

# MOS Transistor Theory

- So far, we have viewed a MOS transistor as an ideal switch (digital operation)
  - Reality: less than ideal

# Introduction

- So far, we have treated transistors as ideal switches
- An ON transistor passes a finite amount of current
  - Depends on terminal voltages
  - Derive current-voltage (I-V) relationships
- Transistor gate, source, drain all have capacitance
  - $I = C (\Delta V / \Delta t) \rightarrow \Delta t = (C / I) \Delta V$
  - Capacitance and current determine speed
- Also explore what a “degraded level” really means

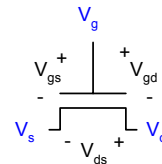


## MOS Transistor Theory

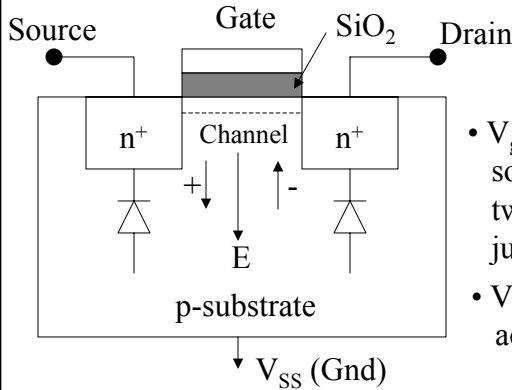
- Study conducting channel between source and drain
  - Modulated by voltage applied to the gate (voltage-controlled device)
  - nMOS transistor: majority carriers are electrons (greater mobility), p-substrate doped (positively doped)
  - pMOS transistor: majority carriers are holes (less mobility), n-substrate (negatively doped)

## Terminal Voltages

- Mode of operation depends on  $V_g$ ,  $V_d$ ,  $V_s$ 
  - $V_{gs} = V_g - V_s$
  - $V_{gd} = V_g - V_d$
  - $V_{ds} = V_d - V_s = V_{gs} - V_{gd}$
- Source and drain are symmetric diffusion terminals
  - By convention, source is terminal at lower voltage
  - Hence  $V_{ds} \geq 0$
- nMOS body is grounded. First assume source is 0 too.
- Three regions of operation
  - *Cutoff*
  - *Linear*
  - *Saturation*



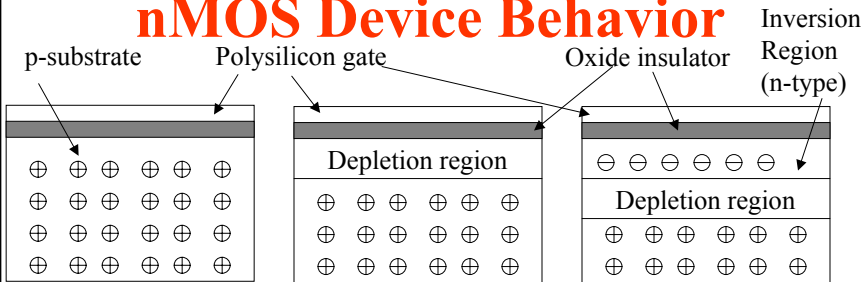
# Gate Biasing



- $V_{gs}=0$ : no current flows from source to drain (insulated by two reverse biased pn junctions)
- $V_{gs}>0$ : electric field created across substrate

- Electrons accumulate under gate: region changes from p-type to n-type
- Conduction path between source and drain

# nMOS Device Behavior



$V_{gs} \ll V_t$

Accumulation mode

$V_{gs} = V_t$

Depletion mode

$V_{gs} > V_t$

Inversion mode

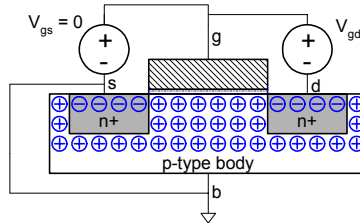
- Enhancement-mode transistor: Conducts when gate bias

$V_{gs} > V_t$

- Depletion-mode transistor: Conducts when gate bias is zero

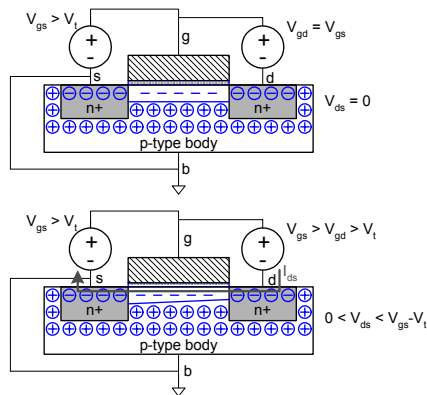
# nMOS Cutoff

- No channel
- $I_{ds} = 0$



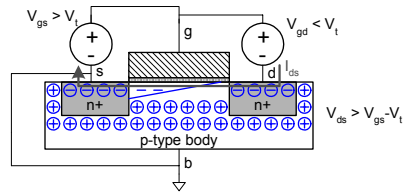
# nMOS Linear

- Channel forms
- Current flows from d to s  
–  $e^-$  from s to d
- $I_{ds}$  increases with  $V_{ds}$
- Similar to linear resistor



## nMOS Saturation

- Channel pinches off
- $I_{ds}$  independent of  $V_{ds}$
- We say current saturates
- Similar to current source

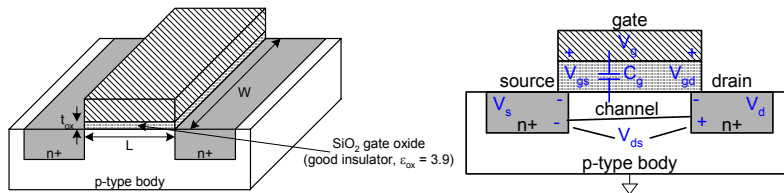


## I-V Characteristics

- In linear region,  $I_{ds}$  depends on
  - How much charge is in the channel?
  - How fast is the charge moving?

# Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel
- $Q_{\text{channel}} =$



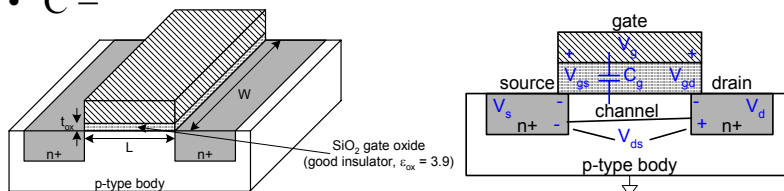
EE 261

Krish Chakrabarty

11

# Channel Charge

- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel
- $Q_{\text{channel}} = CV$
- $C =$



EE 261

Krish Chakrabarty

12

# Channel Charge

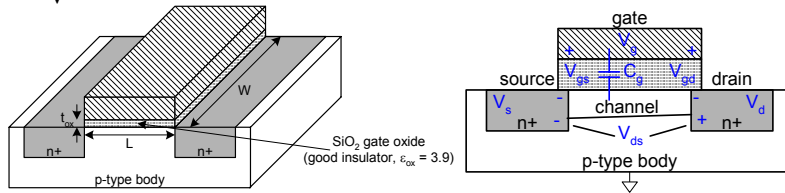
- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel

- $Q_{\text{channel}} = CV$

- $C = C_g = \epsilon_{\text{ox}} WL/t_{\text{ox}} = C_{\text{ox}} WL$

$$C_{\text{ox}} = \epsilon_{\text{ox}} / t_{\text{ox}}$$

- $V =$



# Channel Charge

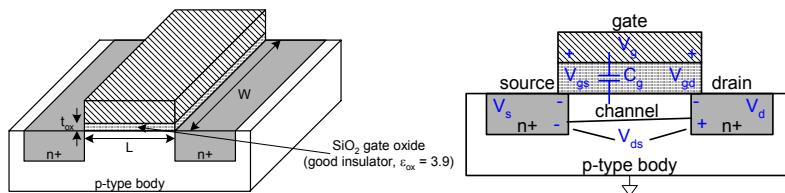
- MOS structure looks like parallel plate capacitor while operating in inversion
  - Gate – oxide – channel

- $Q_{\text{channel}} = CV$

- $C = C_g = \epsilon_{\text{ox}} WL/t_{\text{ox}} = C_{\text{ox}} WL$

$$C_{\text{ox}} = \epsilon_{\text{ox}} / t_{\text{ox}}$$

- $V = V_{\text{gc}} - V_t = (V_{\text{gs}} - V_{\text{ds}}/2) - V_t$



## Carrier velocity

- Charge is carried by e-
- Carrier velocity  $v$  proportional to lateral E-field between source and drain
- $v =$

## Carrier velocity

- Charge is carried by e-
- Carrier velocity  $v$  proportional to lateral E-field between source and drain
- $v = \mu E$                        $\mu$  called mobility
- $E =$



## Carrier velocity

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

## Carrier velocity

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

## nMOS Linear I-V

- Now we know
    - How much charge  $Q_{\text{channel}}$  is in the channel
    - How much time  $t$  each carrier takes to cross
- $$I_{ds} =$$

## nMOS Linear I-V

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

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

=

## nMOS Linear I-V

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

$$\begin{aligned}I_{ds} &= \frac{Q_{\text{channel}}}{t} \\ &= \mu C_{\text{ox}} \frac{W}{L} \left( V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds} \\ &= \beta \left( V_{gs} - V_t - \frac{V_{ds}}{2} \right) V_{ds} \quad \beta = \mu C_{\text{ox}} \frac{W}{L}\end{aligned}$$

## nMOS Saturation I-V

- If  $V_{gd} < V_t$ , channel pinches off near drain
  - When  $V_{ds} > V_{\text{dsat}} = V_{gs} - V_t$
- Now drain voltage no longer increases current

$$I_{ds} =$$

## nMOS Saturation I-V

- If  $V_{gd} < V_t$ , channel pinches off near drain
  - When  $V_{ds} > V_{dsat} = V_{gs} - V_t$
- Now drain voltage no longer increases current

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

## nMOS Saturation I-V

- If  $V_{gd} < V_t$ , channel pinches off near drain
  - When  $V_{ds} > V_{dsat} = V_{gs} - V_t$
- Now drain voltage no longer increases current

$$\begin{aligned} I_{ds} &= \beta \left( V_{gs} - V_t - \frac{V_{dsat}}{2} \right) V_{dsat} \\ &= \frac{\beta}{2} (V_{gs} - V_t)^2 \end{aligned}$$

## nMOS I-V Summary

- Shockley 1<sup>st</sup> order transistor models

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

## Current-Voltage Relations

**Linear Region:  $V_{Ds} \leq V_{GS} - V_T$**

$$I_D = k_n \frac{W}{L} \left( (V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2} \right)$$

with

$$k_n = \mu_n C_{ox} = \frac{\mu_n \epsilon_{ox}}{t_{ox}} \quad \begin{array}{l} \text{Process Transconductance} \\ \text{Parameter} \end{array}$$

**Saturation Mode:  $V_{Ds} \geq V_{GS} - V_T$**  Channel Length Modulation

$$I_D = \frac{k_n W}{2 L} (V_{GS} - V_T)^2 (1 + \lambda V_{DS})$$

# Current-Voltage Relations

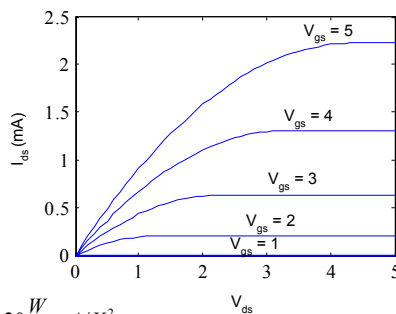
$k'_n$ : transconductance of transistor

$\frac{W}{L}$ : 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

## Example

- For a 0.6  $\mu\text{m}$  process
  - From AMI Semiconductor
  - $t_{\text{ox}} = 100 \text{ \AA}$
  - $\mu = 350 \text{ cm}^2/\text{V}\cdot\text{s}$
  - $V_t = 0.7 \text{ V}$
- Plot  $I_{ds}$  vs.  $V_{ds}$ 
  - $V_{gs} = 0, 1, 2, 3, 4, 5$
  - Use  $W/L = 4/2 \lambda$



$$\beta = \mu C_{\text{ox}} \frac{W}{L} = (350) \left( \frac{3.9 \cdot 8.85 \cdot 10^{-14}}{100 \cdot 10^{-8}} \right) \left( \frac{W}{L} \right) = 120 \frac{W}{L} \mu\text{A}/\text{V}^2$$

## pMOS I-V

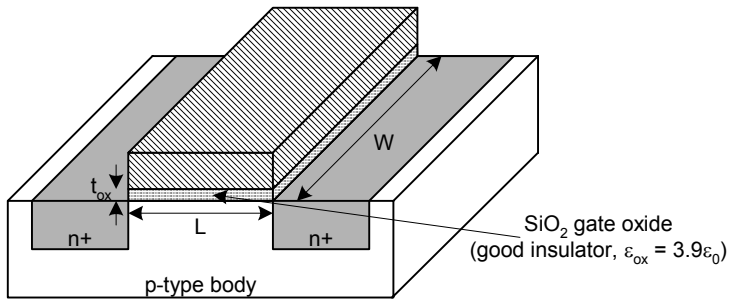
- All dopings and voltages are inverted for pMOS
- Mobility  $\mu_p$  is determined by holes
  - Typically 2-3x lower than that of electrons  $\mu_n$
  - 120 cm<sup>2</sup>/V\*s in AMI 0.6  $\mu$ m process
- Thus pMOS must be wider to provide same current
  - In this class, assume  $\mu_n / \mu_p = 2$  to 3

## Capacitance

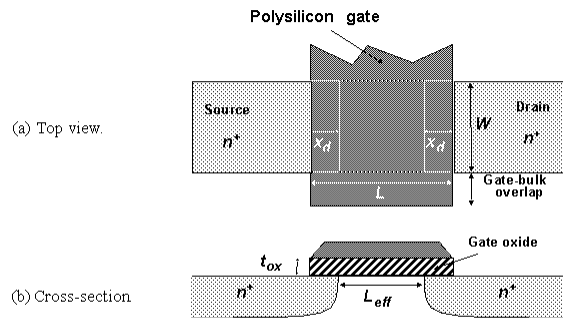
- Any two conductors separated by an insulator have capacitance
- Gate to channel capacitor is very important
  - Creates channel charge necessary for operation
- Source and drain have capacitance to body
  - Across reverse-biased diodes
  - Called diffusion capacitance because it is associated with source/drain diffusion

# Gate Capacitance

- Approximate channel as connected to source
- $C_{gs} = \epsilon_{ox} WL/t_{ox} = C_{ox} WL = C_{permicron} W$
- $C_{permicron}$  is typically about 2 fF/ $\mu\text{m}$



# The Gate Capacitance

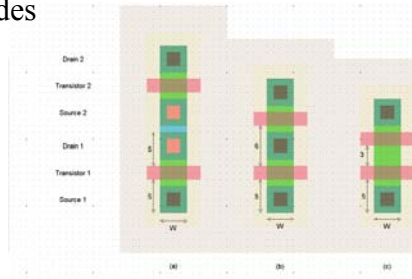


$$C_{gate} = \frac{\epsilon_{ox}}{t_{ox}} WL$$

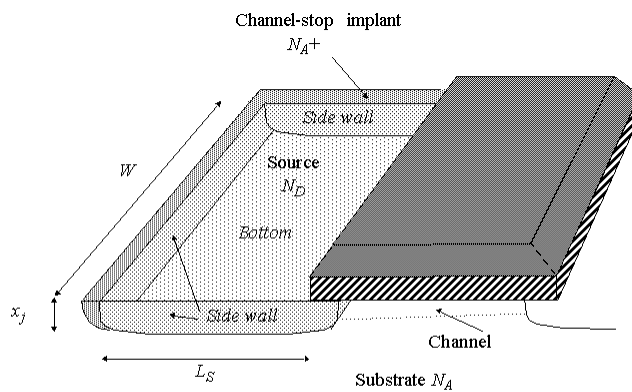


# Diffusion Capacitance

- $C_{sb}$ ,  $C_{db}$
- Undesirable, called *parasitic* capacitance
- Capacitance depends on area and perimeter
  - Use small diffusion nodes
  - Comparable to  $C_g$  for contacted diff
  - $\frac{1}{2} C_g$  for uncontacted
  - Varies with process



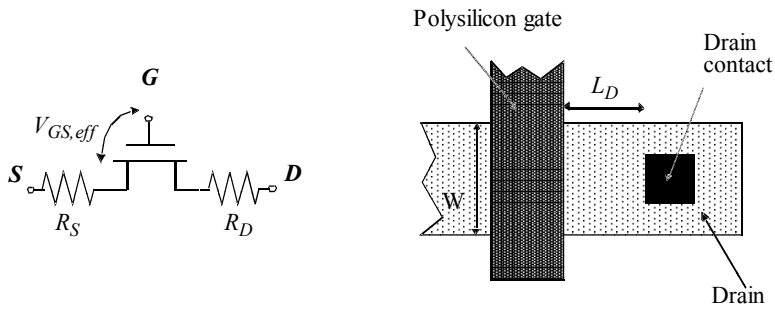
# Diffusion Capacitance



$$C_{diff} = C_{bottom} + C_{sw} = C_j \times AREA + C_{jsw} \times PERIMETER$$

$$= C_j L_S W + C_{jsw} (2L_S + W)$$

## Parasitic Resistances



$$R_S = (L_S/W)R_{\square} + R_C$$

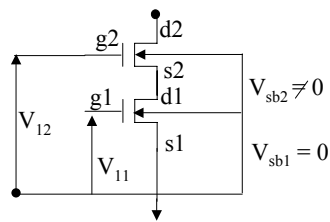
$$R_D = (L_D/W)R_{\square} + R_C$$

$R_C$ : contact resistance

$R_{\square}$ : sheet resistance per square of drain-source diffusion

## Body Effect

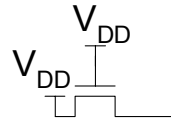
- Many MOS devices on a common substrate
  - Substrate voltage of all devices are normally equal
- But several devices may be connected in series
  - Increase in source-to-substrate voltage as we proceed vertically along the chain



- Net effect: slight increase in threshold voltage  $V_t$   
 $V_{t2} > V_{t1}$

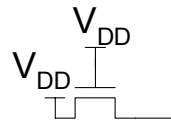
## Pass Transistors

- We have assumed source is grounded
- What if source  $> 0$ ?
  - e.g. pass transistor passing  $V_{DD}$

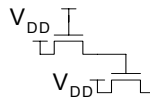
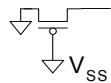
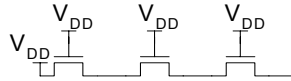
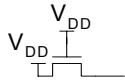


## Pass Transistors

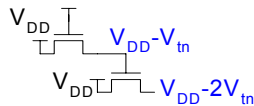
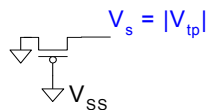
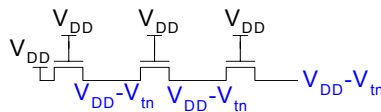
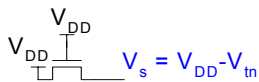
- We have assumed source is grounded
- What if source  $> 0$ ?
  - e.g. pass transistor passing  $V_{DD}$
- $V_g = V_{DD}$ 
  - If  $V_s > V_{DD} - V_t$ ,  $V_{gs} < V_t$
  - Hence transistor would turn itself off
- nMOS pass transistors pull no higher than  $V_{DD} - V_{tn}$ 
  - Called a degraded “1”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



# Pass Transistor Ckts



# Pass Transistor Ckts

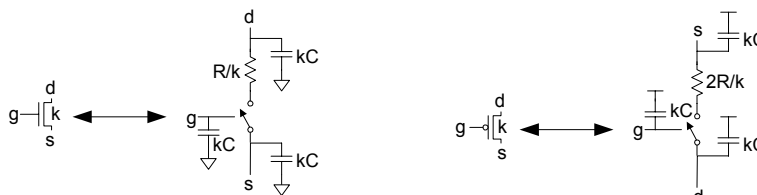


## Effective Resistance

- Shockley models have limited value
  - Not accurate enough for modern transistors
  - Too complicated for much hand analysis
- Simplification: treat transistor as resistor
  - Replace  $I_{ds}(V_{ds}, V_{gs})$  with effective resistance  $R$ 
    - $I_{ds} = V_{ds}/R$
  - $R$  averaged across switching of digital gate
- Too inaccurate to predict current at any given time
  - But good enough to predict RC delay

## RC Delay Model

- Use equivalent circuits for MOS transistors
  - Ideal switch + capacitance and ON resistance
  - Unit nMOS has resistance  $R$ , capacitance  $C$
  - Unit pMOS has resistance  $2R$ , capacitance  $C$
- Capacitance proportional to width
- Resistance inversely proportional to width



## RC Values

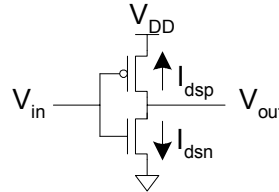
- Capacitance
  - $C = C_g = C_s = C_d = 2 \text{ fF}/\mu\text{m}$  of gate width
  - Values similar across many processes
- Resistance
  - $R \approx 6 \text{ K}\Omega$  in  $0.6\mu\text{m}$  process
  - Improves with shorter channel lengths
- Unit transistors
  - May refer to minimum contacted device ( $4/2 \lambda$ )
  - Or maybe  $1 \mu\text{m}$  wide device
  - Doesn't matter as long as you are consistent

## Activity

- 1) If the width of a transistor increases, the current will  
increase                      decrease                      not change
- 2) If the length of a transistor increases, the current will  
increase                      decrease                      not change
- 3) If the supply voltage of a chip increases, the maximum transistor current will  
increase                      decrease                      not change
- 4) If the width of a transistor increases, its gate capacitance will  
increase                      decrease                      not change
- 5) If the length of a transistor increases, its gate capacitance will  
increase                      decrease                      not change
- 6) If the supply voltage of a chip increases, the gate capacitance of each transistor will  
increase                      decrease                      not change

## DC Response

- DC Response:  $V_{out}$  vs.  $V_{in}$  for a gate
- Ex: Inverter
  - When  $V_{in} = 0$   $\rightarrow V_{out} = V_{DD}$
  - When  $V_{in} = V_{DD}$   $\rightarrow V_{out} = 0$
  - 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

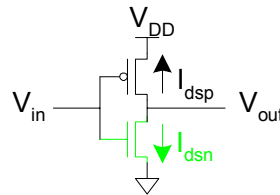


## Transistor Operation

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

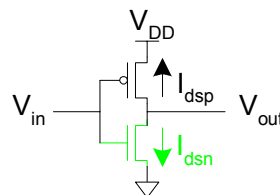
## nMOS Operation

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



## nMOS Operation

Cutoff	Linear	Saturated
$V_{gsn} < V_{tn}$	$V_{gsn} > V_{tn}$	$V_{gsn} > V_{tn}$
	$V_{dsn} < V_{gsn} - V_{tn}$	$V_{dsn} > V_{gsn} - V_{tn}$



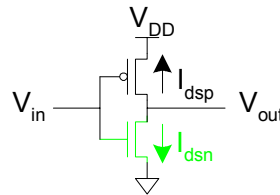


## nMOS Operation

Cutoff	Linear	Saturated
$V_{gsn} < V_{tn}$	$V_{gsn} > V_{tn}$	$V_{gsn} > V_{tn}$
	$V_{dsn} < V_{gsn} - V_{tn}$	$V_{dsn} > V_{gsn} - V_{tn}$

$$V_{gsn} = V_{in}$$

$$V_{dsn} = V_{out}$$

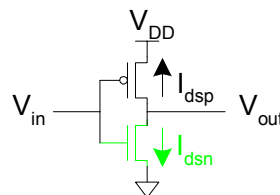


## nMOS Operation

Cutoff	Linear	Saturated
$V_{gsn} < V_{tn}$	$V_{gsn} > V_{tn}$	$V_{gsn} > V_{tn}$
$V_{in} < V_{tn}$	$V_{in} > V_{tn}$	$V_{in} > V_{tn}$
	$V_{dsn} < V_{gsn} - V_{tn}$	$V_{dsn} > V_{gsn} - V_{tn}$
	$V_{out} < V_{in} - V_{tn}$	$V_{out} > V_{in} - V_{tn}$

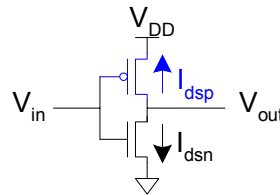
$$V_{gsn} = V_{in}$$

$$V_{dsn} = V_{out}$$



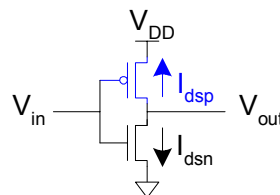
## pMOS Operation

Cutoff	Linear	Saturated
$V_{gsp} >$	$V_{gsp} <$	$V_{gsp} <$
	$V_{dsp} >$	$V_{dsp} <$



## pMOS Operation

Cutoff	Linear	Saturated
$V_{gsp} > V_{tp}$	$V_{gsp} < V_{tp}$	$V_{gsp} < V_{tp}$
	$V_{dsp} > V_{gsp} - V_{tp}$	$V_{dsp} < V_{gsp} - V_{tp}$



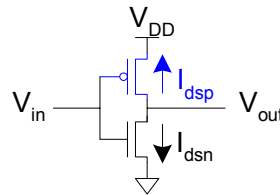
## pMOS Operation

Cutoff	Linear	Saturated
$V_{gsp} > V_{tp}$	$V_{gsp} < V_{tp}$	$V_{gsp} < V_{tp}$
	$V_{dsp} > V_{gsp} - V_{tp}$	$V_{dsp} < V_{gsp} - V_{tp}$

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

$$V_{tp} < 0$$

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



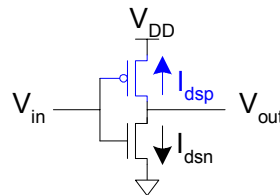
## pMOS Operation

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

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

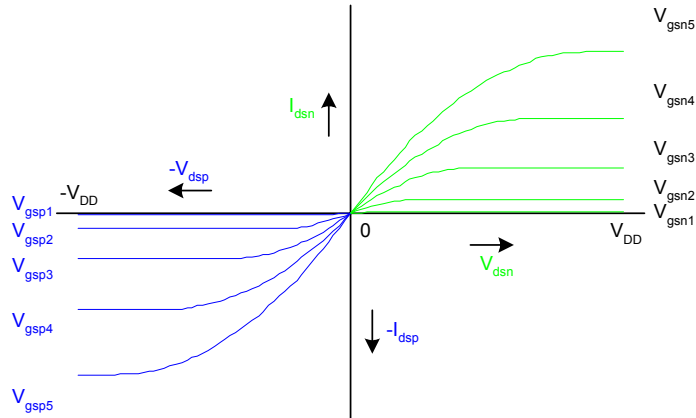
$$V_{tp} < 0$$

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



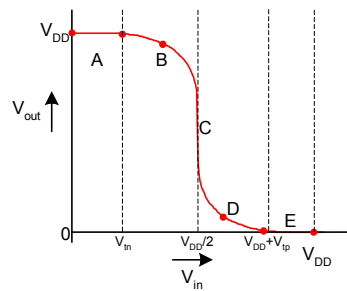
## I-V Characteristics

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



## DC Transfer Curve

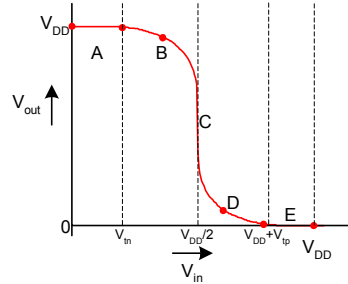
- Transcribe points onto  $V_{in}$  vs.  $V_{out}$  plot



# Operating Regions

- Revisit transistor operating regions

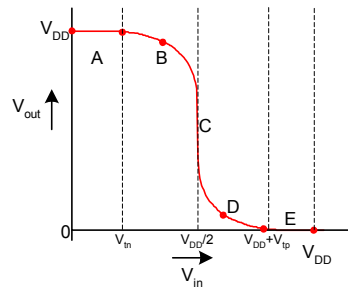
Region	nMOS	pMOS
A		
B		
C		
D		
E		



# Operating Regions

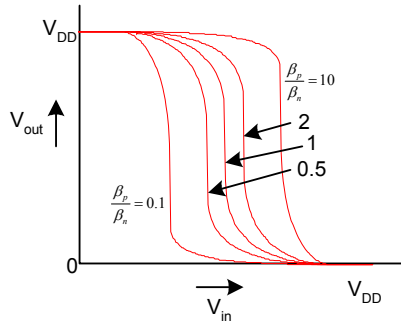
- Revisit transistor operating regions

Region	nMOS	pMOS
A	Cutoff	Linear
B	Saturation	Linear
C	Saturation	Saturation
D	Linear	Saturation
E	Linear	Cutoff



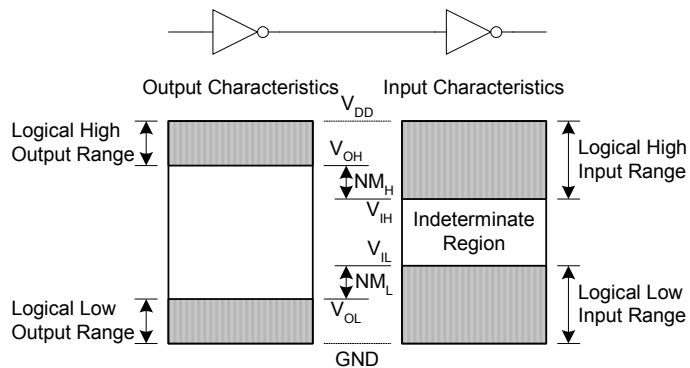
## Beta Ratio

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



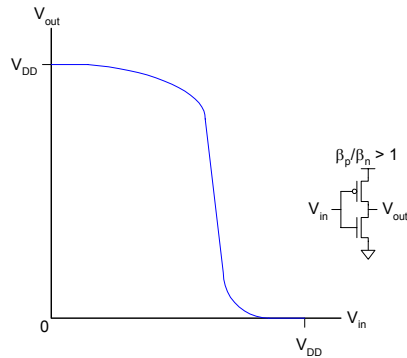
## Noise Margins

- 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



# Logic Levels

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

