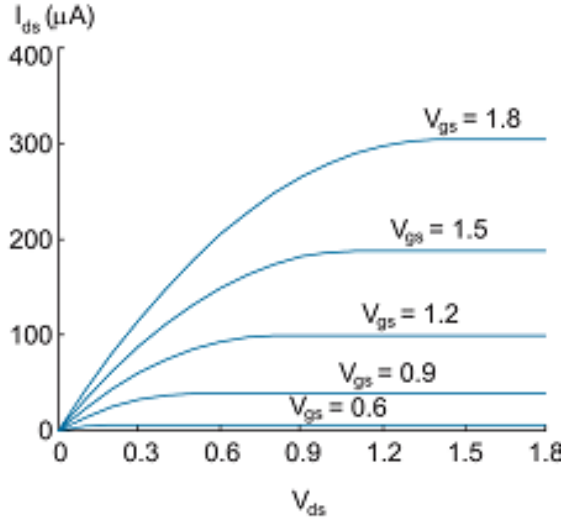


FIG 2.7 I-V characteristics of ideal nMOS transistor

for nMOS

$$\beta = \beta_n = \mu_n \frac{\varepsilon_{ox} W}{t_{ox} L}$$
$$V_t = V_{tn}$$

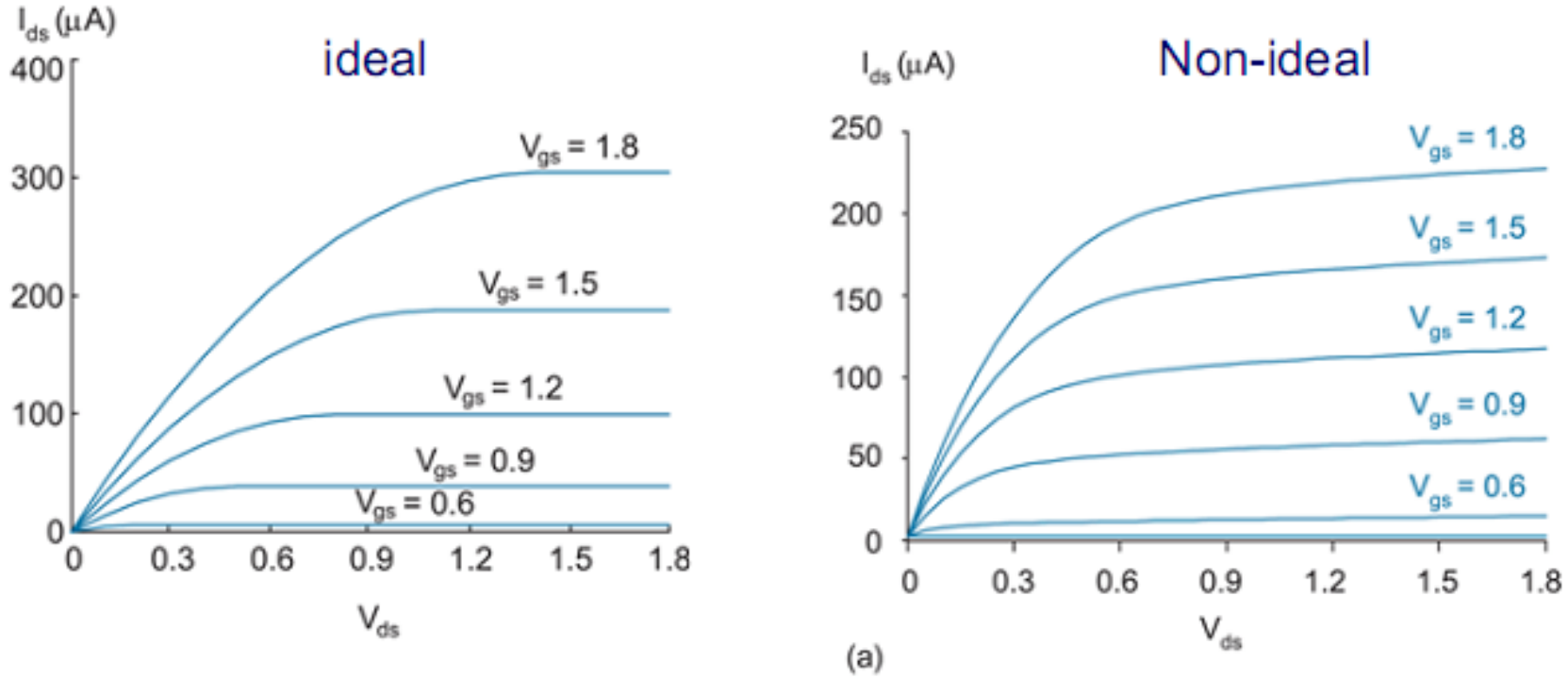
for pMOS

$$\beta = \beta_p = \mu_p \frac{\varepsilon_{ox} W}{t_{ox} L}$$
$$V_t = V_{tp}$$

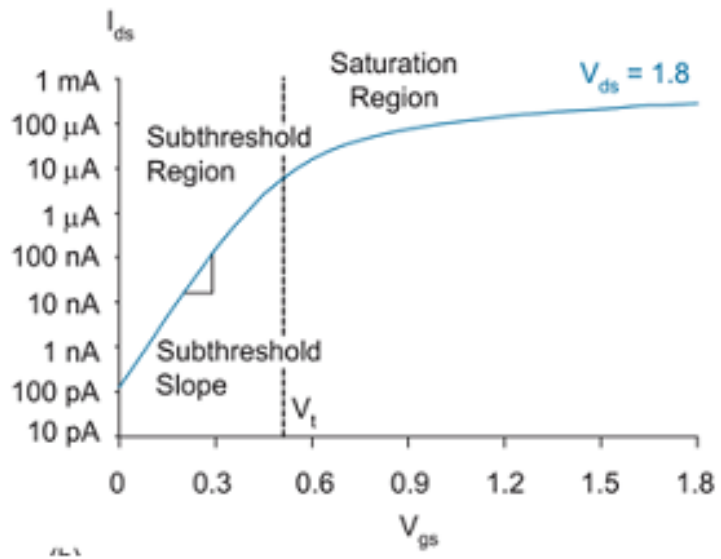
$$\mu_p < \mu_n (\mu_n \approx 2 \times \mu_p)$$

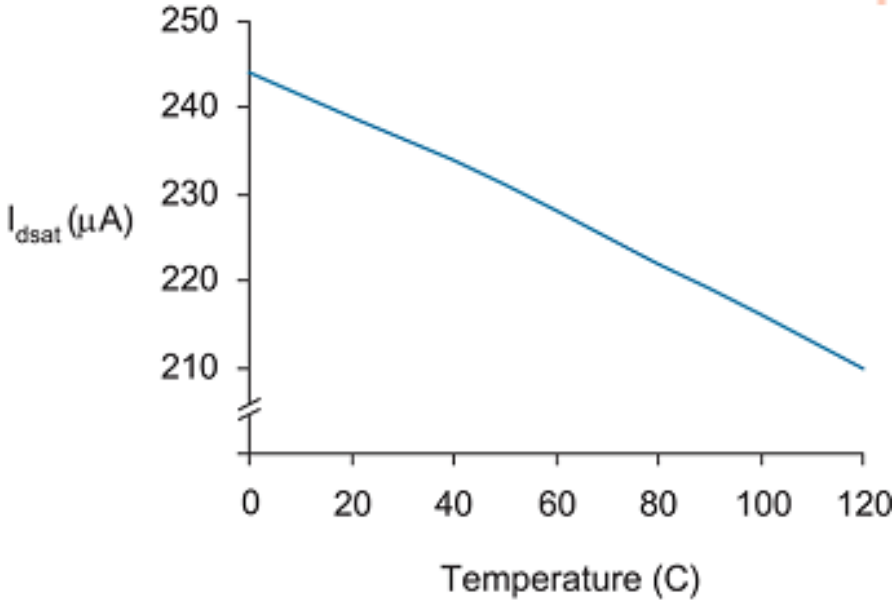
Covered ideal (long channel) operation (Shockley model) of transistor

IDEAL VS. NON-IDEAL



- Saturation current does not increase quadratically with Vgs
- Saturation current lightly increases with increase in Vds





- There is leakage current when the transistor is in cut off
- Ids depends on the temperature

VELOCITY SATURATION

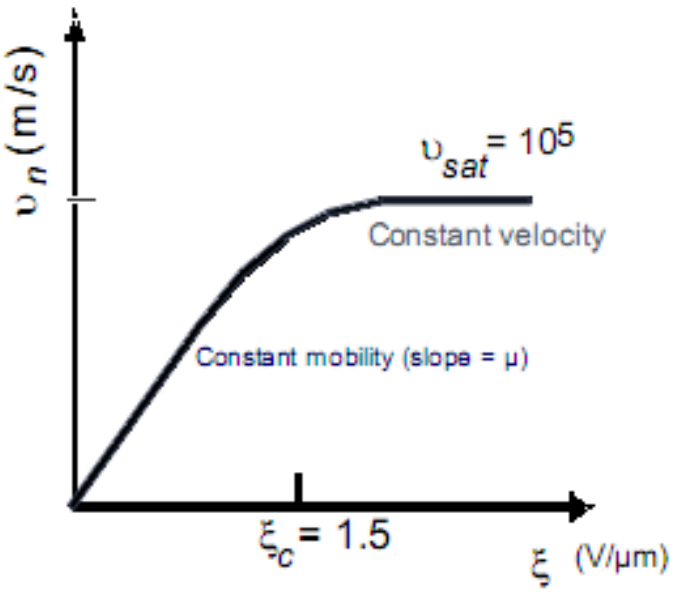
At high electric field, drift velocity rolls of due to carrier scattering

$$v = \mu_n E \text{ for } E \leq E_c$$
$$= v_{sat} = \mu_n E_c \text{ for } E \geq E_c$$

$$I_{ds} = \mu C_{ox} \frac{W}{L} \frac{Vgs - Vt}{2} V_{ds}$$
$$I_{ds} = C_{ox} W \frac{Vgs - Vt}{2} v_{SAT}$$

Empirically:

$I_{ds} \propto (V_{gs} - V_t)^{\alpha}$
where α is close to 1



ALPHA MODEL

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t & \text{cutoff} \\ I_{dsat}\dfrac{V_{ds}}{V_{dsat}} & V_{ds} < V_{dsat} & \text{linear} \\ I_{dsat} & V_{ds} > V_{dsat} & \text{saturation} \end{cases}$$

$$I_{dsat} = P_c \frac{\beta}{2}\left(V_{gs} - V_t\right)^{\alpha}$$

$$V_{dsat} = P_v \left(V_{gs} - V_t\right)^{\alpha/2}$$

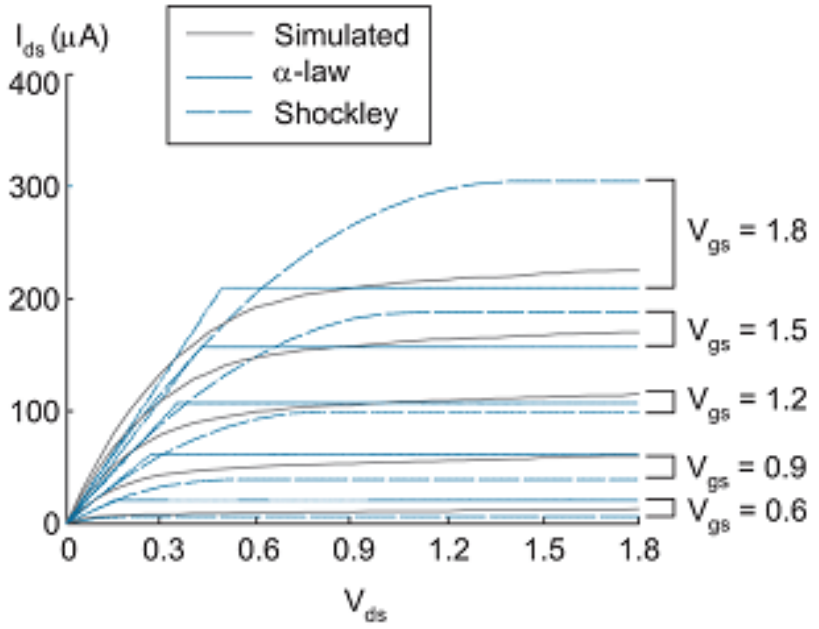


FIG 2.17 I-V characteristics for nMOS transistor with velocity saturation

Pc, Pv and alpha are found by fitting the model to the empirical modeling results

CHANNEL LENGTH MODULATION

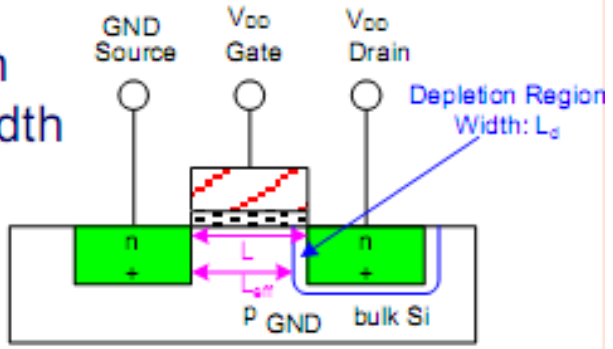
- The reverse-bias p-n junction between drain and body forms a depletion region with a width L_d that increases with V_{db}

$$L = L - L_d$$

- Increasing V_{ds}
  - ➢ increases depletion width
  - ➢ decreases effective channel length
  - ➢ increases current

$$I_{ds} = \beta \frac{(V_{gs} - Vt)^2}{2}(1 + \lambda V_{ds})$$

Channel length modulation factor (empirical factor)



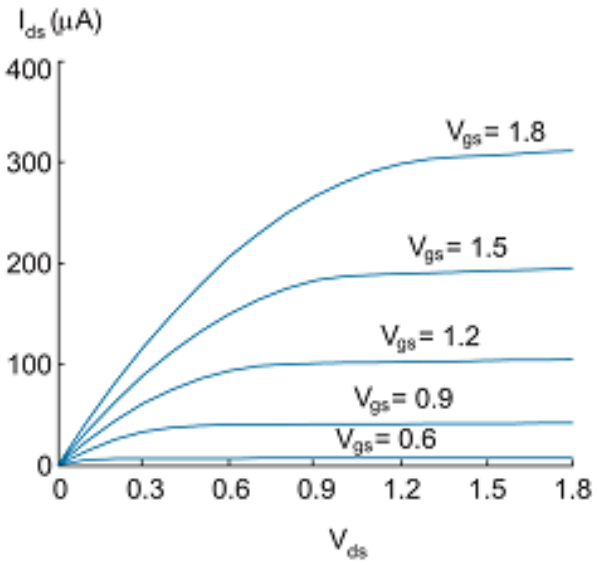
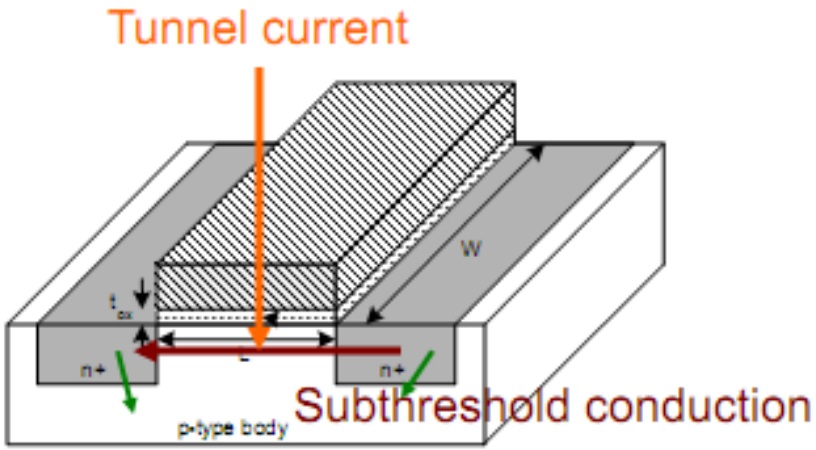
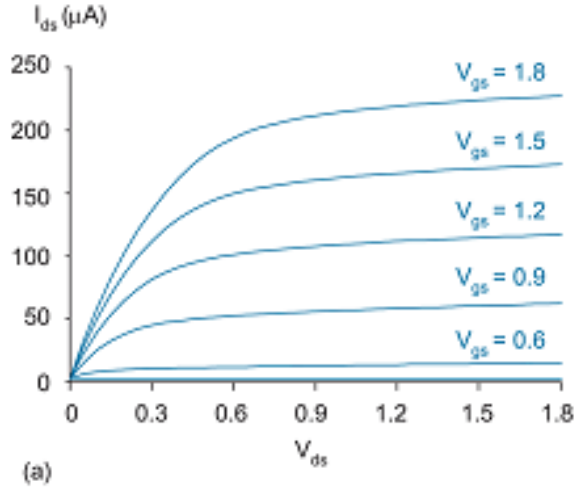


FIG 2.18 I-V characteristics of nMOS transistor with channel length modulation

LEAKAGE CURRENT: SUBTHRESHOLD



Junction leakage



- Subthreshold leakage is the biggest source in modern transistors

$$I_{ds} = I_{ds0} e^{\frac{V_{gs}-V_t}{nv_T}} \left(1 - e^{\frac{-V_{ds}}{v_T}}\right)$$

$$I_{ds0} = \beta v_T^2 e^{1.8}$$

n = 1.4-15

Boltzmann distribution

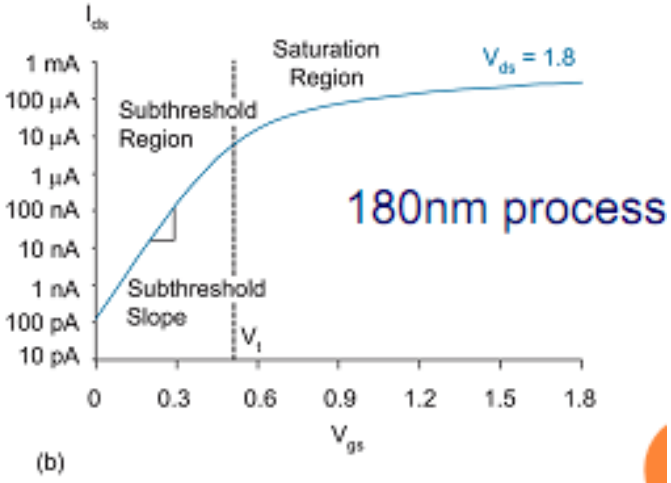
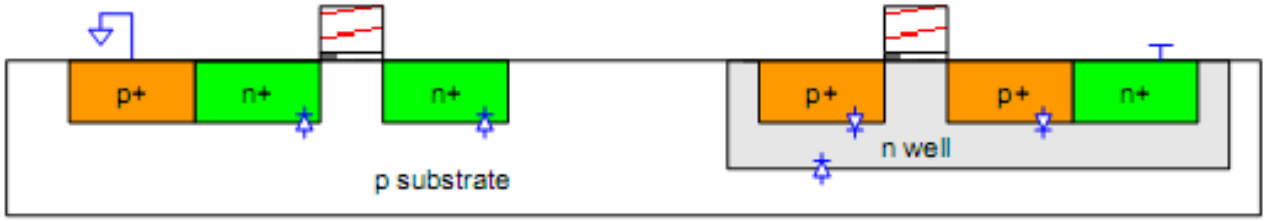


FIG 2.15 Simulated I-V characteristics

Leakage current: junction leakage and tunneling

Junction leakage: reverse-biased p-n junctions have some leakage.
I_s depends on doping levels and area and perimeter of diffusion regions

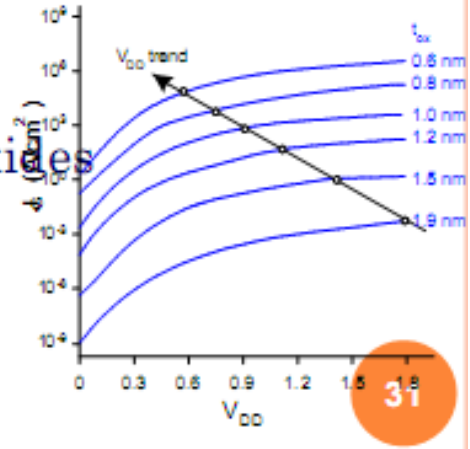
$$I_D = I_S\left(e^{\frac{V_D}{v_T}} - 1\right)$$



Tunneling leakage:

- Carriers may tunnel thorough very thin gate oxides
- Negligible for older processes

(and future processes with high-k dielectrics!)



IMPACT OF TEMPERATURE

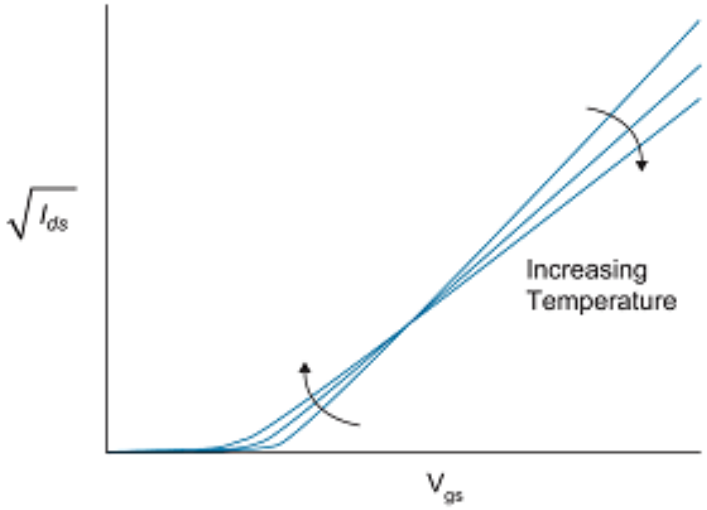


FIG 2.21 I–V characteristics of nMOS transistor in saturation at various temperatures

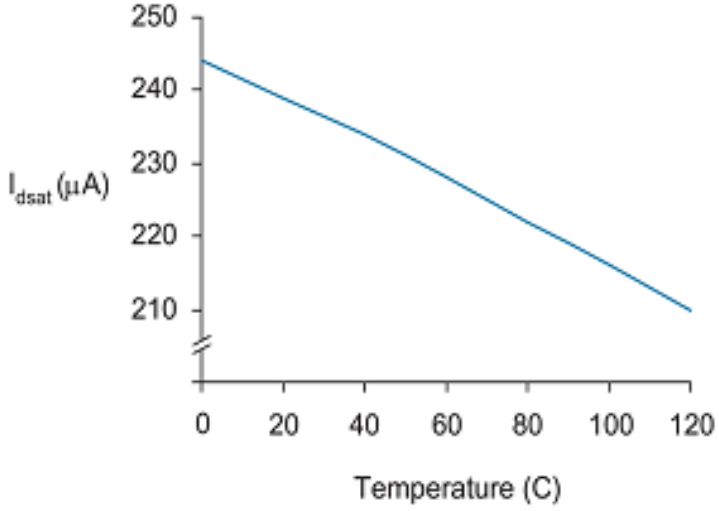


FIG 2.22 I_{dsat} vs. temperature

- Carrier mobility ***decreases*** with T°↑
- Threshold voltage ***decreases*** nearly linearly with T° ↑
- Junction leakage ***increases*** with T° ↑
- ON current ***decreases*** and OFF current ***increases*** with T°↑

→ Circuit performance is generally ***worst*** at high T°↑

→ ***negative temperature coefficient***

IMPACT OF TEMPERATURE (CONT.)

Circuit performance can be improved by cooling

- Subthreshold leakage ***decreases*** with T°↓
- Velocity saturation ***increases*** with T°↓ → more current
- Mobility ***increases*** with T° ↓ → save power
- Depletion regions become ***wider*** with T°↓ → ***less*** junction capacitance

BODY EFFECT

- V_t is sensitive to Vsb -> *body effect*

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

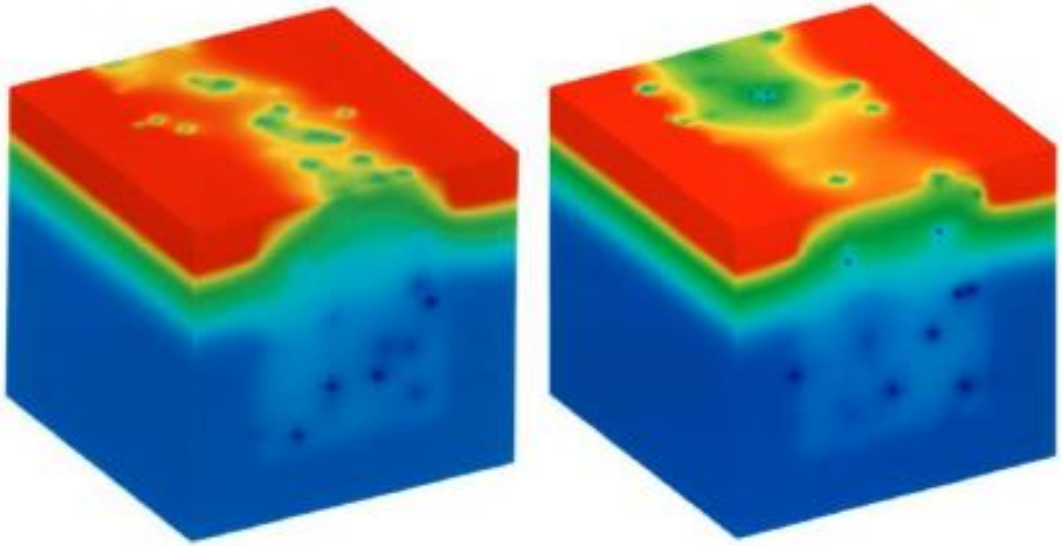
$$\phi_s = 2v_T \ln\frac{N_A}{n_i}$$

$$\gamma = \frac{t_{ox}}{\varepsilon_{ox}}\sqrt{2q\varepsilon_{si}N_A} = \frac{\sqrt{2q\varepsilon_{si}N_A}}{C_{ox}}$$

- What is the impact on Vt if we increase/decrease the body bias?

PROCESS VARIATIONS

Both MOSFETs have 30nm channel with 130 dopant atoms in the channel depletion region



threshold voltage 0.97V

threshold voltage 0.57V

Process variations impact gate length, threshold voltage, and oxide thickness

SUMMARY OF TRANSISTOR OPERATION

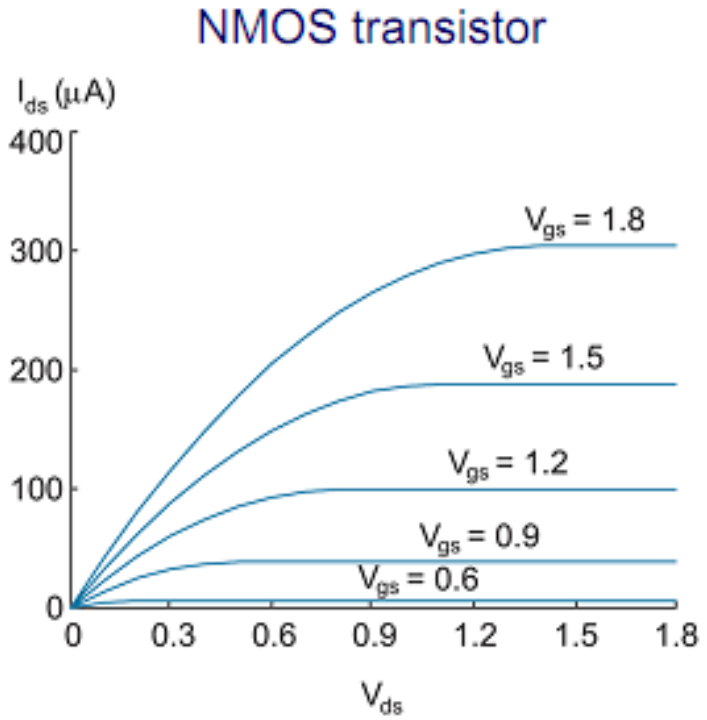


FIG 2.7 I-V characteristics of ideal nMOS transistor

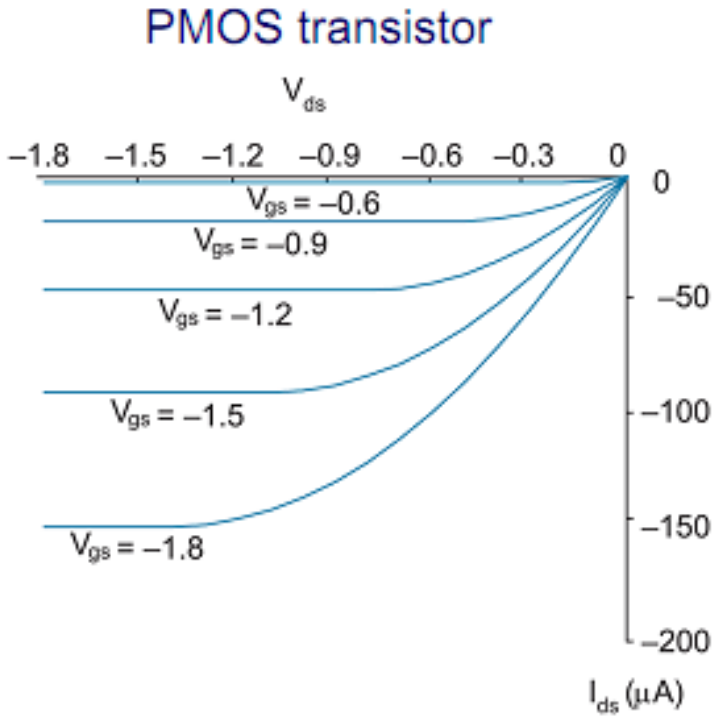
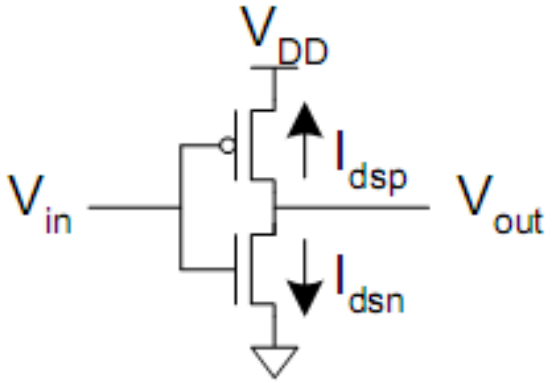


FIG 2.8 I-V characteristics of ideal pMOS transistor

DC Response

- DC Response: V_{out} vs. V_{in} for a gate
- Ex: Inverter
  - When $V_{in} = 0$ -> $V_{out} = V_{DD}$
  - When $V_{in} = V_{DD}$ -> $V_{out} = 0$
  - In between, V_{out} depends on transistor size and current
  - By KCL, must settle such that $I_{dsn} = |I_{dsp}|$
  - We could solve equations
  - But graphical solution gives more insight

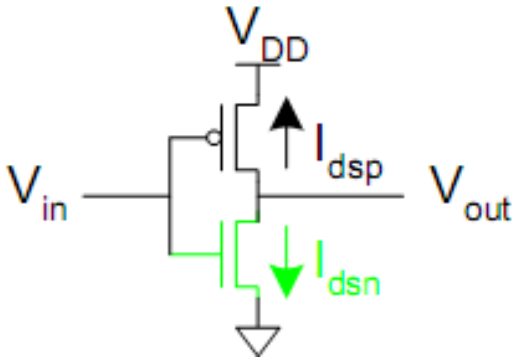


Transistor Operation

- Current depends on region of transistor behavior
- For what V_{in} and V_{out} are nMOS and pMOS in
  - Cutoff?
  - Linear?
  - Saturation?

nMOS Operation

Cutoff	Linear	Saturated
V_{gsn} <	V_{gsn} >	V_{gsn} >
	V_{dsn} <	V_{dsn} >



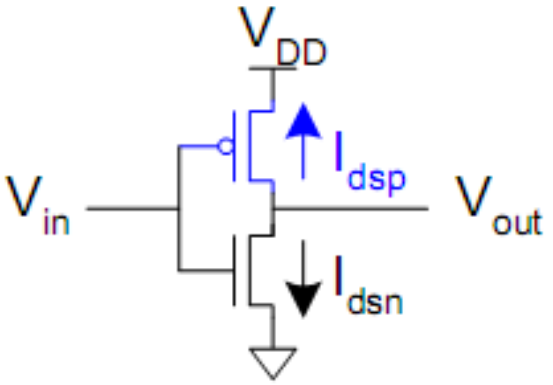
pMOS Operation

Cutoff	Linear	Saturated
$V_{gsp} > V_{tp}$ $V_{in} > V_{DD} + V_{tp}$	$V_{gsp} < V_{tp}$ $V_{in} < V_{DD} + V_{tp}$ $V_{dsp} > V_{gsp} - V_{tp}$ $V_{out} > V_{in} - V_{tp}$	$V_{gsp} < V_{tp}$ $V_{in} < V_{DD} + V_{tp}$ $V_{dsp} < V_{gsp} - V_{tp}$ $V_{out} < V_{in} - V_{tp}$

$V_{gsp} = V_{in} - V_{DD}$

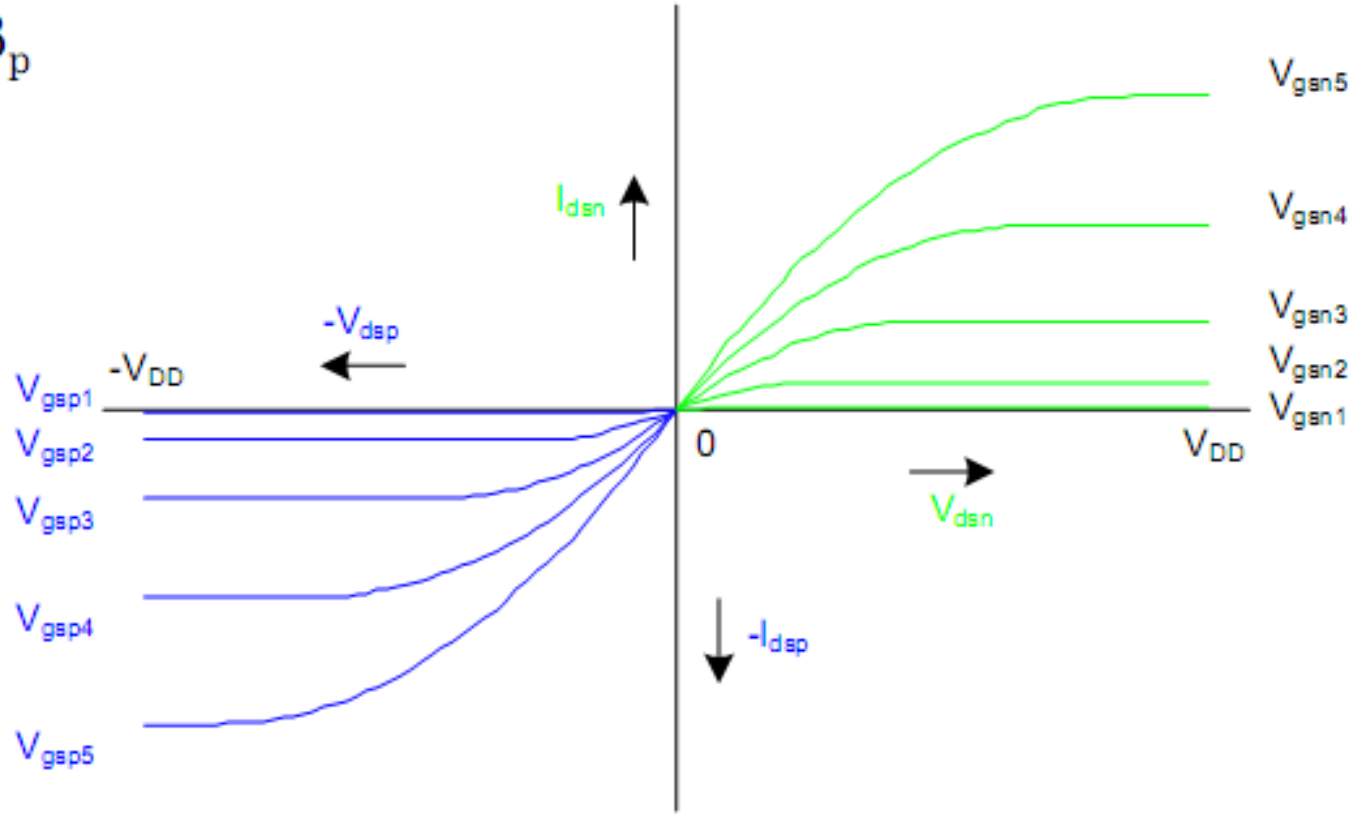
$V_{dsp} = V_{out} - V_{DD}$

$V_{tp} < 0$

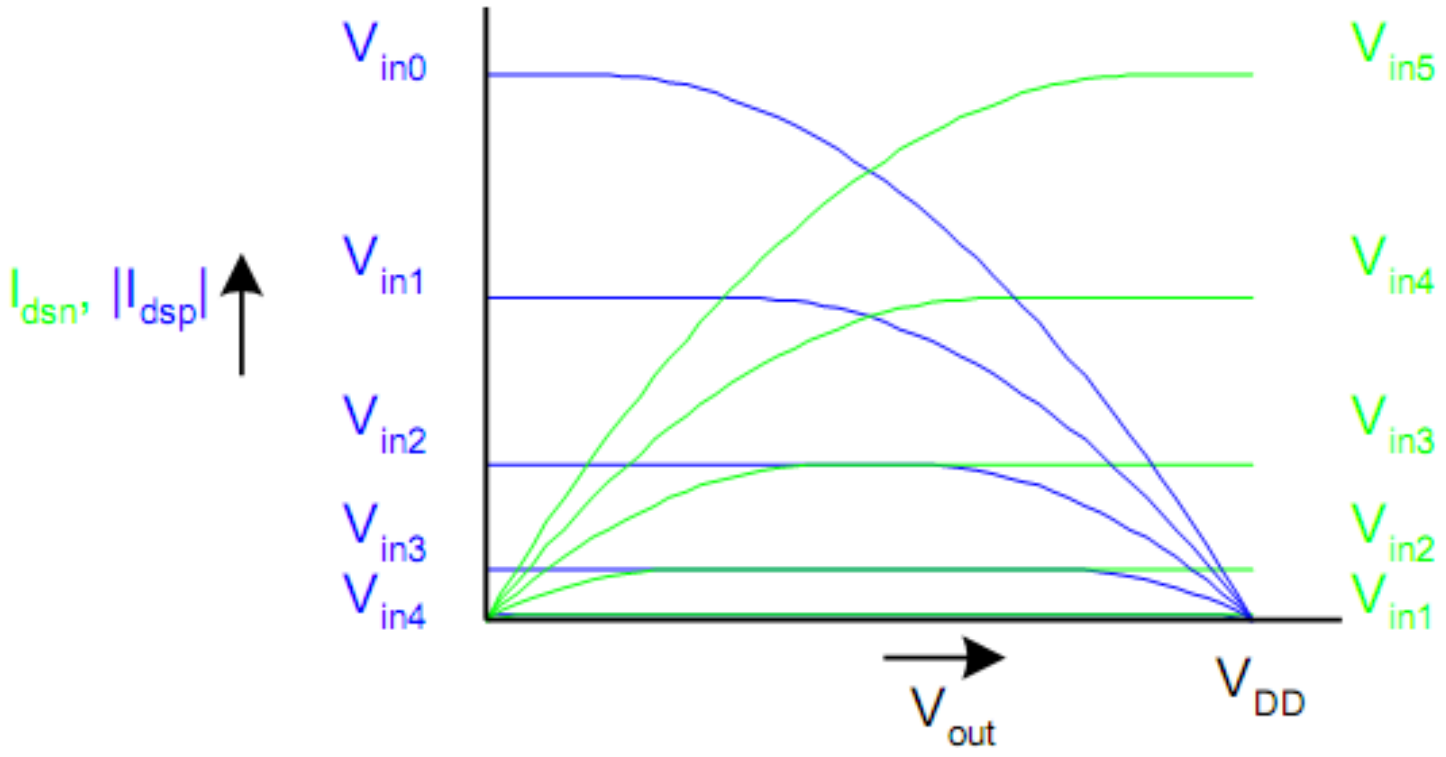


I-V Characteristics

- Make pMOS is wider than nMOS such that $\beta_n = \beta_p$

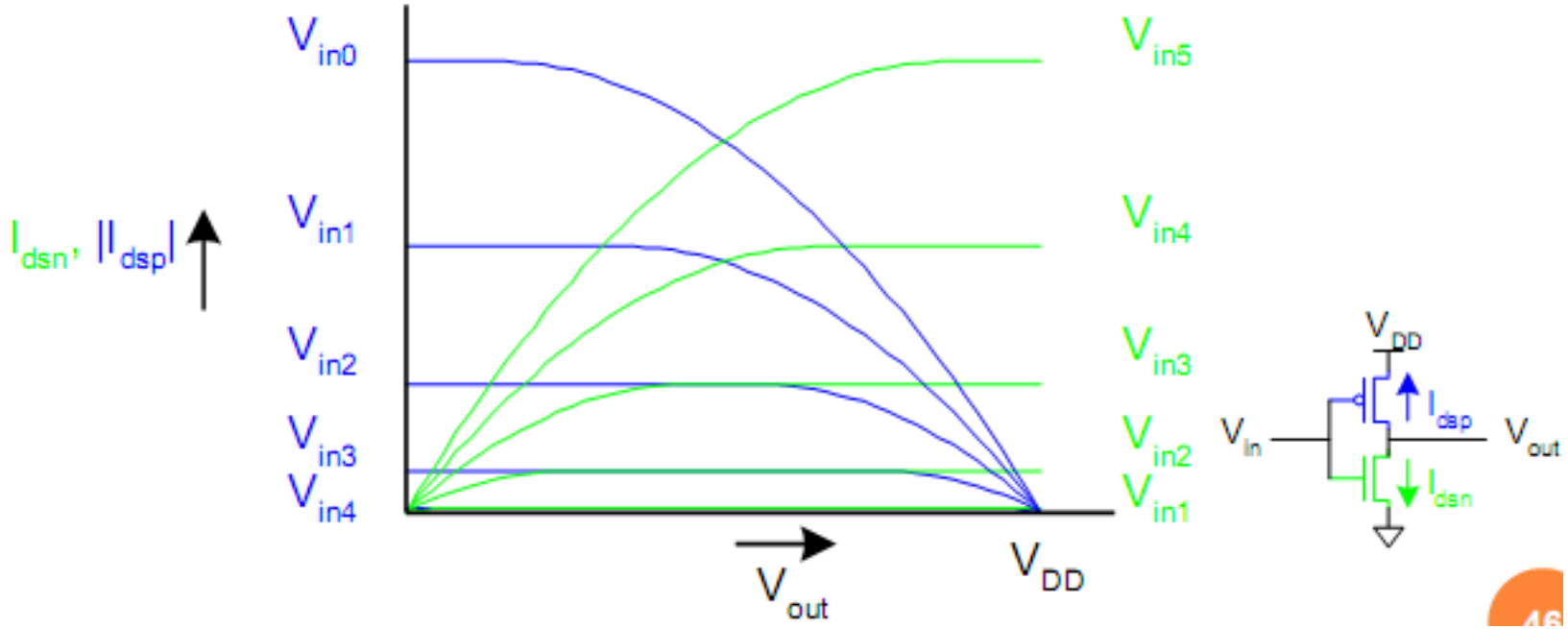


Current vs. V_{OUT}, V_{IN}



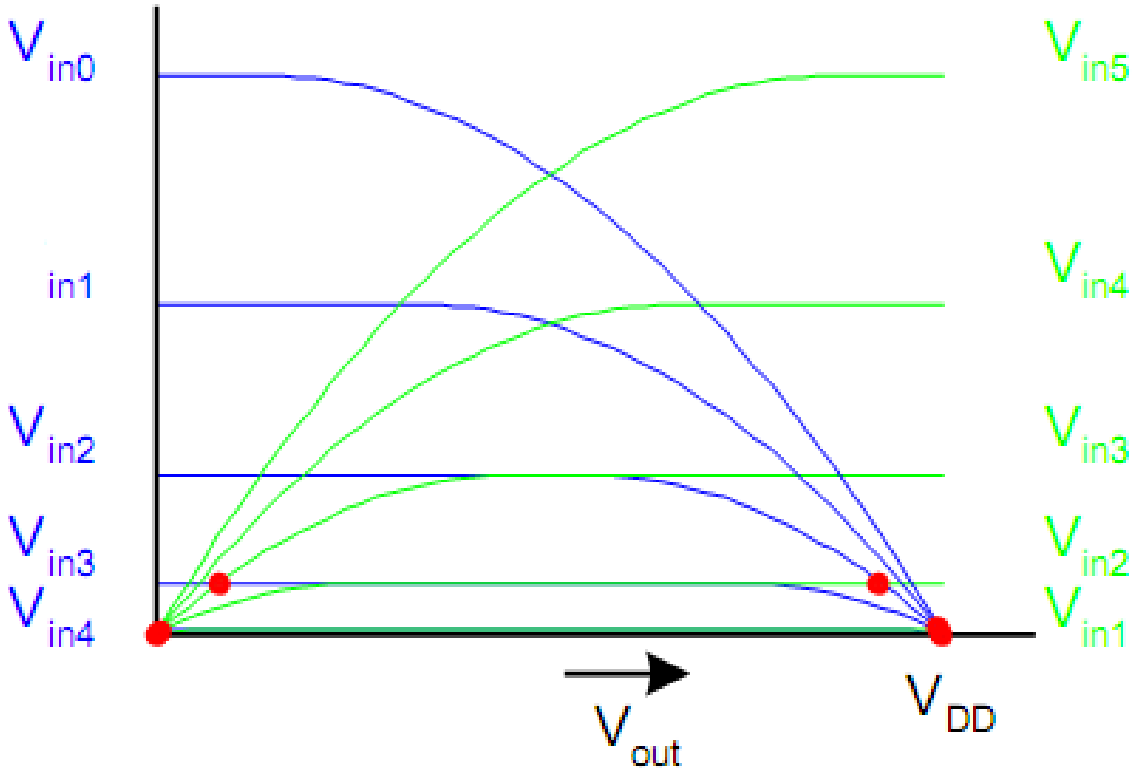
Load Line Analysis

- For a given V_{in}:
  - Plot I_{dsn}, I_{dsp} vs. V_{out}
  - V_{out} must be where |currents| are equal in



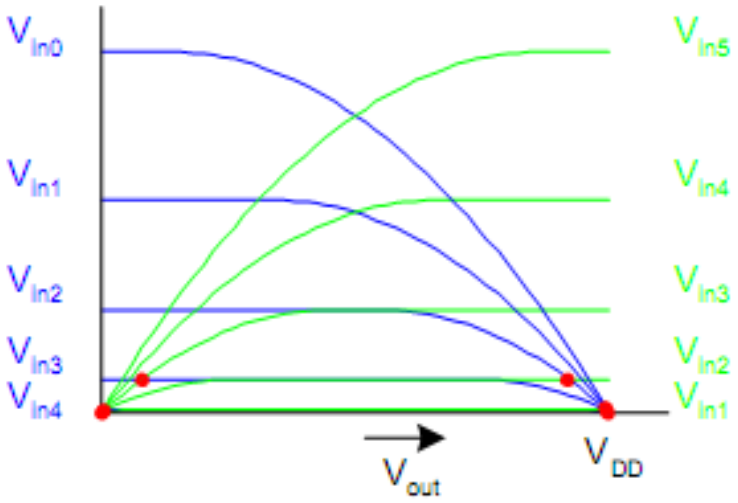
Load Line Analysis

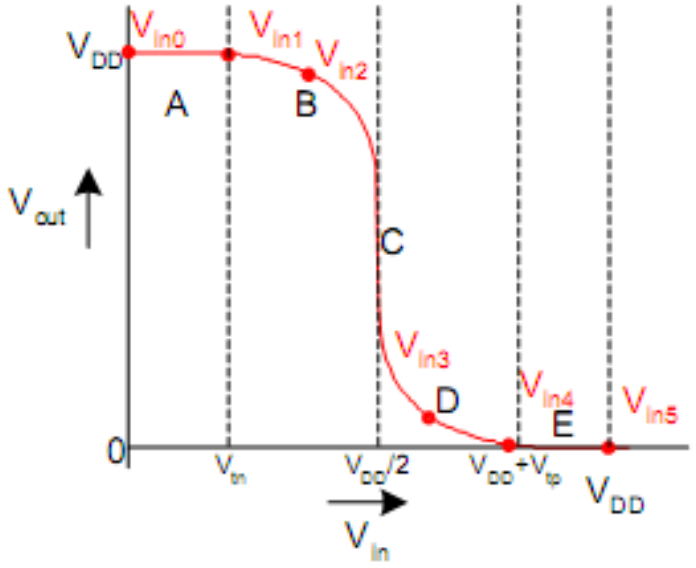
$V_{in} = V_{DD}$



DC Transfer Curve

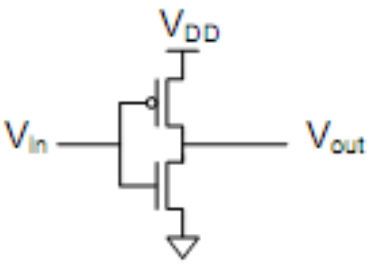
- Transcribe points onto V_{in} vs. V_{out} plot



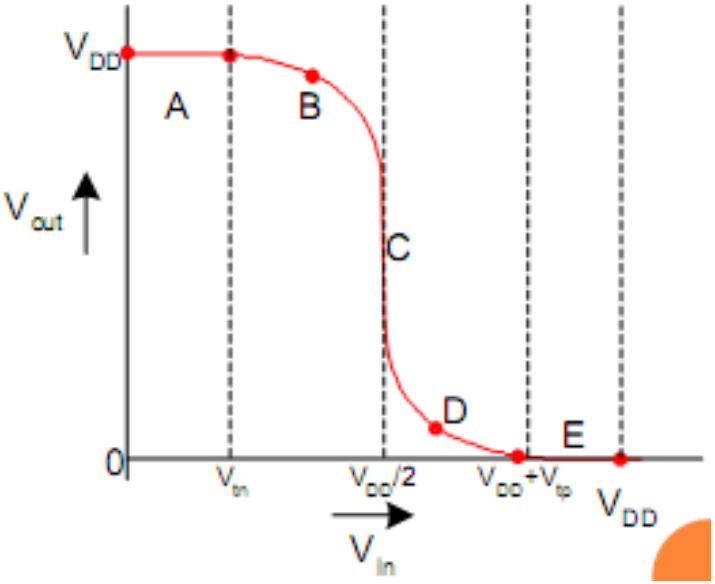


Operating Regions

- Revisit transistor operating regions

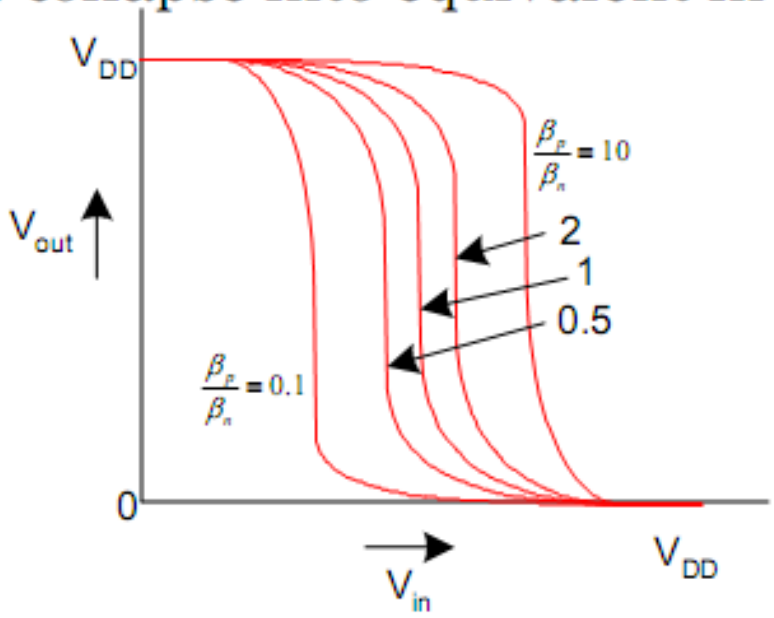


Region	nMOS	pMOS
A		
B		
C		
D		
E		



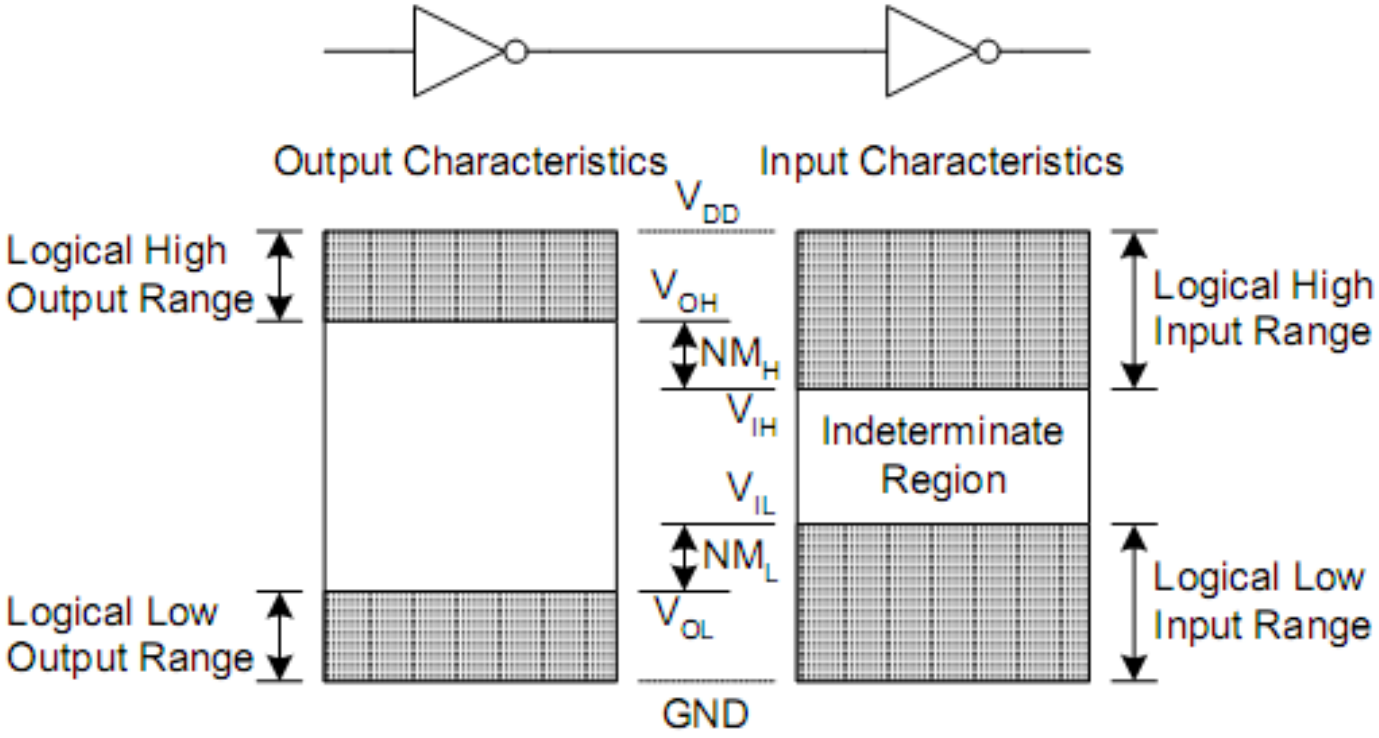
Beta Ratio

- If $\beta_p / \beta_n \neq 1$, switching point will move from $V_{DD}/2$
- Called *skewed* gate
- Other gates: collapse into equivalent inverter



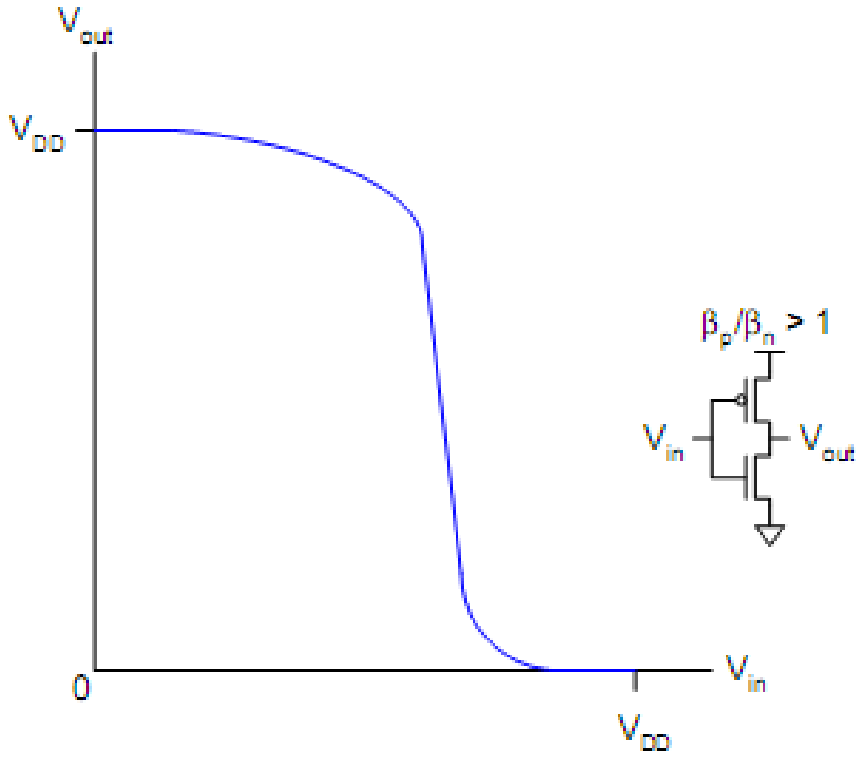
Noise Margins

- How much noise can a gate input see before it does not recognize the input?



Logic Levels

- To maximize noise margins, select logic levels at
  - unity gain point of DC transfer characteristic

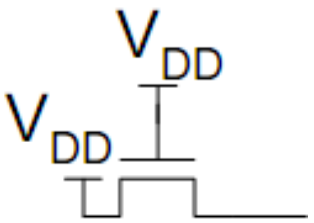


Transient Response

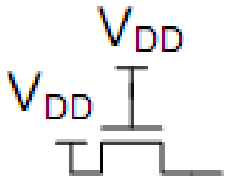
- *DC analysis* tells us V_{out} if V_{in} is constant
- *Transient analysis* tells us $V_{out}(t)$ if $V_{in}(t)$ changes
  - Requires solving differential equations
- Input is usually considered to be a step or ramp
  - From 0 to V_{DD} or vice versa

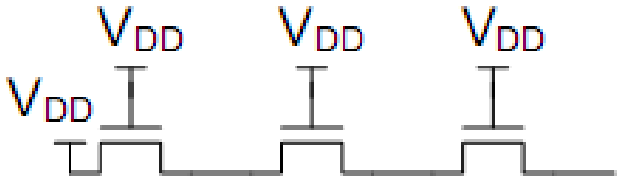
Pass Transistors

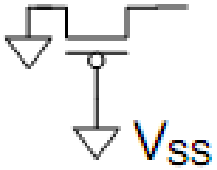
- We have assumed source is grounded
- What if source > 0?
  - e.g. pass transistor passing V_{DD}
- $V_g = V_{DD}$
  - If $V_s > V_{DD}-V_t$, $V_{gs} < V_t$
  - Hence transistor would turn itself off
- nMOS pass transistors pull no higher than $V_{DD}-V_{tn}$
  - Called a degraded "1"
  - Approach degraded value slowly (low I_{ds})
- pMOS pass transistors pull no lower than V_{tp}
- Transmission gates are needed to pass both 0 and 1

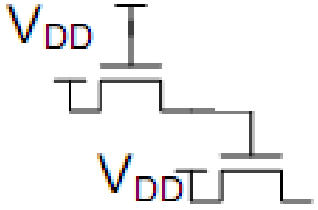


Pass Transistor Ckts

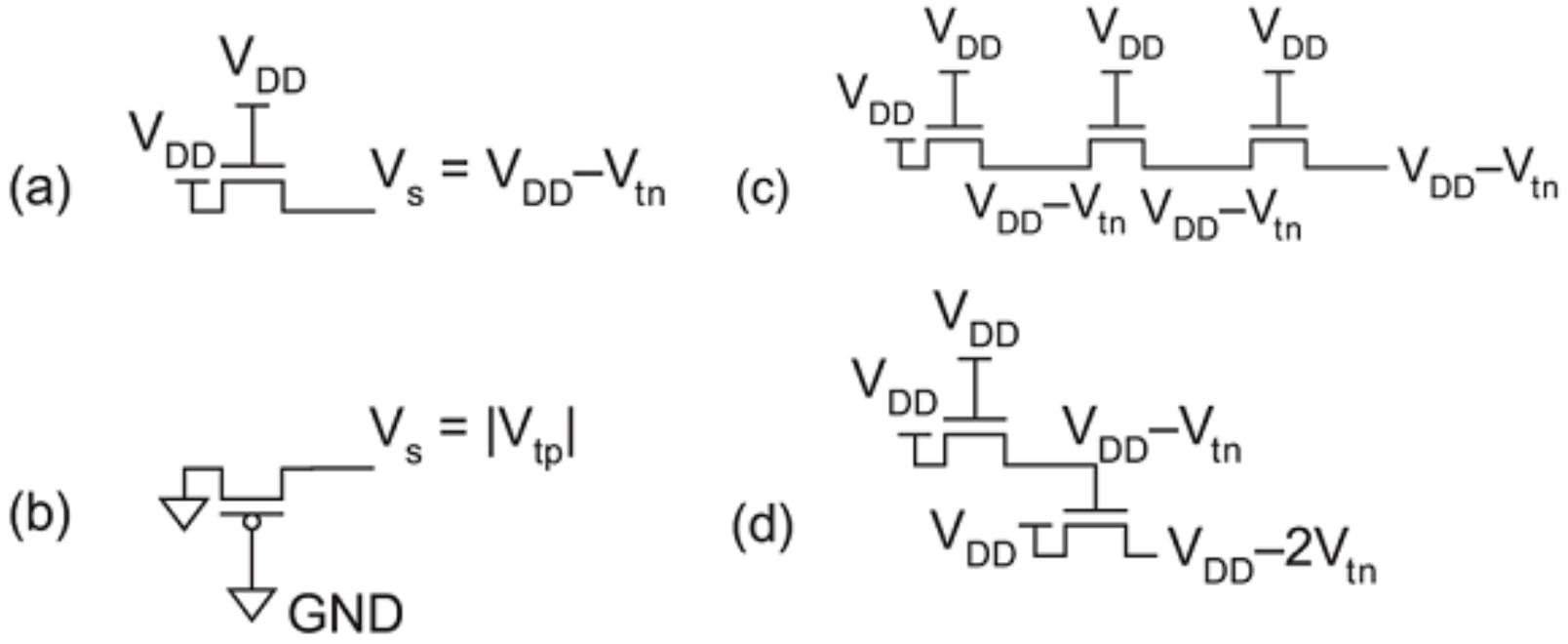








PASS TRANSISTOR DC CHARACTERISTICS



- As the source can rise to within a threshold voltage of the gate, the output of several transistors in series is no more degraded than that of a single transistor

Summary

- Covered ideal (long channel) operation (Shockley model) of transistor
- Short-channel transistors
- TA