EE3230 Lecture 3: MOS Transistor Theory

Ping-Hsuan Hsieh (謝秉璇)

Delta Building R908

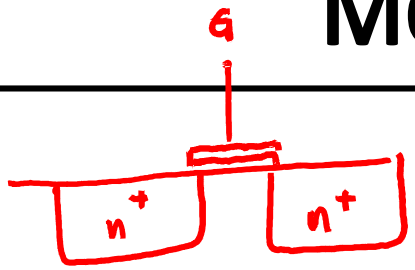
EXT 42590

phsieh@ee.nthu.edu.tw

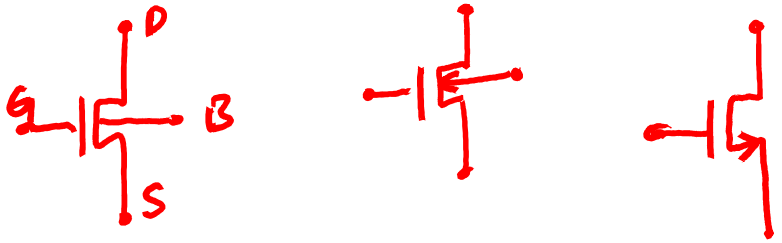
Outline

- **Introduction**
- Ideal I-V characteristics
- Nonideal Effects
- DC transfer characteristics
- C-V characteristics
- Switch-level RC delay models

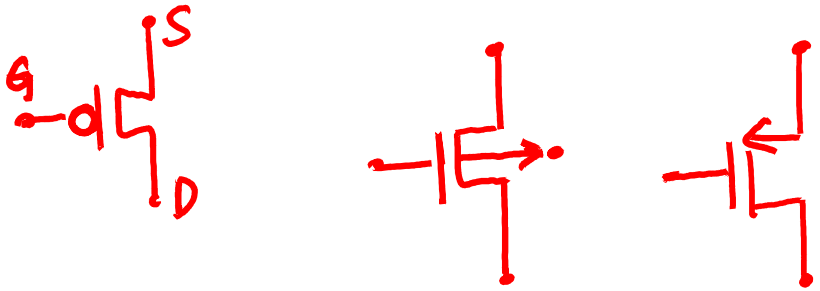
MOS Transistor Symbols



NMOS

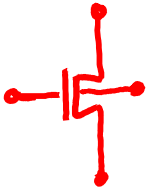


PMOS



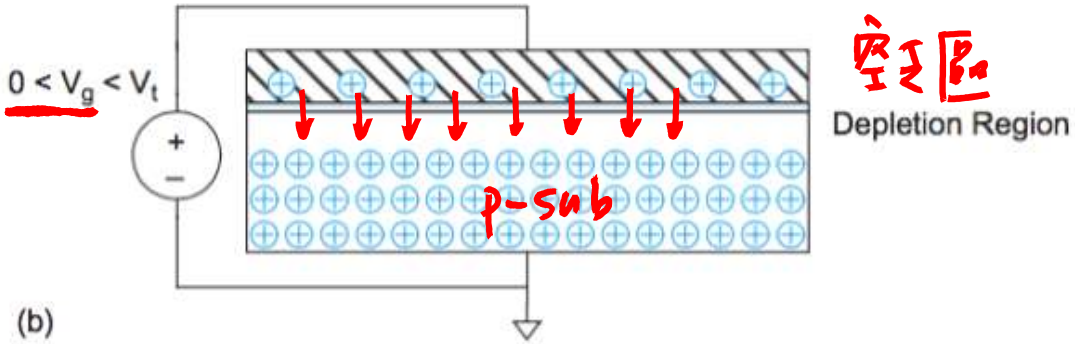
MOS Structure

NMOS



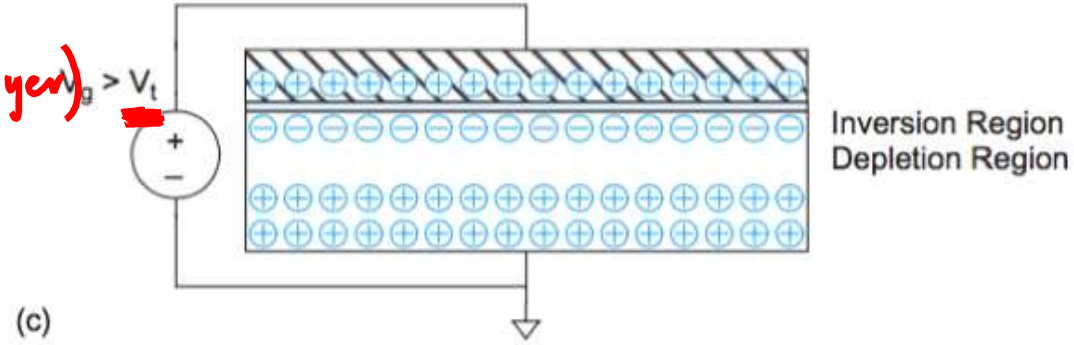
$V_{gb} > 0$

- Depletion mode



- Inversion mode

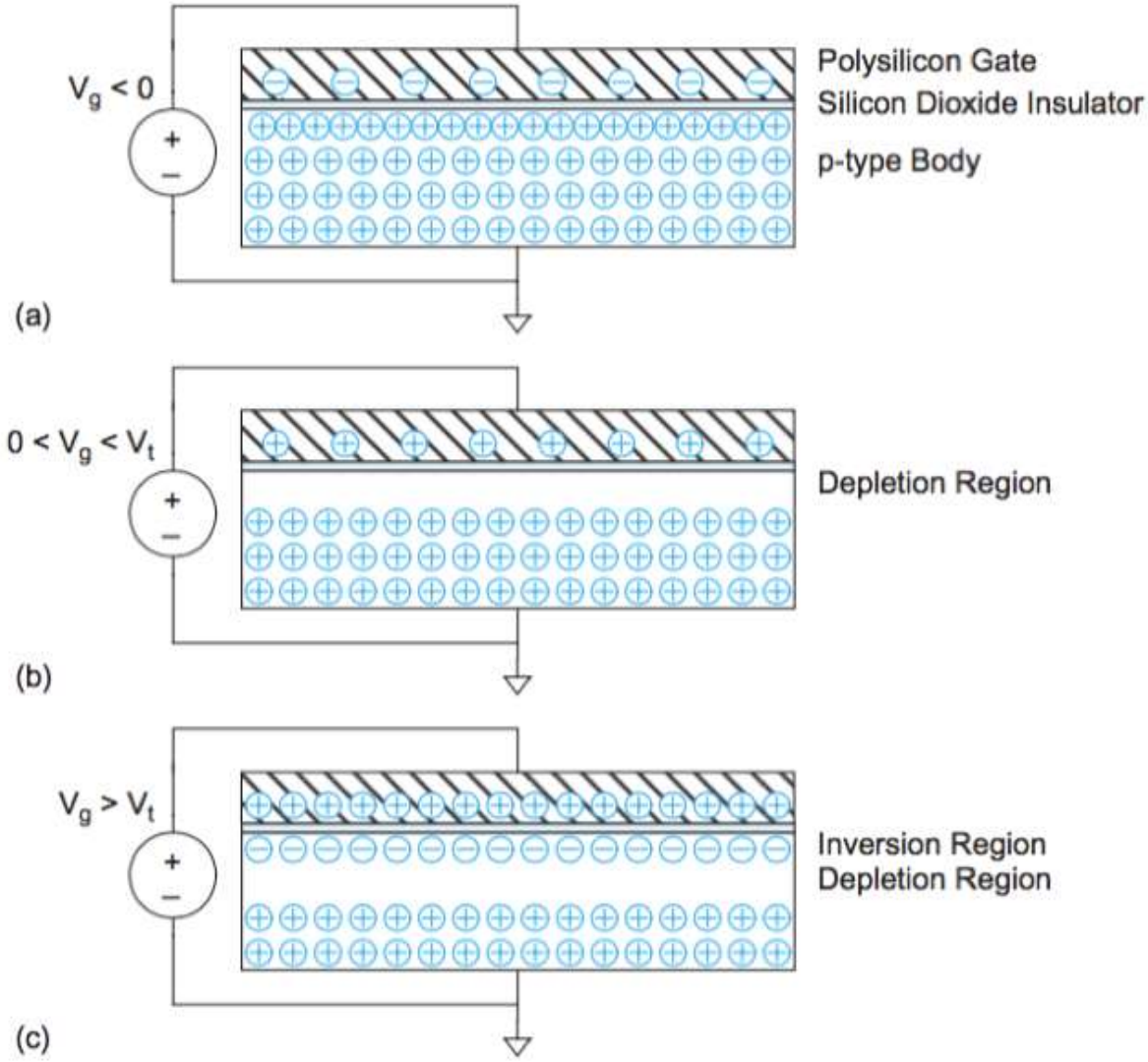
形成 channel (inversion layer)



process dependant
(doping level
oxide thickness

MOS Structure

- Accumulation mode
- Depletion mode
- Inversion mode



NMOS Operating Regions (I, II)

- Cutoff

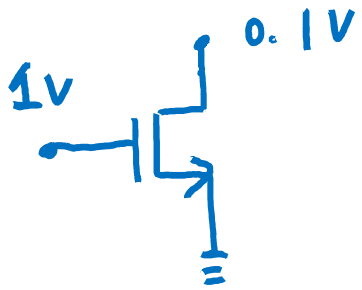
  $V_{gs} < V_{th}$

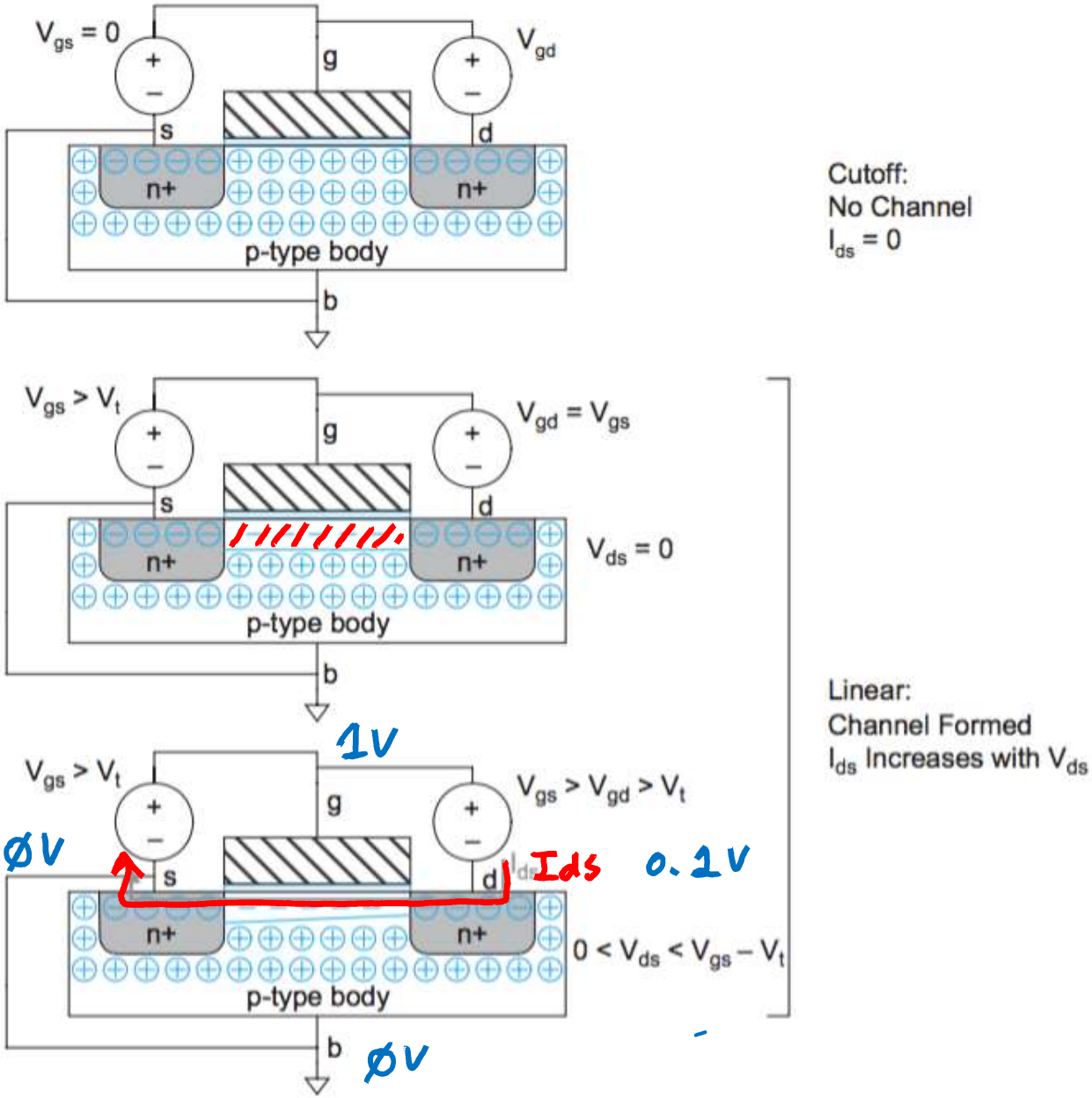
  no channel

  inversion mode

- Linear triode

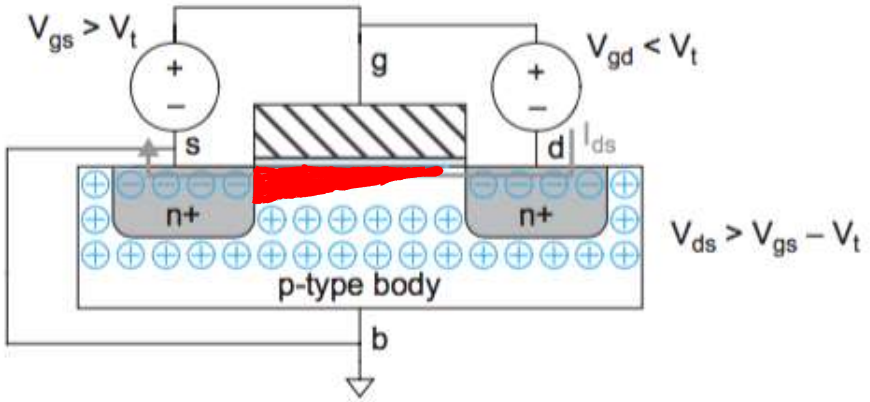
  $V_{gs} > V_{th}$





NMOS Operating Regions (III)

- Saturation



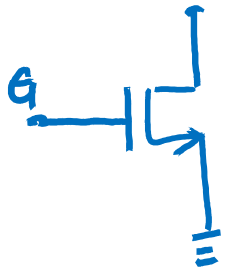
Saturation:
Channel Pinched Off
I_{ds} Independent of V_{ds}

$V_{gs} > V_{th}$

$V_{gd} < V_{th}$

$\rightarrow V_{ds} > (V_{gs} - V_{th})$

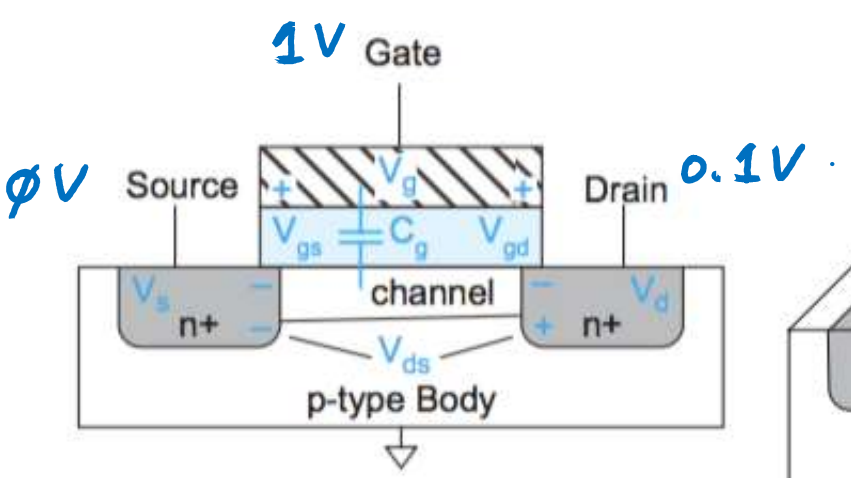
overdrive voltage : V_{ov}, V_{GT}

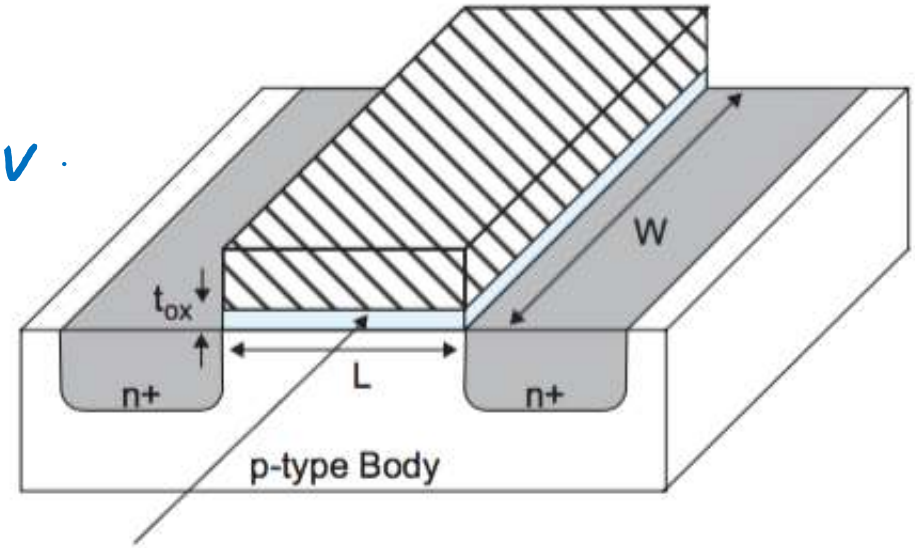


Outline

- Introduction
- **Ideal I-V characteristics**
- Nonideal Effects
- DC transfer characteristics
- C-V characteristics
- Switch-level RC delay models

MOS Channel Charge





Average gate to channel potential:

$$V_{gc} = (V_{gs} + V_{gd})/2 = V_{gs} - V_{ds}/2$$

$I_{ds} = f(Q_{ch}, \mu \cdot V_{ds})$

$(V_{gs} - V_{th}) \cdot W$

$= \frac{Q_{ch}}{L} \cdot v$ (where $v = \mu \cdot E = \mu \frac{V_{ds}}{L}$)

$Q_{ch} = C_{ch} \cdot (V_{gs} - \frac{1}{2} V_{ds} - V_{th})$

$C_{ch} = \varepsilon_{ox} \cdot \frac{W \cdot L}{t_{ox}} = \frac{\varepsilon_{ox}}{t_{ox}} \cdot WL = C_{ox} \cdot WL$

↑ cap/area

$fF/(\mu m)^2$

Ideal I-V Equations

$$I_{ds} = \frac{C_{ox} \cdot W \cdot L \left(V_{gs} - \frac{1}{2} V_{ds} - V_{th}\right)}{L} \mu \cdot \frac{V_{ds}}{L}$$

$$= \mu C_{ox} \frac{W}{L} \left(V_{gs} - V_{th} - \frac{1}{2} V_{ds}\right) \cdot V_{ds}$$ in linear region (triode)

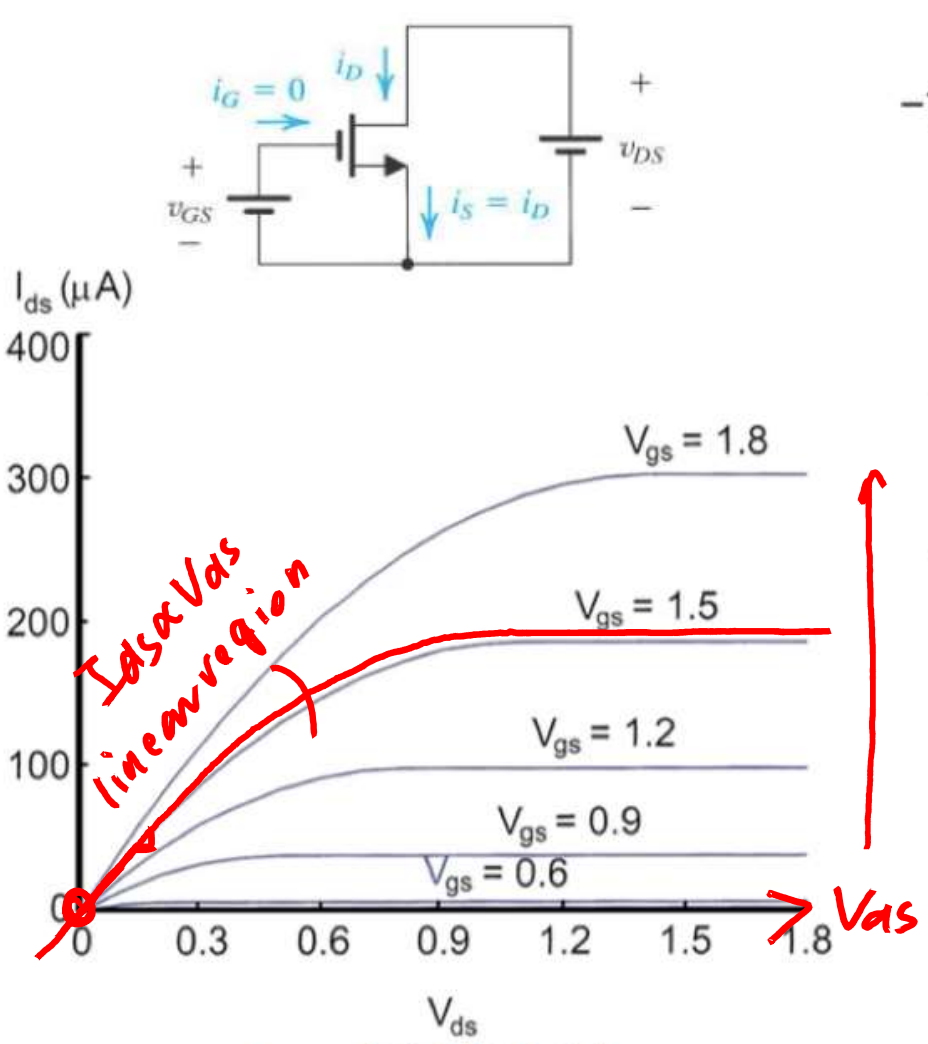
要進入 saturation

$$V_{ds} = V_{gs} - V_{th}$$

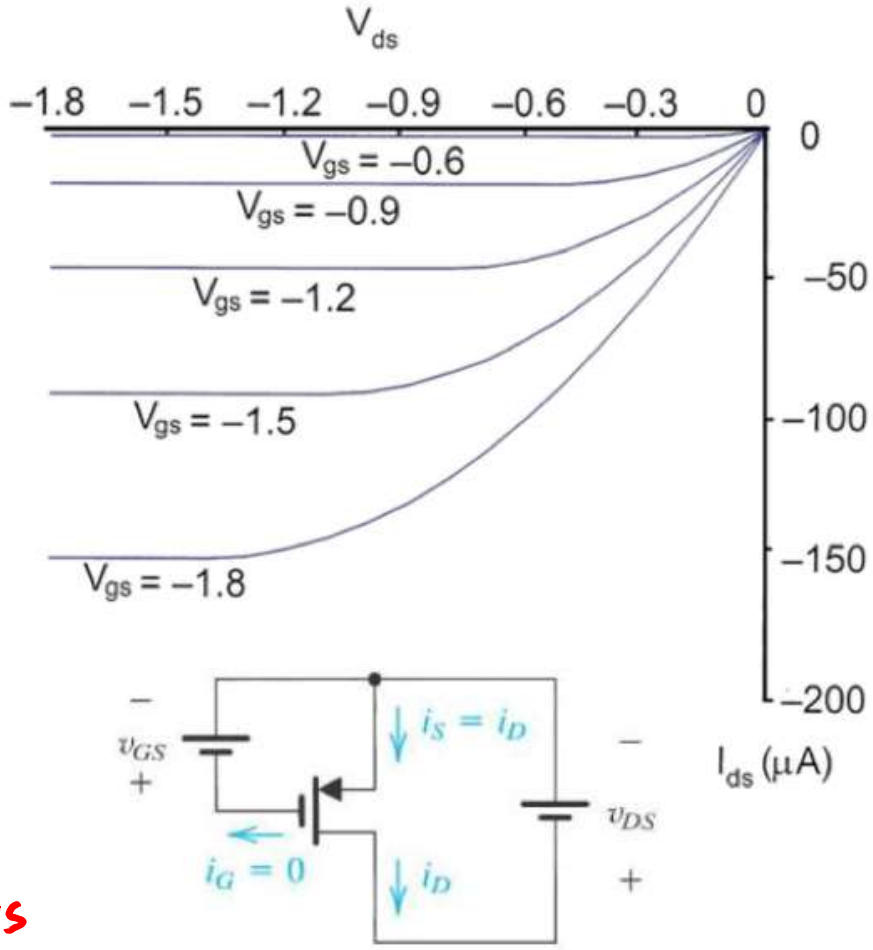
$$I_{ds,sat} = \frac{1}{2} \mu C_{ox} \frac{W}{L} \left(V_{gs} - V_{th}\right)^2$$ super-linear relation

β

Ideal I-V Characteristics



- nMOS I-V



- pMOS I-V

Outline

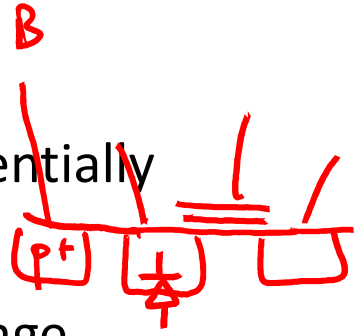
- Introduction
- Ideal I-V characteristics
- **Nonideal Effects**
- DC transfer characteristics
- C-V characteristics
- Switch-level RC delay models

Nonideal I-V Effects

- **Channel length modulation** at high V_{ds} ↑, depletion of S/D ↑, effective L ↓, I_{ds} increase ↑.
- **Body effect** threshold voltage Vth influenced by Vbs (body-to-source voltage).

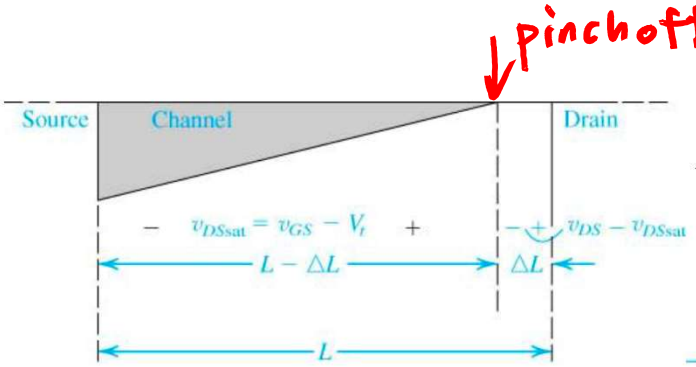
- **Velocity saturation** at high V_{ds} ↑, the carrier velocity is no longer proportional to lateral field. I_{ds} decrease ↓.
- **Mobility degradation** at high V_{gs} ↑, the carrier scatter more and mobility decreases. I_{ds} decrease ↓.

- **Subthreshold conduction** Vgs < Vth, Ids is exponentially dropoff instead of abruptly becoming zero
- **Drain/source leakage** reverse diode junction leakage
- **Non-zero gate current Ig** carriers tunneling effect

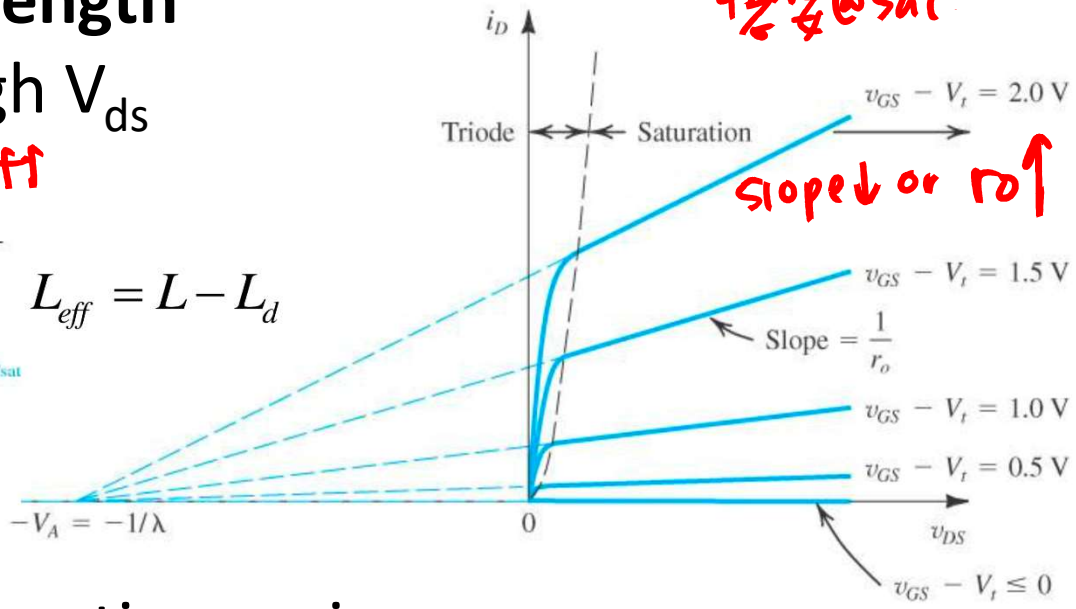


Channel Length Modulation

- **Effective channel length** reduced due to high V_{ds}



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



- I-V equation at saturation region:
  - λ' empirical parameter

$$I_{ds} = \beta \frac{(V_{gs} - V_t)^2}{2}(1 + \lambda V_{ds}), \lambda = \lambda' / L$$

- With shorter L ↓, λ ↑, resulting in output resistance ↓, MOSFET intrinsic gain ↓

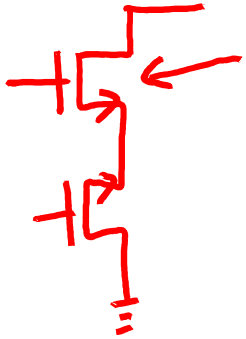
Body Effect

- Threshold voltage Vth increased with positive V_{sb}
- $V_{sb} < 0$ → Vth ↓, OFF leakage ↑ (design trade-off)

$$V_t = V_{t0} + \gamma[\sqrt{2\varphi_s + V_{sb}} - \sqrt{2\varphi_s}],$$

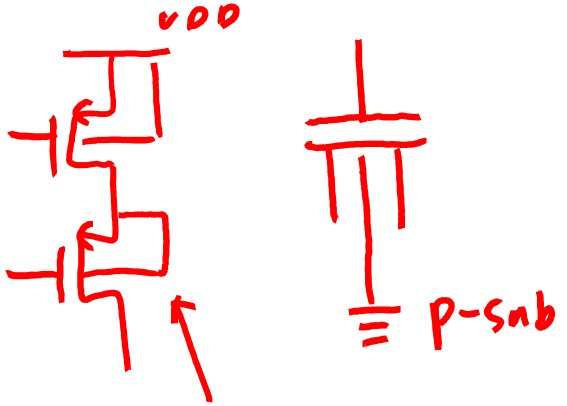
$$\varphi_s = 2\upsilon_T \ln\frac{N_A}{n_i}$$

V_{t0}: threshold voltage for $V_{sb} = 0$
ϕ_f: process-dependent parameter



suffered from body effect

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



PMOS 的 N well 可以接到 source

γ: body-effect coefficient (process-depedent)
N_A: doping concentration of p-substrate
ε_s: permittivity of silicon = 11.7 ε_0

Velocity Saturation

- **Carrier velocity** nonlinearly proportional to lateral electrical field before velocity saturation

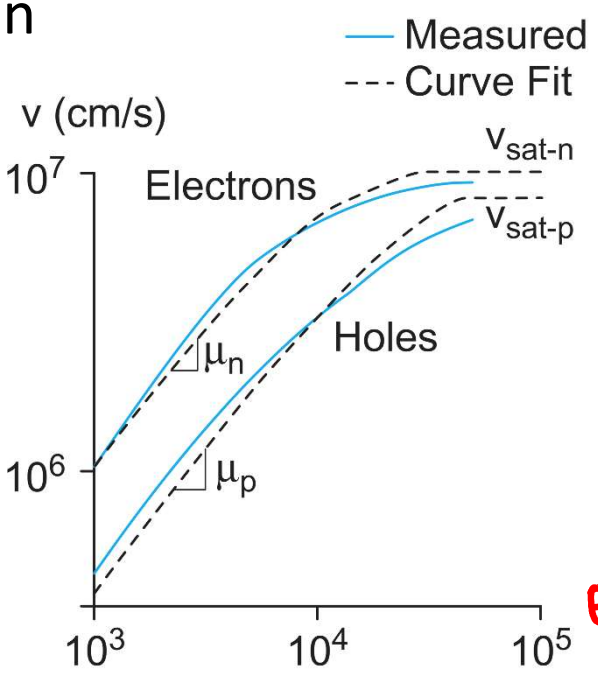
$$\upsilon = \mu E_{lat} / \left(1 + E_{lat} / E_{sat}\right)$$

υ : carrier velocity

μ : mobility

$E_{lat} = V_{ds} / L$: lateral electrical field

$E_{sat} = \upsilon_{sat} / \mu$



- I_{ds} will saturate to velocity saturation, depending on channel length **L** and applied $\mathbf{V_{ds}}$

$$I_{ds} = \frac{Q_{channel}}{t_{channel}} = \frac{Q_{channel}}{L / \upsilon_{sat}} = C_{ox} W \left(V_{gs} - V_t\right) \upsilon_{sat}$$

α-power Law Model

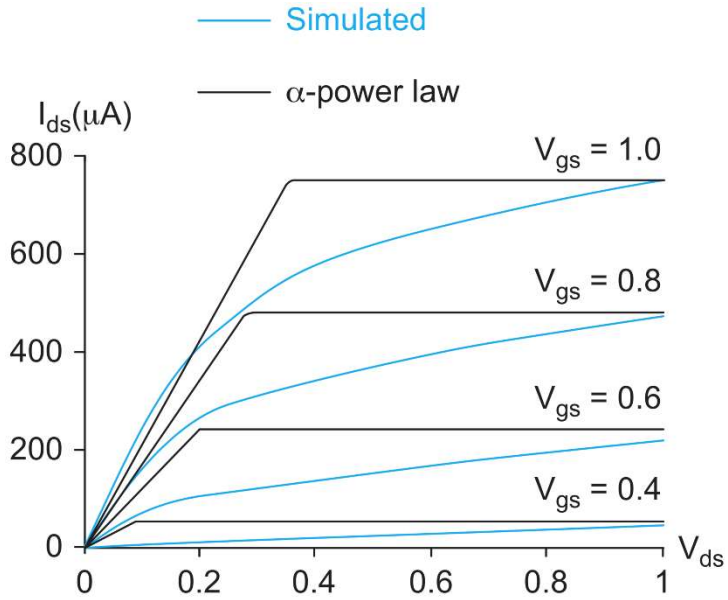
- Piecewise linear model to illustrate MOSFET's I-V characteristic with velocity saturation

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

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

Empirical parameters: P_c, P_v, α

→ gm degradation



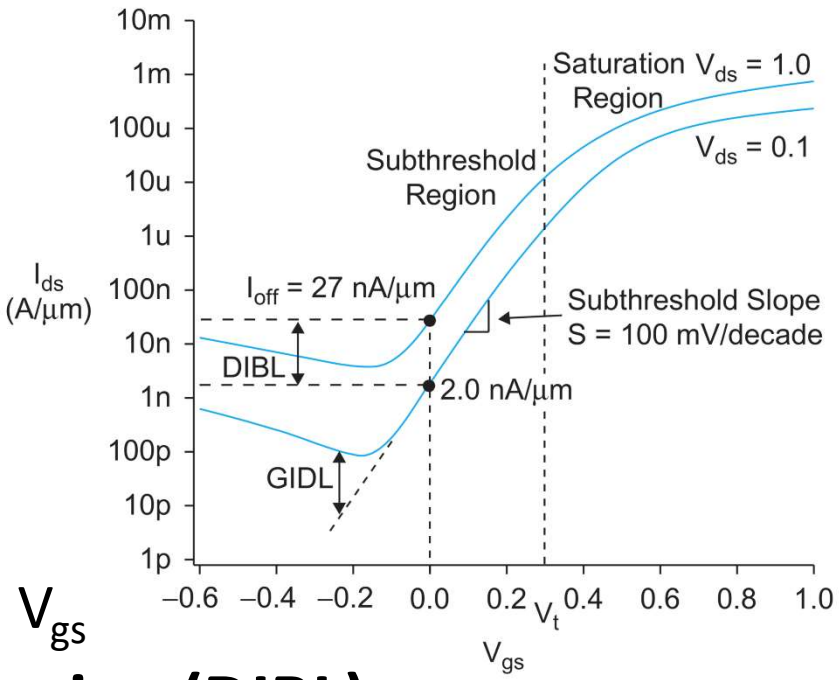
- **Because $\mu_p < \mu_n$**, PMOS experiences less velocity saturation than NMOS → $\alpha_p > \alpha_n$
- **Mobility degradation** is modeled by a $\mu_{eff} < \mu$, and it can be included in to the parameter α

Subthreshold Conduction

- Leakage current at subthreshold region
  - $V_{gs} < V_{th}$: weak inversion

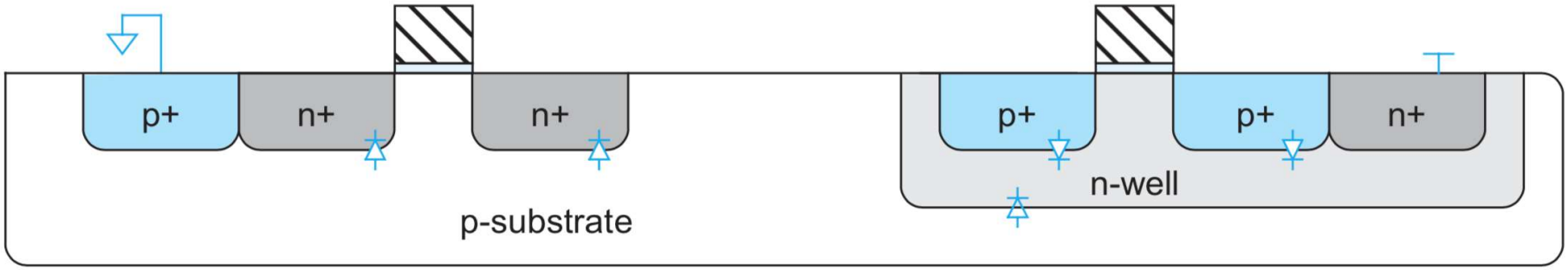
$$I_{ds} = I_{ds0} e^{\frac{V_{gs}-V_t}{n\upsilon_T}} \left(1 - e^{\frac{-V_{ds}}{\upsilon_T}}\right)$$

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



- $I_d = 0$ when $V_{ds} = 0$
  - Increase exponentially with V_{gs}
- **Drain-induced barrier lowering (DIBL)**
  - V_{th} will reduce with positive V_{ds}
  - Worsens leakage at subthreshold
  - Like channel-length modulation at active mode

Junction Leakage

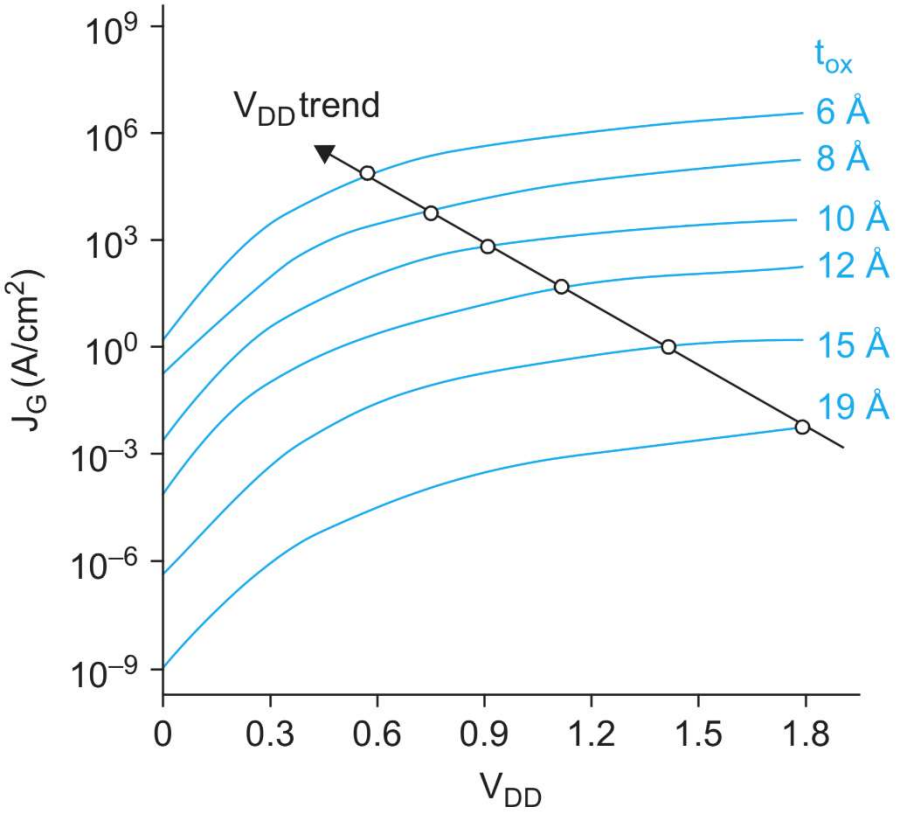


- S/D junction leakage from a reverse-biased diode

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

- Junction leakage used to be the limitation for storage time. In modern processes, subthreshold leakage becomes dominant

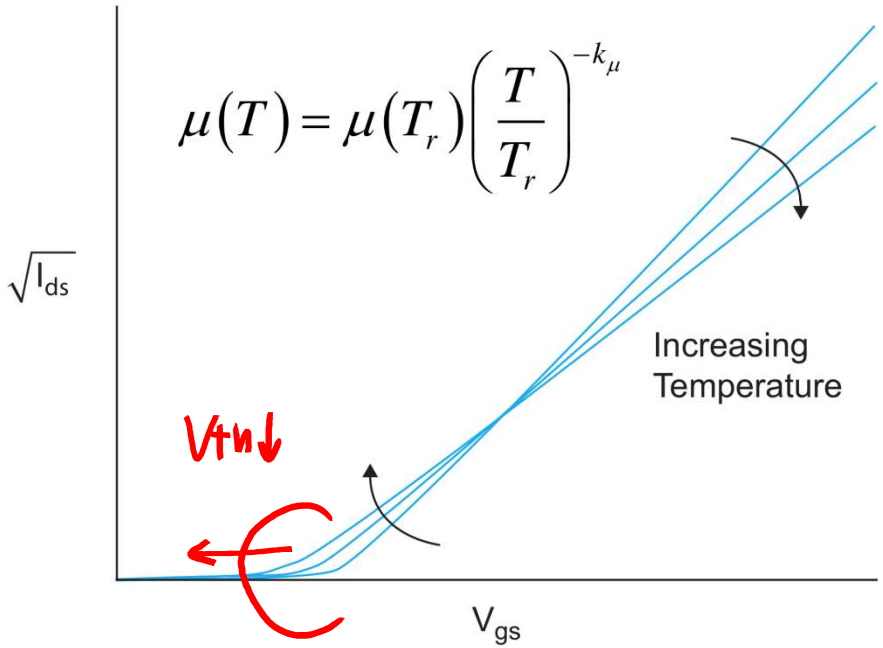
Tunneling Effect



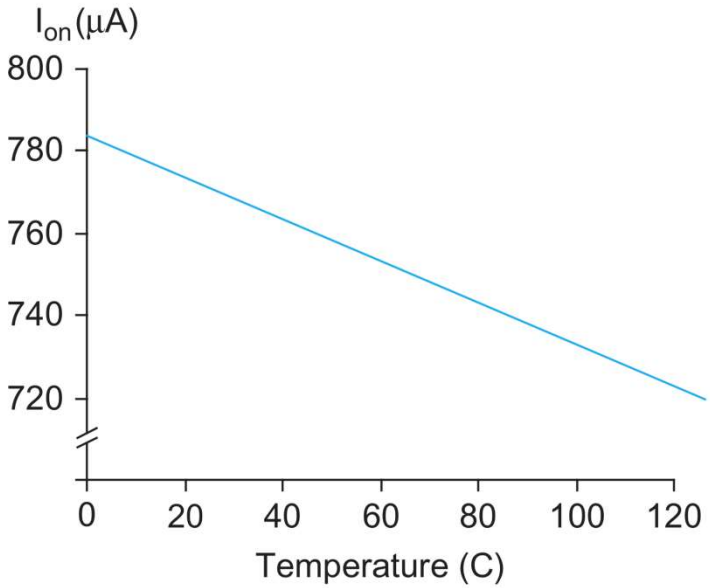
- Gate leakage: from carriers' tunneling through gate oxide. Exponentially inversely proportional to gate oxide thickness.
- High-*k* (dielectric constant) gate insulator used

Temperature Dependence

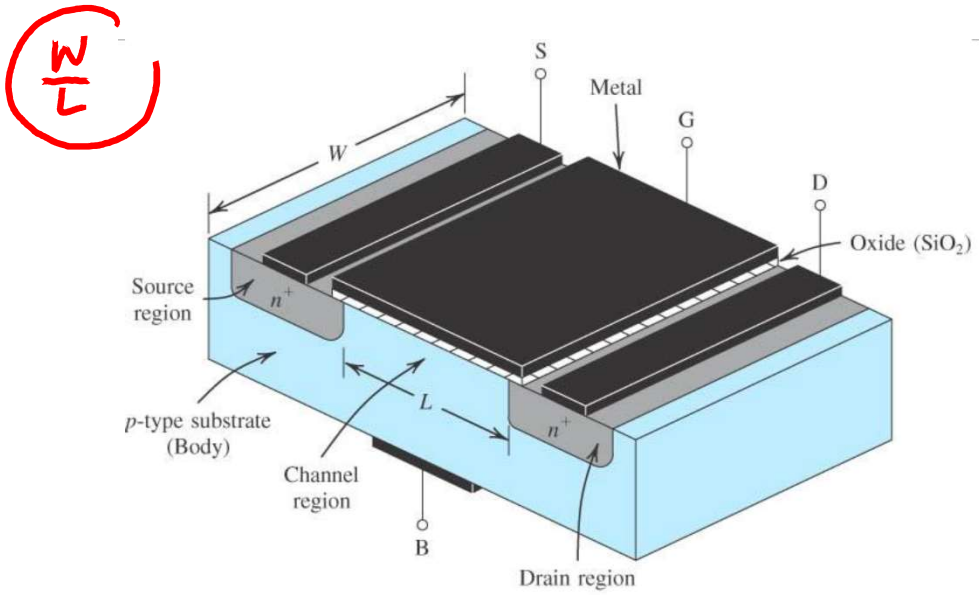
- T ↑ IOFF ↑ ION
- Circuit performance improve with T ↓: subthreshold leakage ↓, saturation velocity ↑, mobility ↑, junction capacitance ↓, but breakdown voltage ↓.

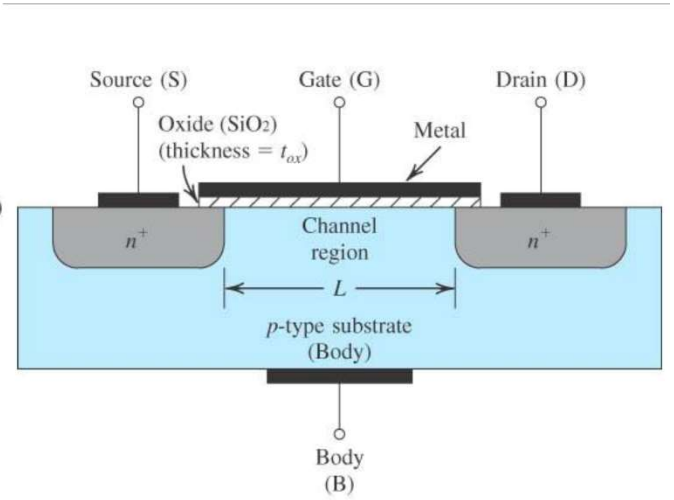


$$V_t(T) = V_t(T_r) - k_{vt}(T - T_r)$$



Geometry Dependence





- Effective channel length and width

$$L_{eff} = L_{drawn} + X_L - 2L_D \qquad X_L, X_W : \text{Poly over-etch}$$

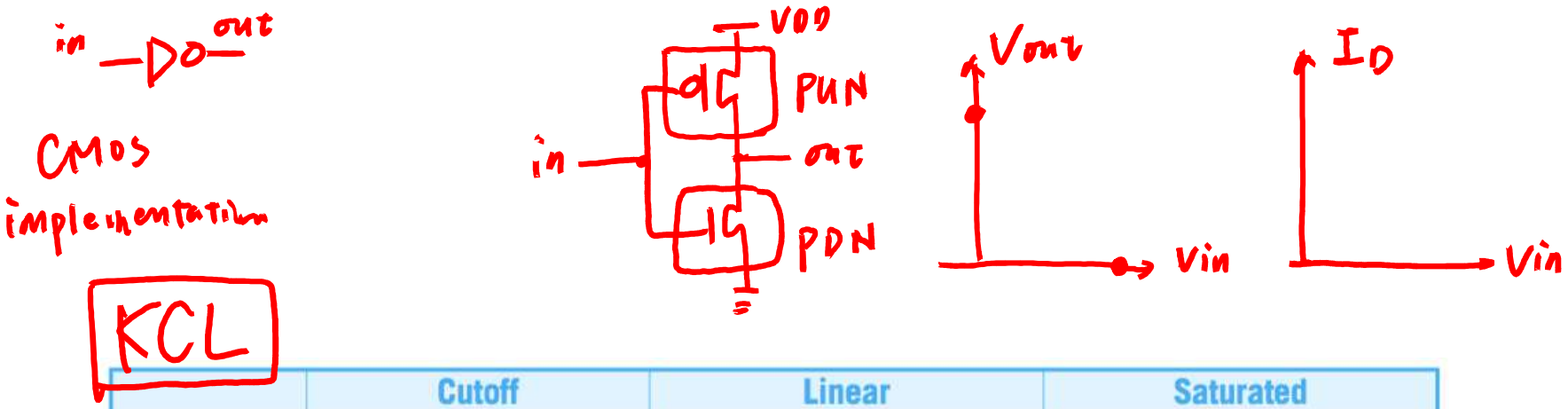
$$W_{eff} = W_{drawn} + X_W - 2W_D \qquad L_D, W_D : \text{Source-drian lateral diffusion}$$

- Use identical and same orientation for MOSFETs for good matching – **EX:** differential pair, current mirror

Outline

- Introduction
- Ideal I-V characteristics
- Nonideal Effects
- **DC transfer characteristics**
- C-V characteristics
- Switch-level RC delay models

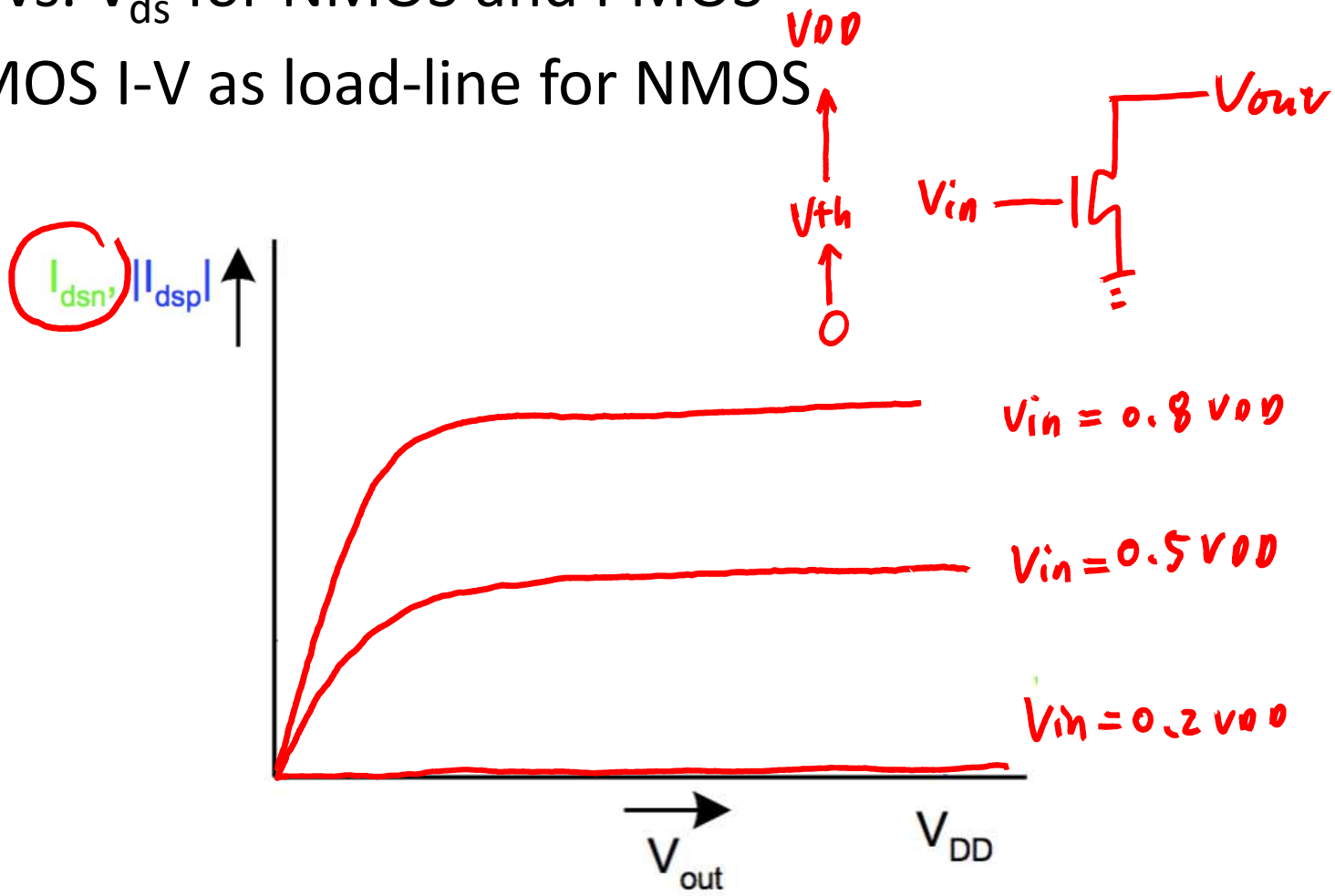
CMOS Inverter DC Characteristics (I)



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

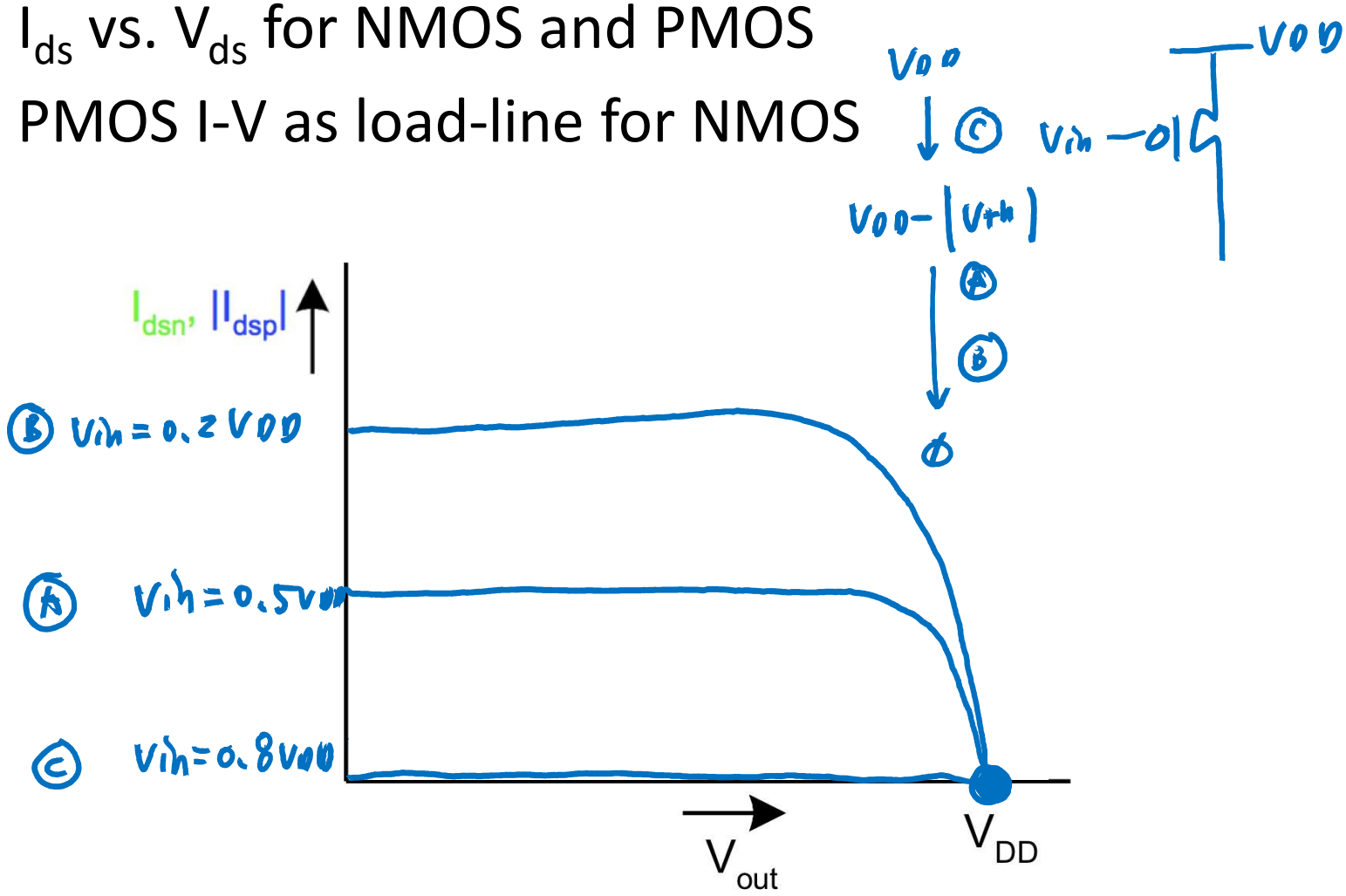
CMOS Inverter DC Characteristics (II)

- I_{ds} vs. V_{ds} for NMOS and PMOS
- PMOS I-V as load-line for NMOS



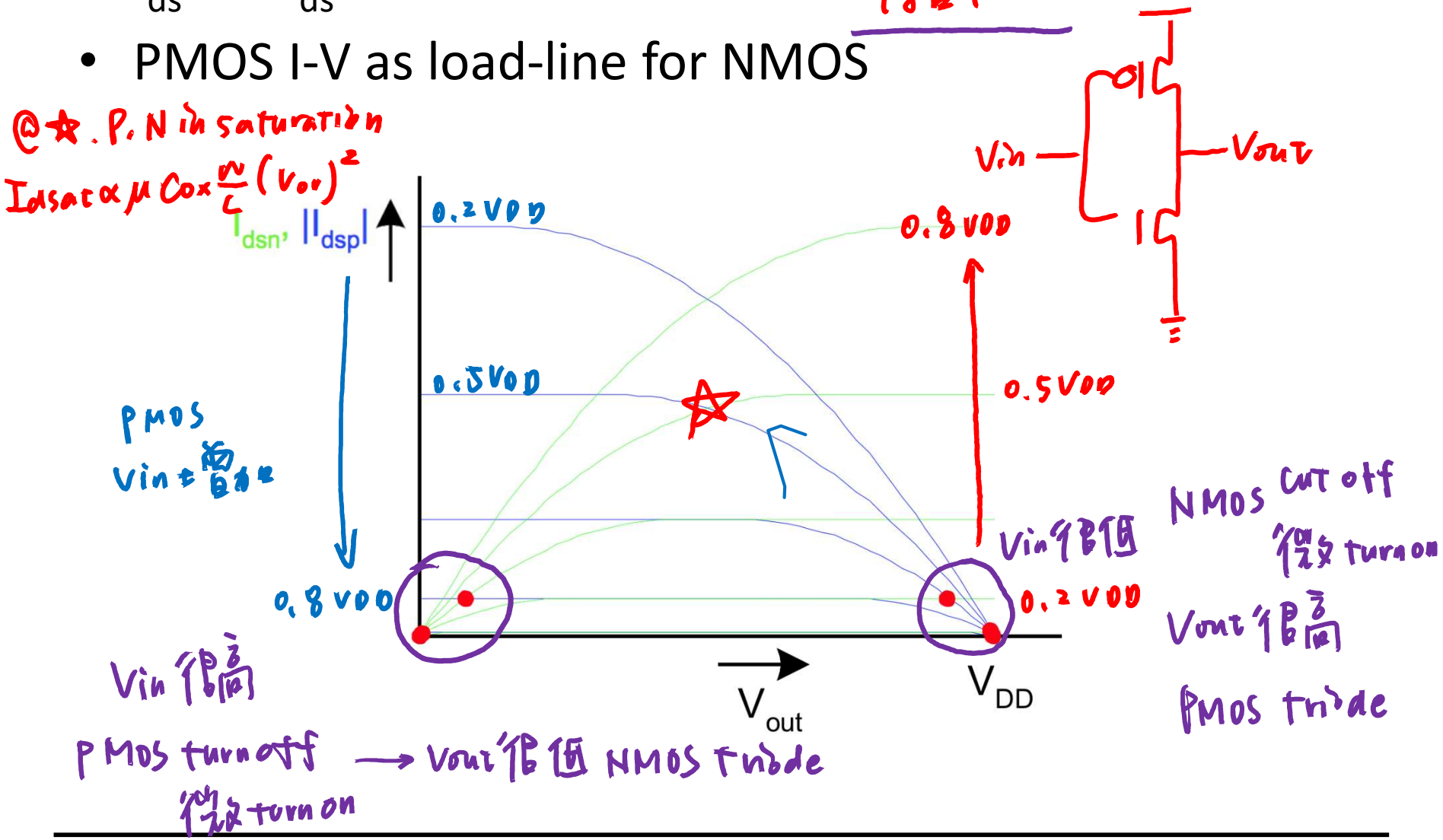
CMOS Inverter DC Characteristics (III)

- I_{ds} vs. V_{ds} for NMOS and PMOS
- PMOS I-V as load-line for NMOS



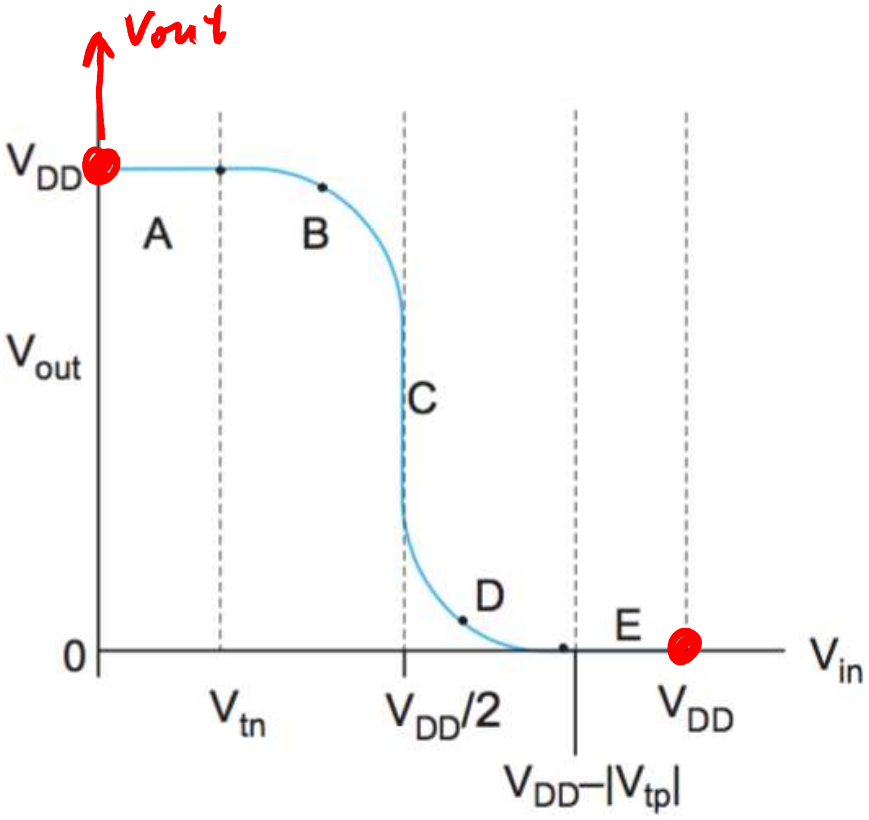
CMOS Inverter DC Characteristics (IV)

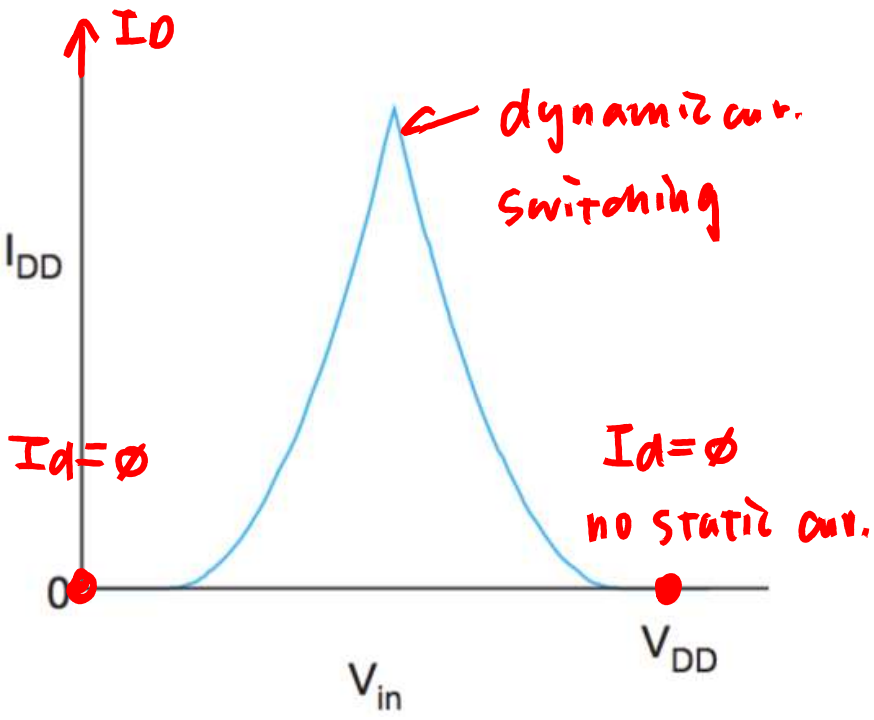
- I_{ds} vs. V_{ds} for NMOS and PMOS
- PMOS I-V as load-line for NMOS



CMOS Inverter DC Characteristics (V)

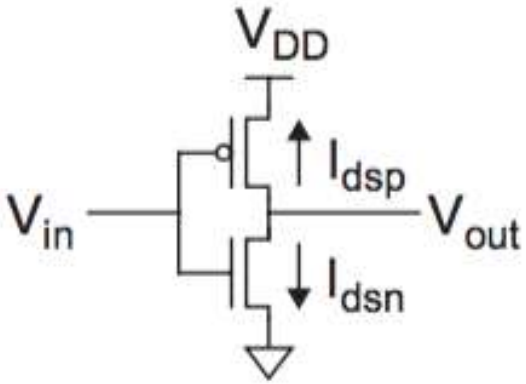
- V_{in}-V_{out} DC transfer curve
- Rail-to-rail operation
- V_{in}-I_{DD} DC transfer curve
- Dynamic power dissipation





DC Response

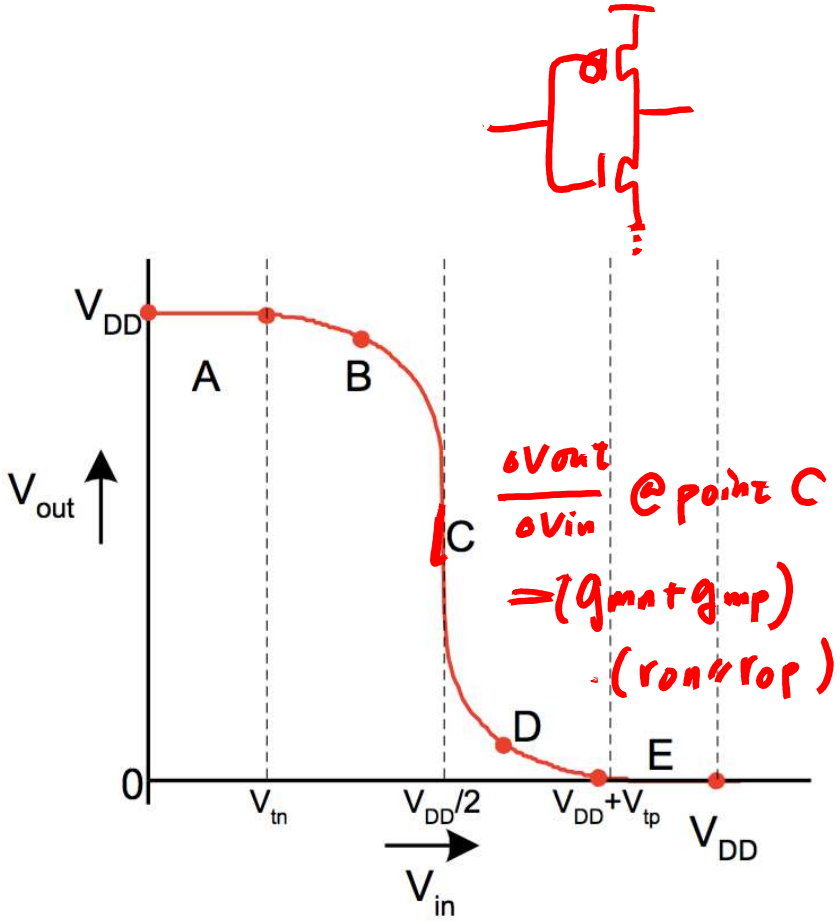
- DC response: V_{out} vs. V_{in} for a gate
- **Ex: inverter**
  - $V_{in} = 0 \rightarrow V_{out} = V_{DD}$
  - $V_{in} = V_{DD} \rightarrow V_{out} = 0$
  - In between, V_{out} depends on transistor sizes and currents
  - By KCL must settle such that $I_{dsn} = |I_{dsp}|$
  - We could solve equations
  - Graphical solution gives more insight



Transistor Operating Regions

DC characteristics

	NMOS	PMOS
A	cut off	linear
B	sat ∥ ON	linear
C	sat	sat
D	linear	sat ∥ ON
E	linear	cut off



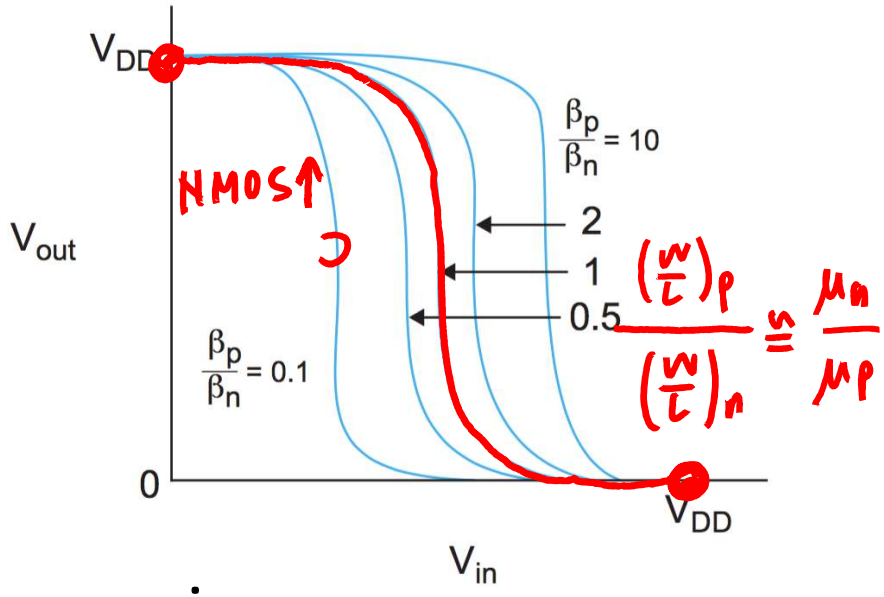
Beta Ratio Effect

- β ratio and parameters

$$\beta_p = \mu_p C_{ox} (W/L)_p$$

$$\beta_n = \mu_n C_{ox} (W/L)_n$$

$$\beta \text{ ratio} = \beta_p / \beta_n$$



- β ratio = 1 → largest noise margin
  - $\mu_n > \mu_p$, choose $(W/L)_p > (W/L)_n$ to make β ratio = 1
- β ratio > 1 → HI-skewed inverter, switching threshold > 0.5 V_{DD}
- β ratio < 1 → LO-skewed inverter, switching threshold < 0.5 V_{DD}

Noise Margin (I)

- The allowable noise voltage on the input that the output won't be corrupted

$$NM_L = V_{IL} - V_{OL}$$

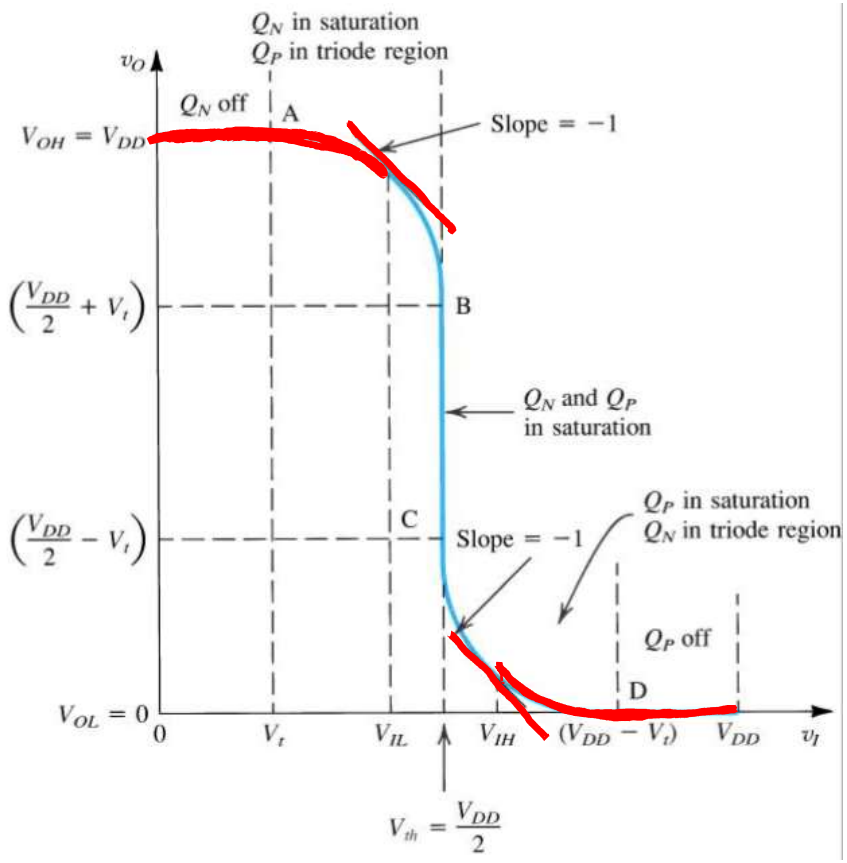
$$NM_H = V_{OH} - V_{IH}$$

V_{IH} = minimum HIGH input voltage

V_{IL} = maximum LOW input voltage

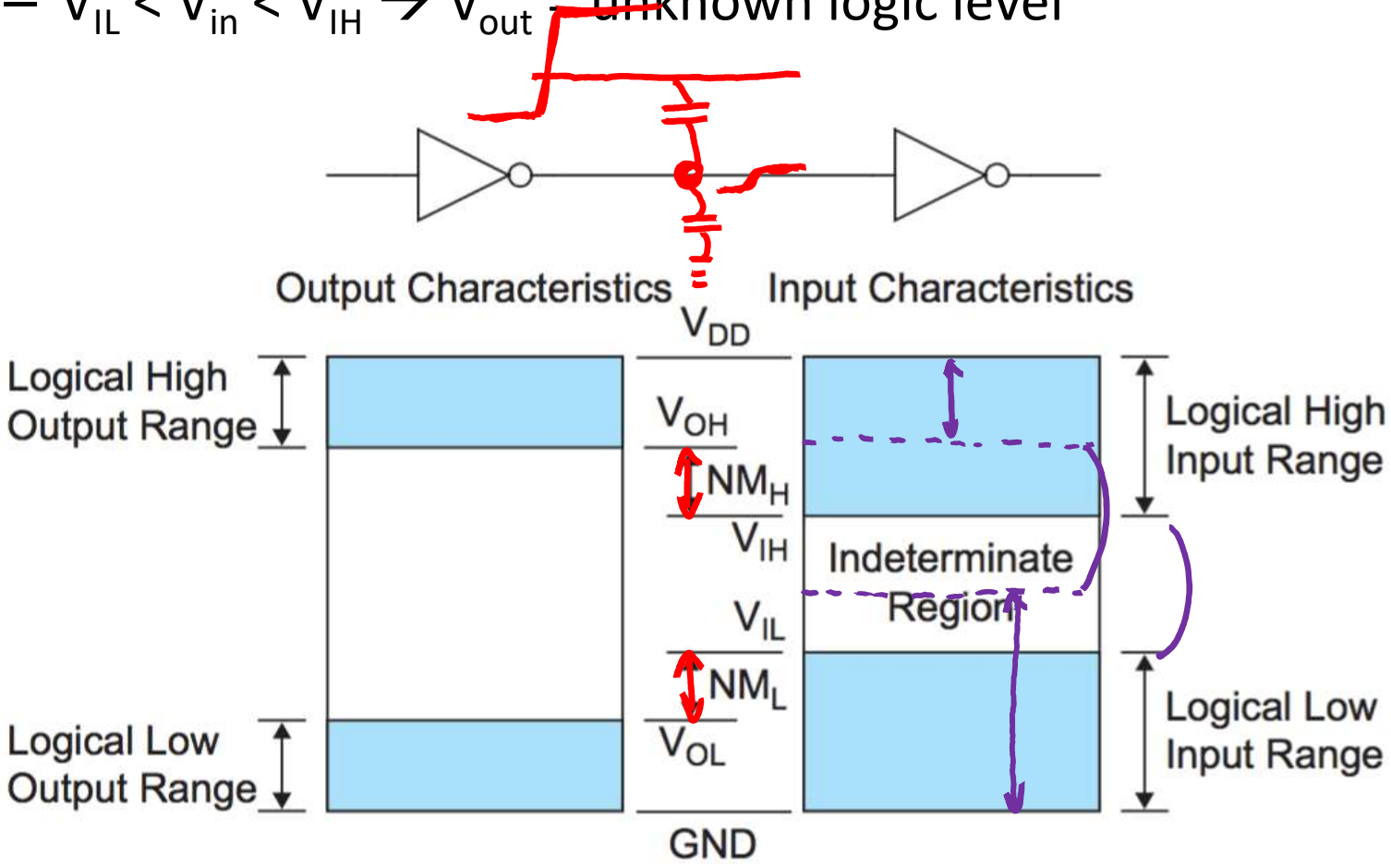
V_{OH} = minimum HIGH output voltage

V_{OL} = maximum LOW output voltage



Noise Margin (II)

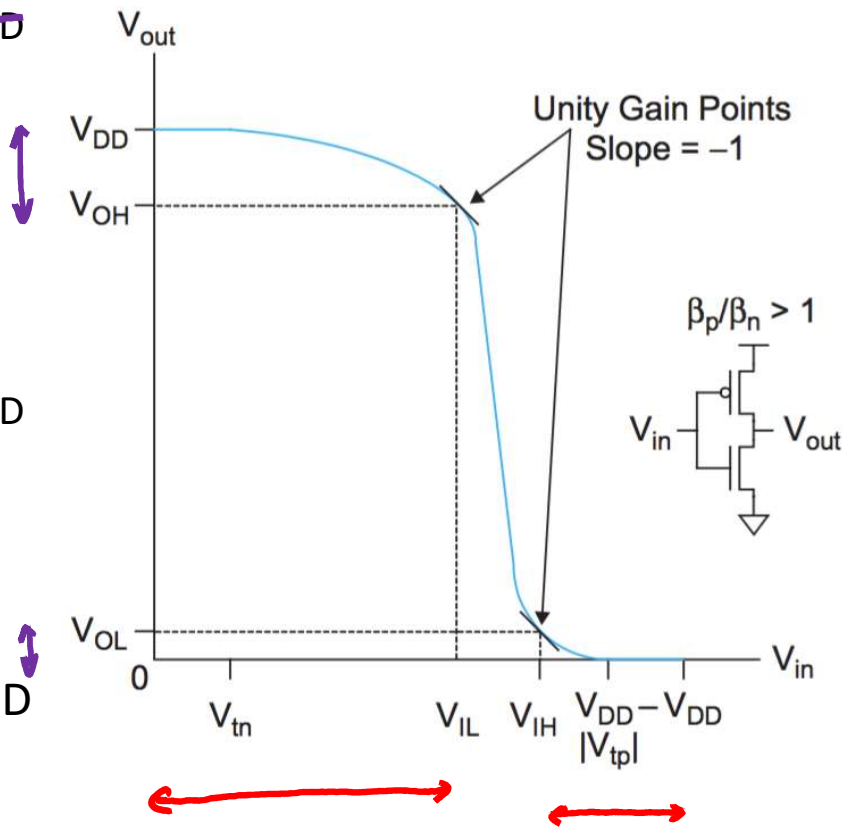
- Indeterminate region (*forbidden zone*)
  - $V_{IL} < V_{in} < V_{IH} \rightarrow V_{out}$ = unknown logic level



Beta Ratio and Noise Margin

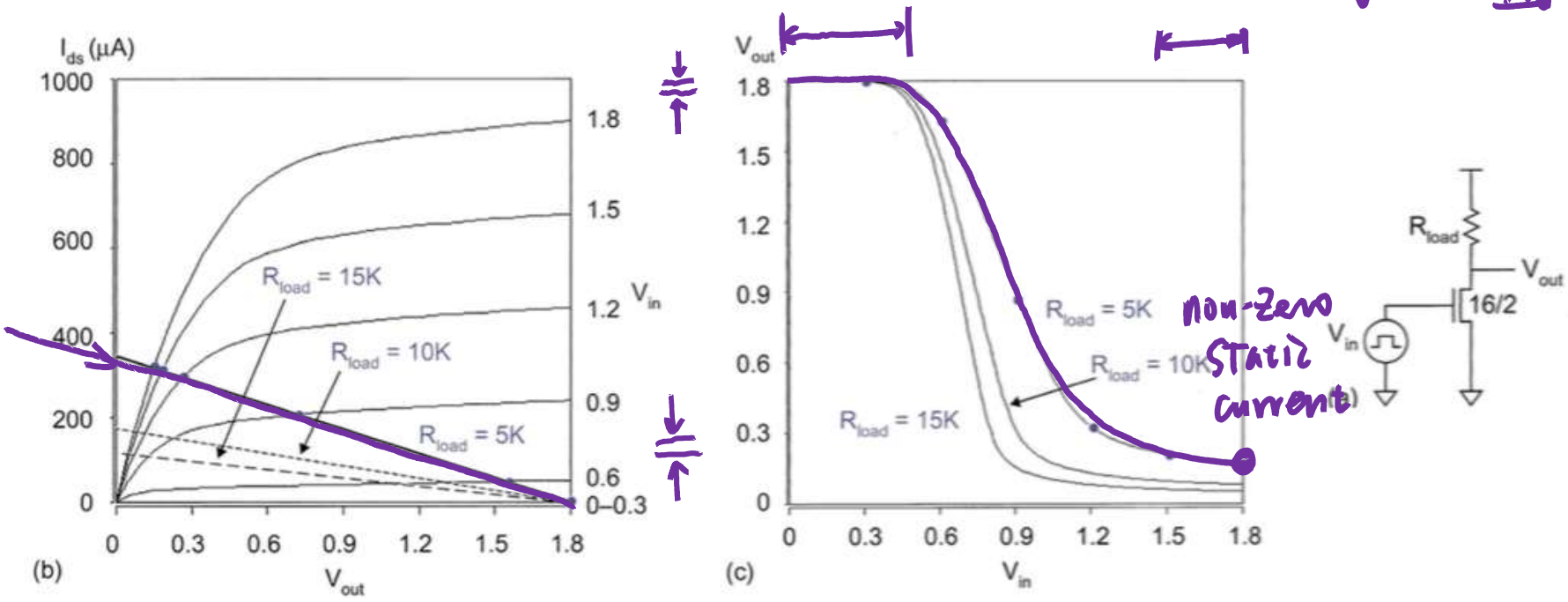
- β ratio > 1
  - Switching threshold > 0.5 V_{DD}
  - VIH ↑ NMH ↓
  - VIL ↑ NML ↑
- β ratio < 1
  - Switching threshold < 0.5 V_{DD}
  - VIH ↓ NMH ↑
  - VIL ↓ NML ↓
- Noise tend to scale with V_{DD}
  - As VDD ↓
    smaller NM is acceptable

curve往右移



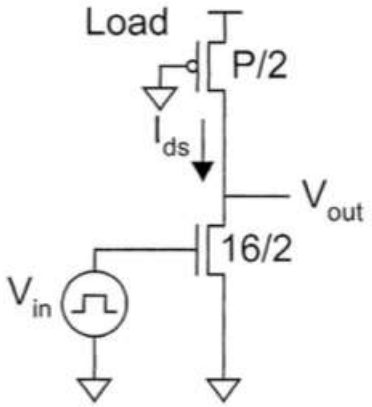
Ratioed Inverter Transfer Function (I)

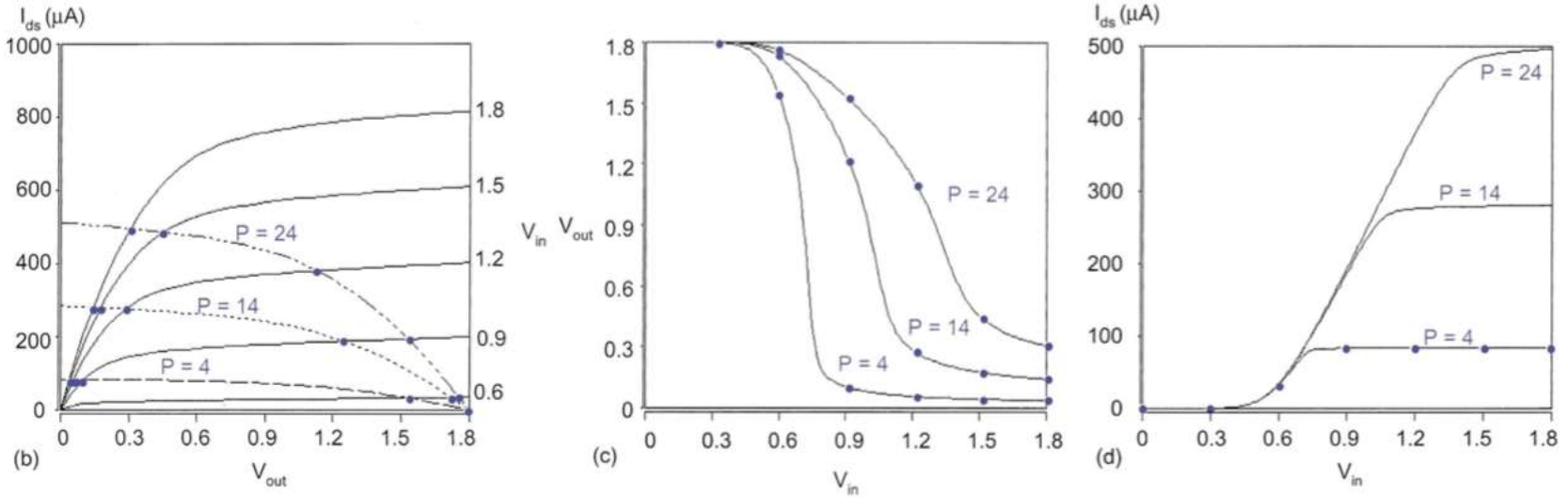
- NMOS inverters with resistive or constant current-source load
  - Transfer function depends on the ratio of pull-down to the pull-up transistor (static load)



Ratioed Inverter Transfer Function (II)

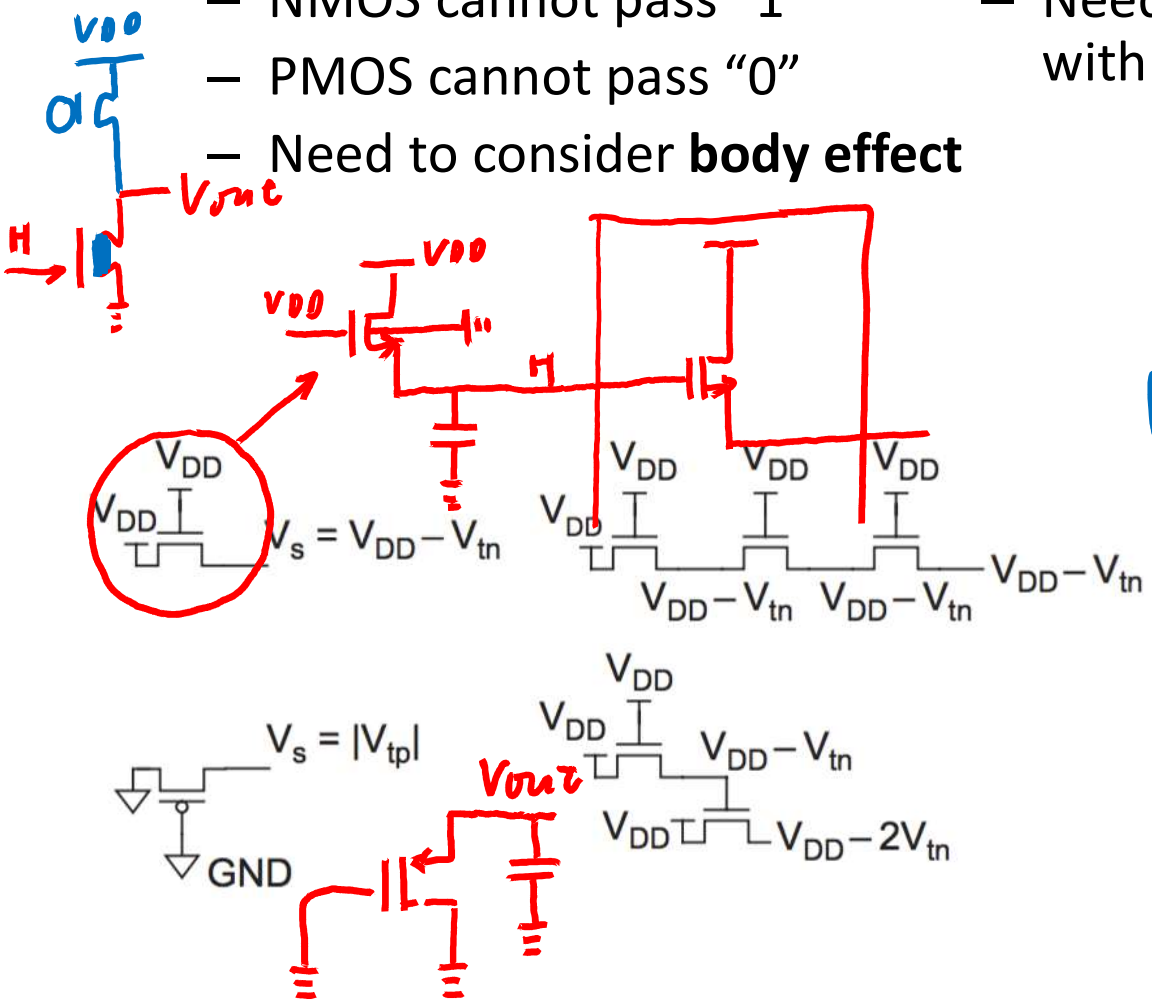
- NMOS inverters with turn-ON PMOS as load
  - Turn-ON PMOS is made by a depletion mode NMOS in pure NMOS process
  - Dissipating static power when V_{out} = LOW
  - Poor NM but smaller area and input capacitance loading

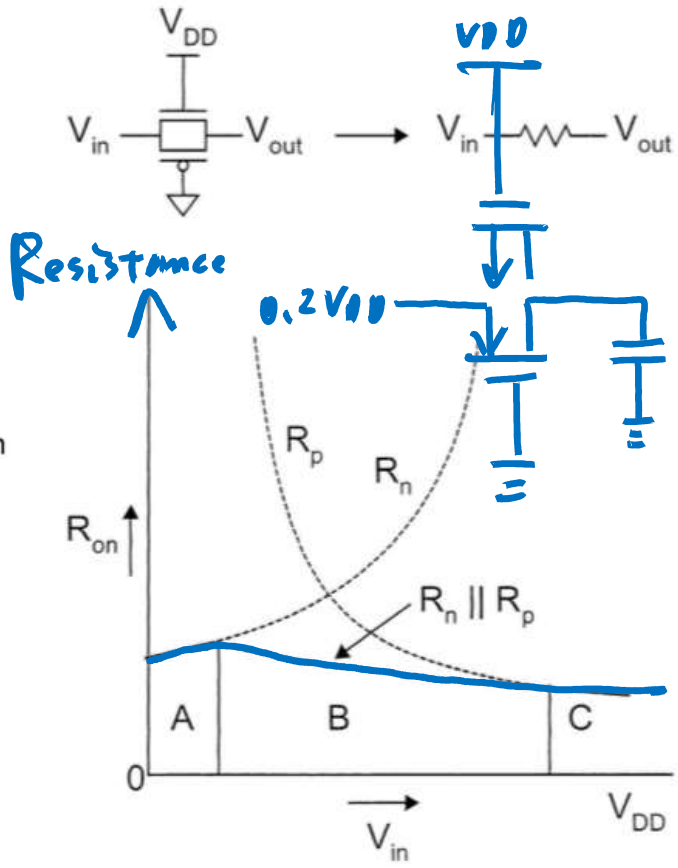




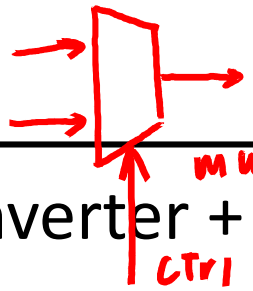
Pass Transistor DC Characteristics

- Due to V_{th}
  - NMOS cannot pass "1"
  - PMOS cannot pass "0"
  - Need to consider **body effect**
- ON-resistance depends on V_{in}
  - Need to boost gate voltage with small V_{DD}

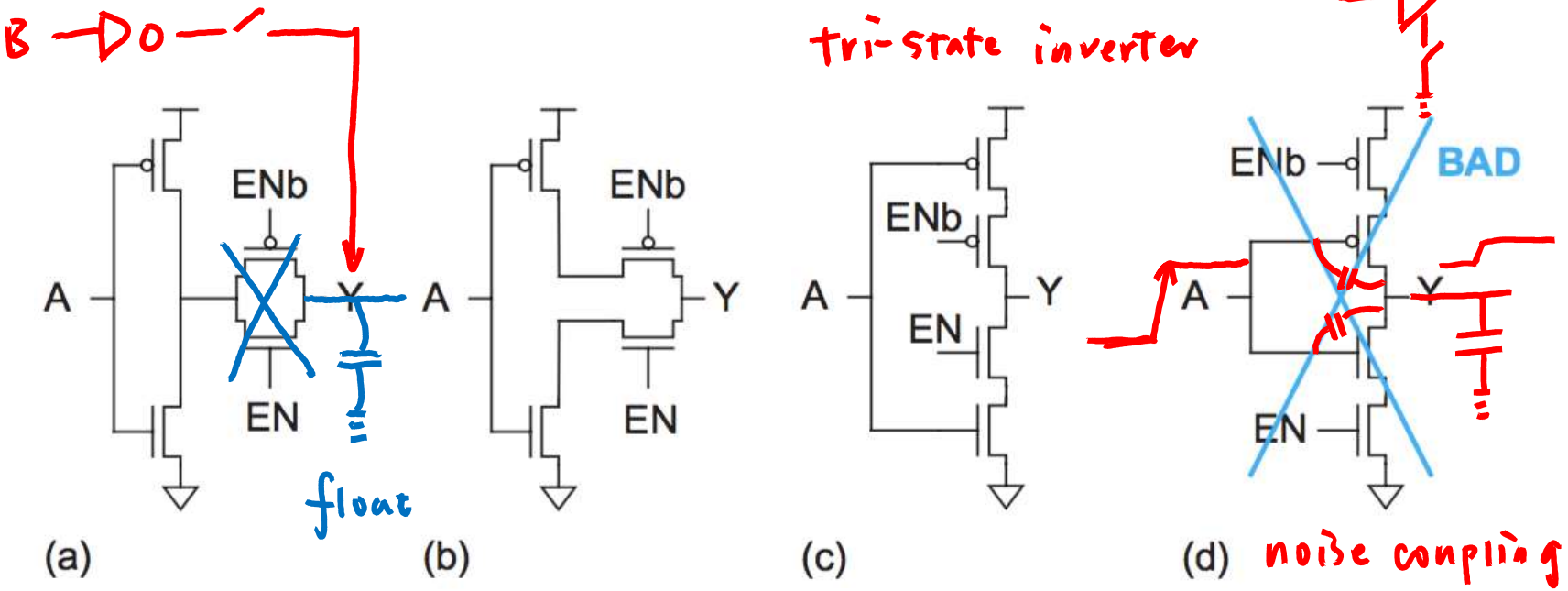




Tri-state Inverter

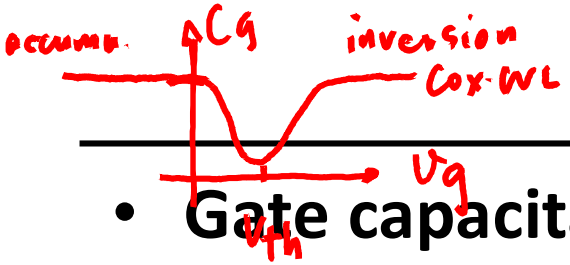


- Inverter + transmission gate
  - Approximately half the speed of CMOS inverter for the same n and p device sizes
  - The structure in (d) suffer from A's toggling in tristate
  - Need to consider **body effect**



Outline

- Introduction
- Ideal I-V characteristics
- Nonideal Effects
- DC transfer characteristics
- **C-V characteristics**
- Switch-level RC delay models

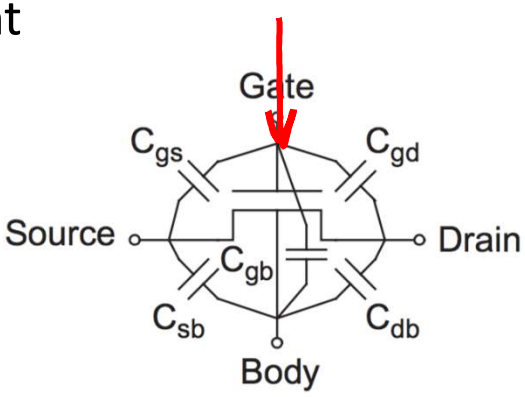


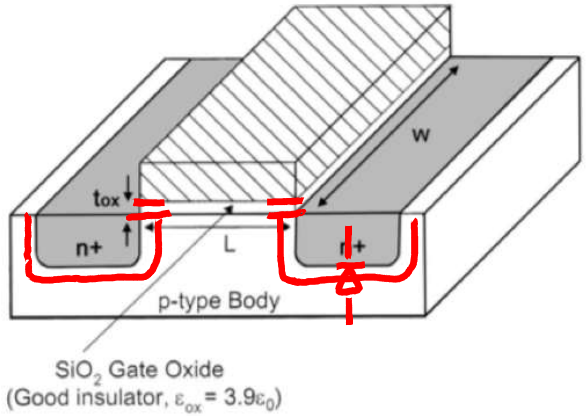
C-V Characteristics

- **Gate capacitance:** with advanced technology, t_{ox} ↓, L ↓, $C_{permicron}$ keeps constant

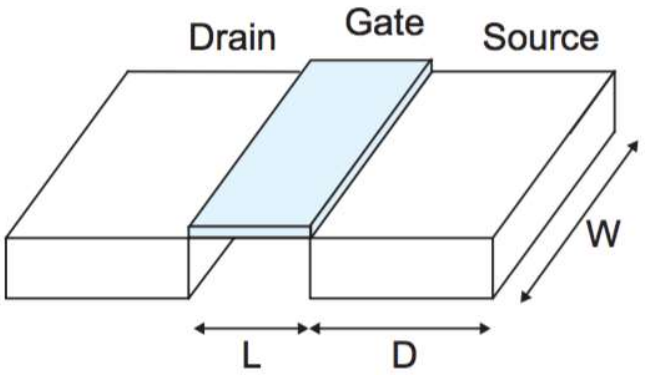
$$C_g = C_{ox}WL = C_{permicron}W$$

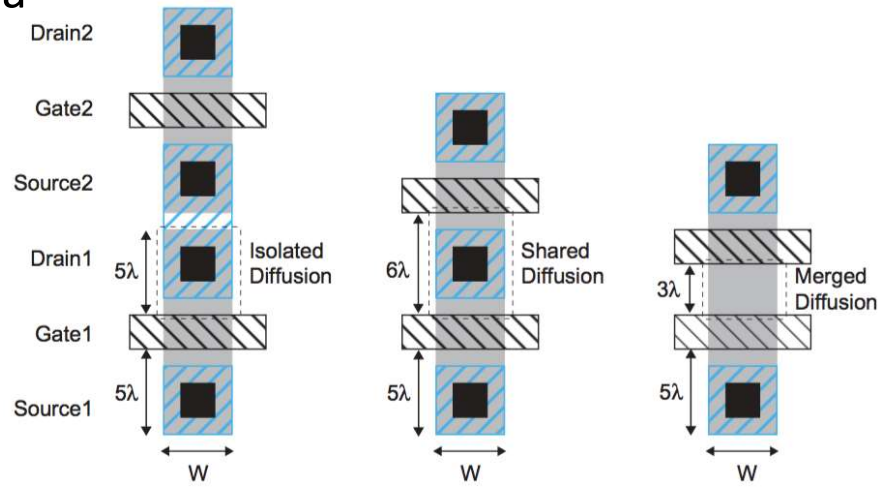
$$C_{permicron} = C_{ox}L = \frac{\varepsilon_{ox}}{t_{ox}}L$$





- **Parasitic capacitance:** C_{db} and C_{sb} from reverse-biased p-n junction and proportional to S/D area





MOS Gate Capacitance Model

- **Gate capacitance:** vary with channel behavior at different operation regions

old tech: 2 fF/μm

advanced tech: 1 fF/μm

$$C_0 = C_{ox}WL$$

width

Parameter	Cutoff	Linear	Saturation
C_{gb}	$\leq C_0$	0	0
C_{gs}	0	$C_0/2$	$2/3\ C_0$
C_{gd}	0	$C_0/2$	0
$C_g = C_{gs} + C_{gd} + C_{gb}$	C_0	C_0	$2/3\ C_0$

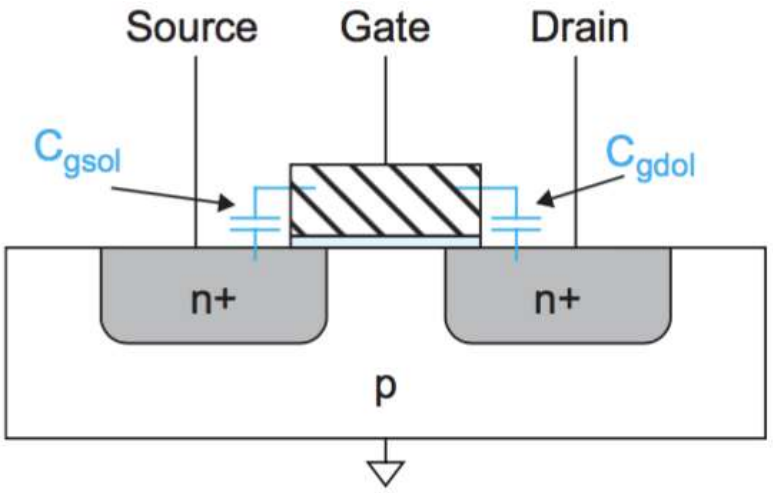
- **S/D overlap capacitance:** Cgs(overlap) and Cgd(overlap) from S/D lateral diffusion

面單位寬度

$$C_{gs(overlap)} = C_{gsol}W$$

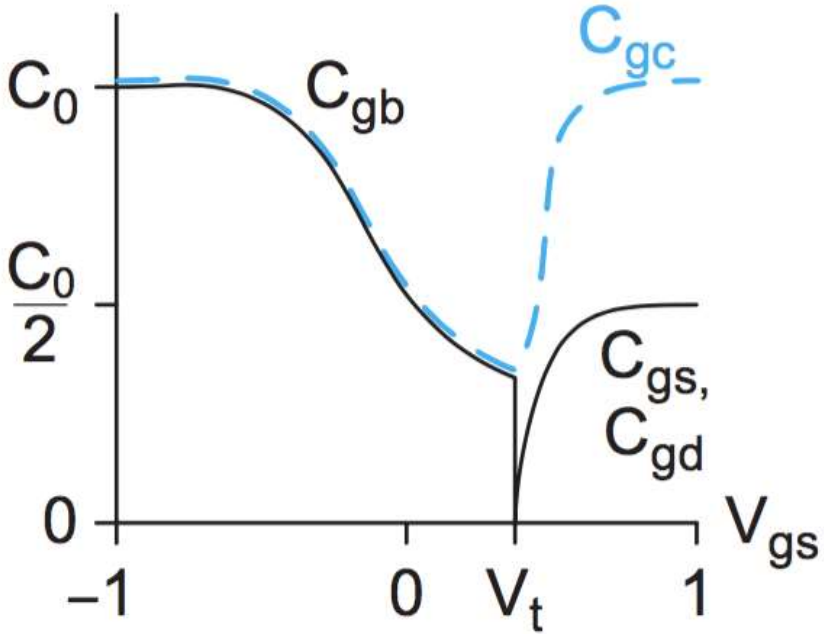
$$C_{gd(overlap)} = C_{gdol}W$$

1 fF/μm

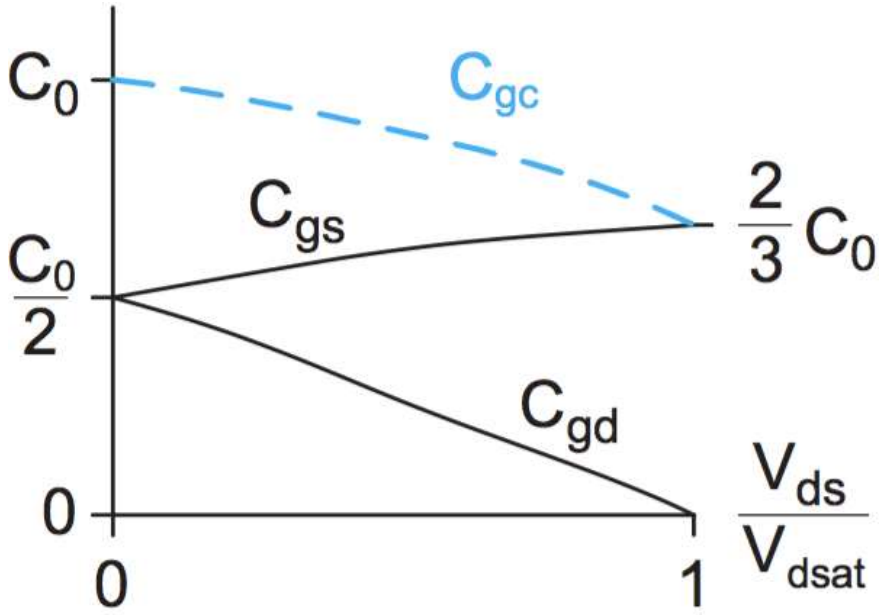


C_{gc} vs. V_{gs} and V_{ds}

- C_{gc} vs. V_{gs}

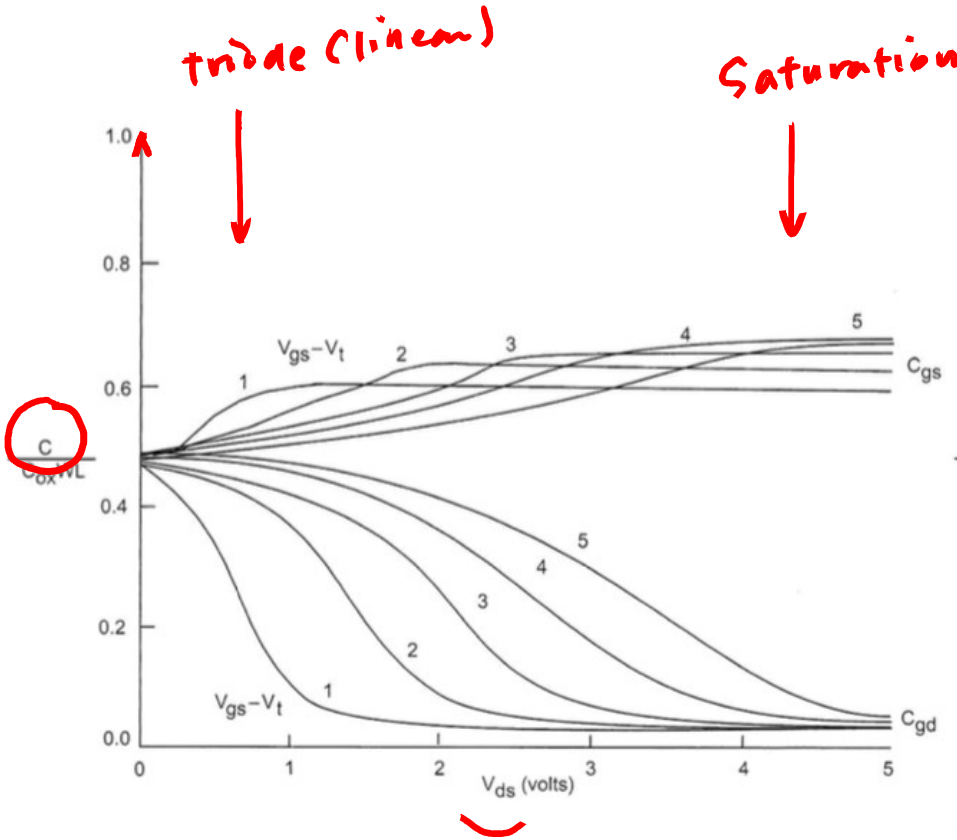


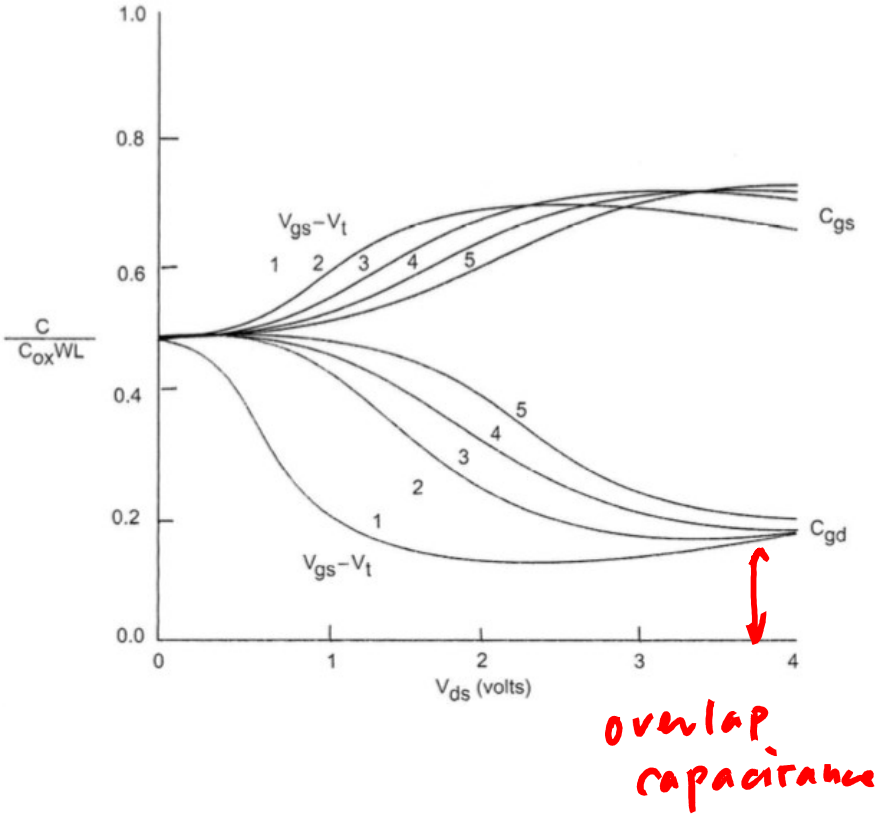
- C_{gc} vs. V_{ds}



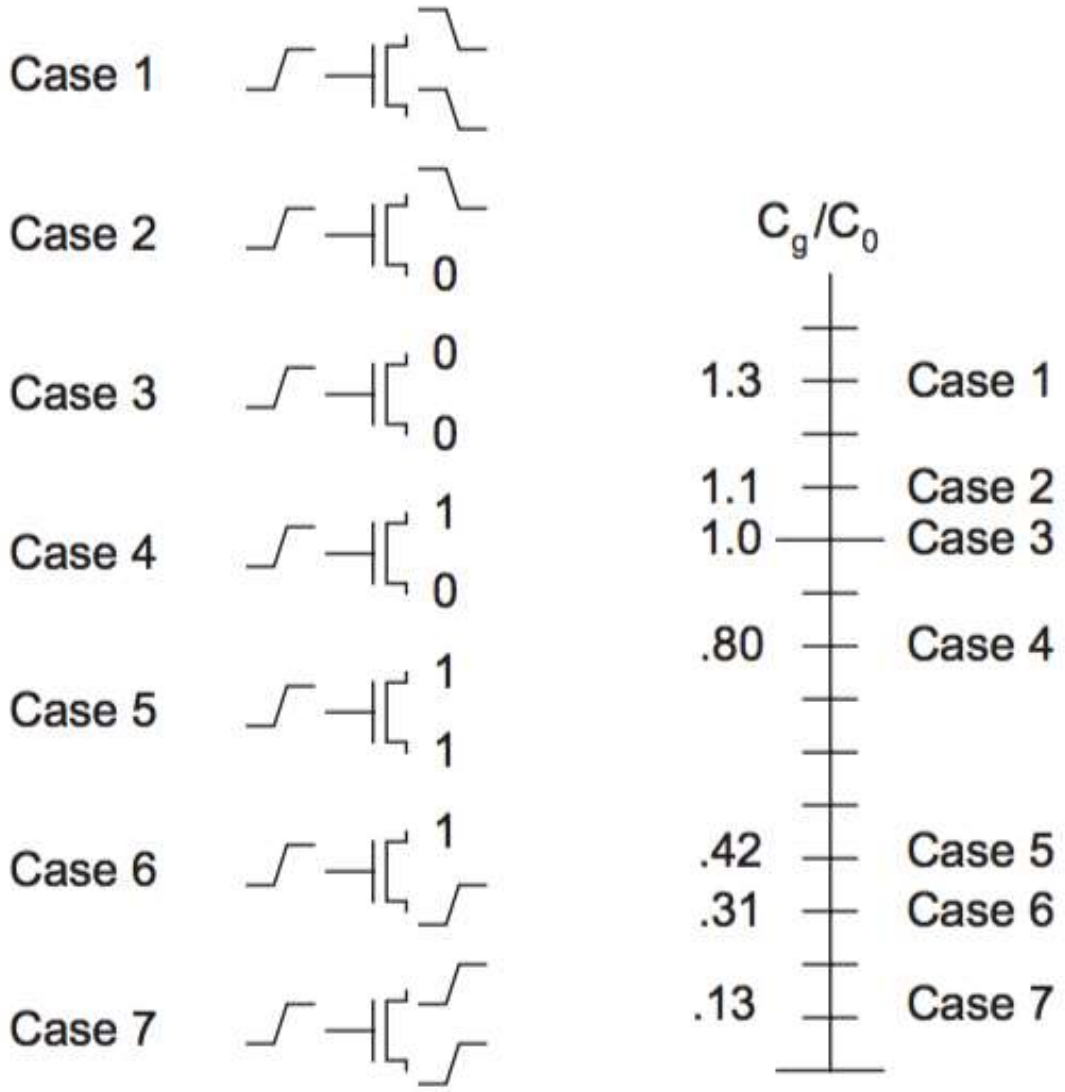
Gate Capacitance vs V_{ds}

- **Long-channel device:**
  C_{gd} becomes ~0 at saturation
- **Short-channel device:**
  more $C_{gd(overlap)}$ and $C_{gs(overlap)}$ factor





Data-Dependent Gate Capacitance

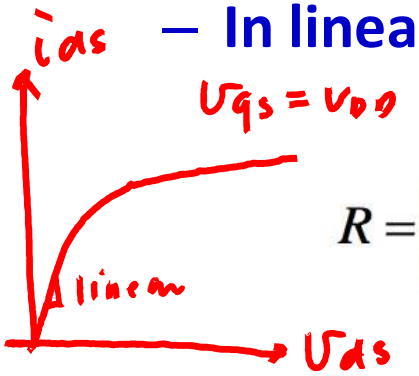


Outline

- Introduction
- Ideal I-V characteristics
- Nonideal Effects
- DC transfer characteristics
- C-V characteristics
- **Switch-level RC delay models**

Effective Resistance *R*

- ***R*: effective resistance** of unit NMOS (W_{min} and L_{min})
  - Unit PMOS has ***2R*** (or ***3R***) due to lower mobility
  - **In linear region**, inversely proportional to W/L and V_{gs}



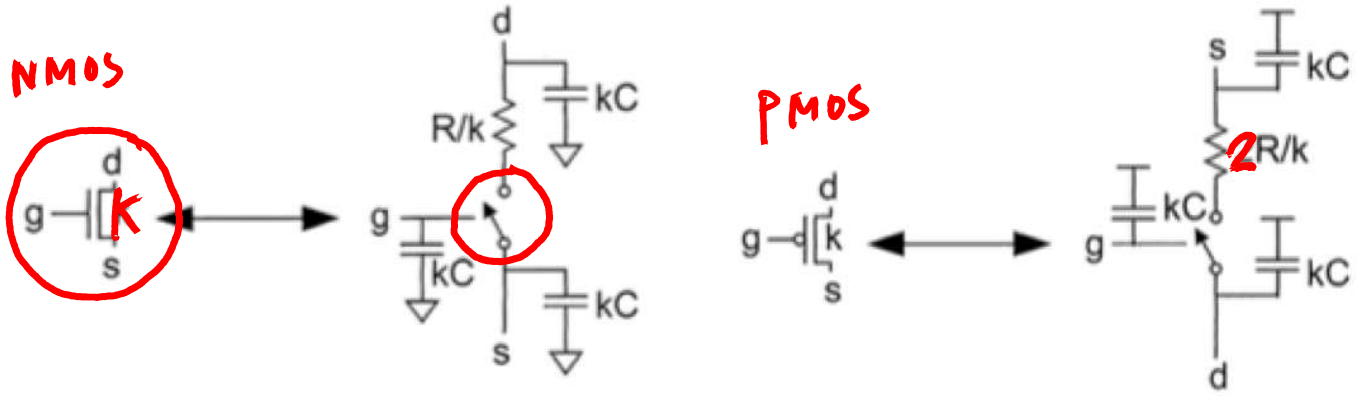
$$R = \left(\frac{\partial I_{ds}}{\partial V_{ds}}\right)^{-1} = \frac{1}{\beta\left(V_{gs} - V_t\right)} = \frac{1}{\mu C_{ox}} \frac{L}{W} \frac{1}{\left(V_{gs} - V_t\right)}$$

$V_{gs} \approx V_{DD}$

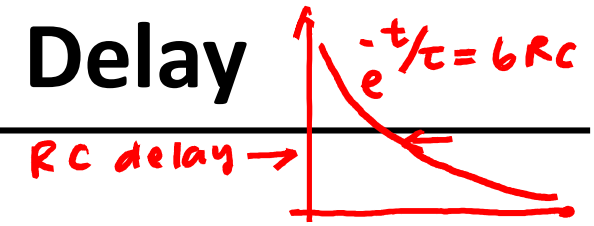
- ***C*: gate capacitance** of unit transistor (NMOS and PMOS)
  - Proportional to gate area W*L
- ***C*: S/D junction capacitance** of unit transistor
  - Proportional to gate width W

RC Circuit Model

- NMOS of k times unit width has resistance of ***R/k***, gate capacitance of ***kC***, and S/D capacitance of ***kC***
  - NMOS parasitic capacitance referenced to GND (p-sub)
- PMOS of k times unit width has resistance of ***2R/k***, gate capacitance of ***kC***, and S/D capacitance of ***kC***
  - PMOS parasitic capacitance referenced to V_{DD} (n-well)

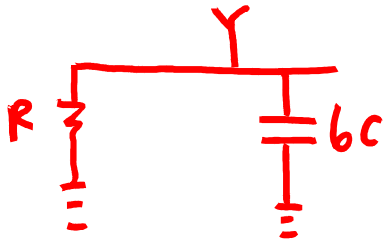


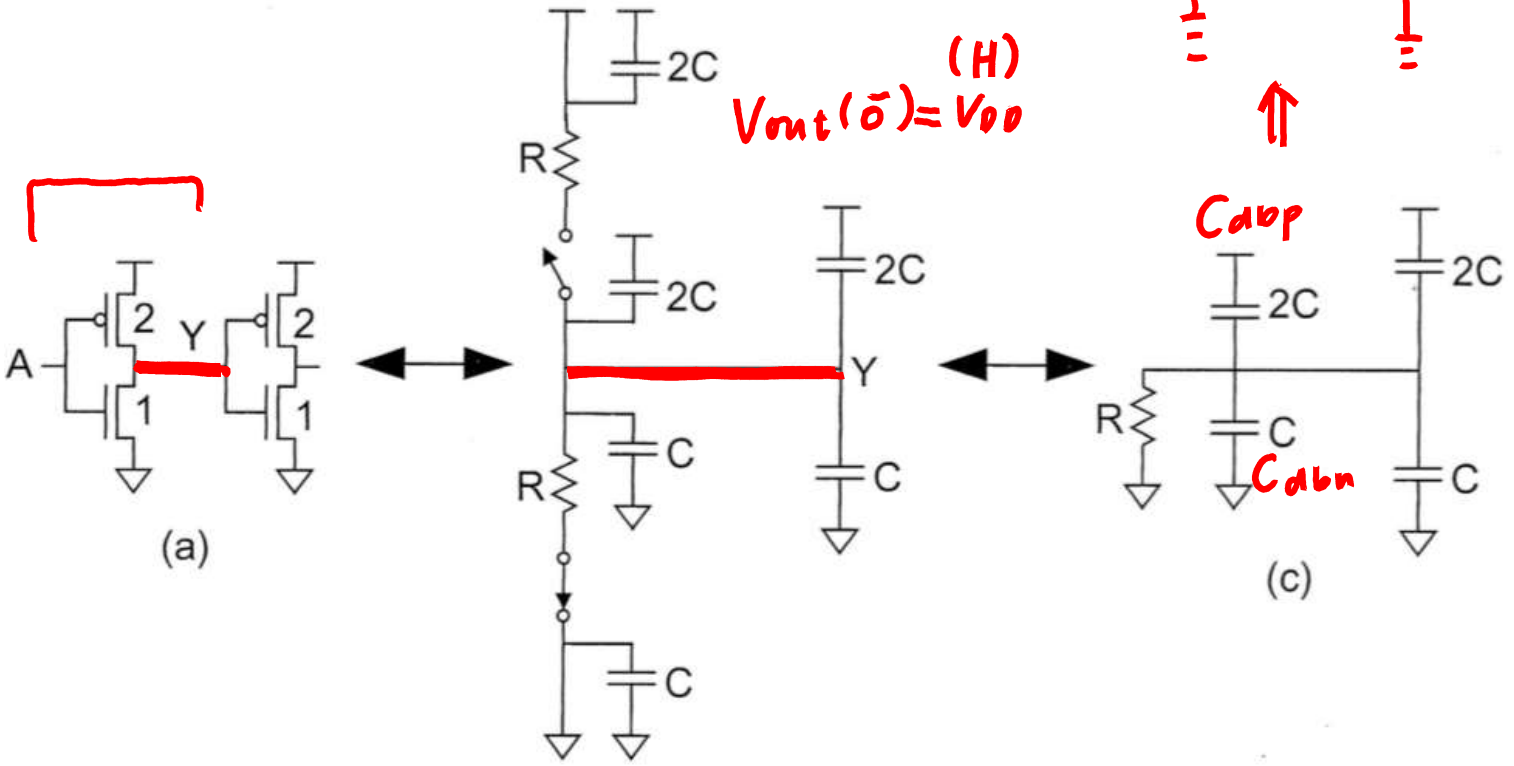
Inverter Propagation Delay



- Fanout-of-1 inverter
  - Choose PMOS width to be 2x~3x of NMOS width
  - $t_{pd} = R(6C) = 6RC$

$V_{out}(\bar{0}) = V_{DD}$

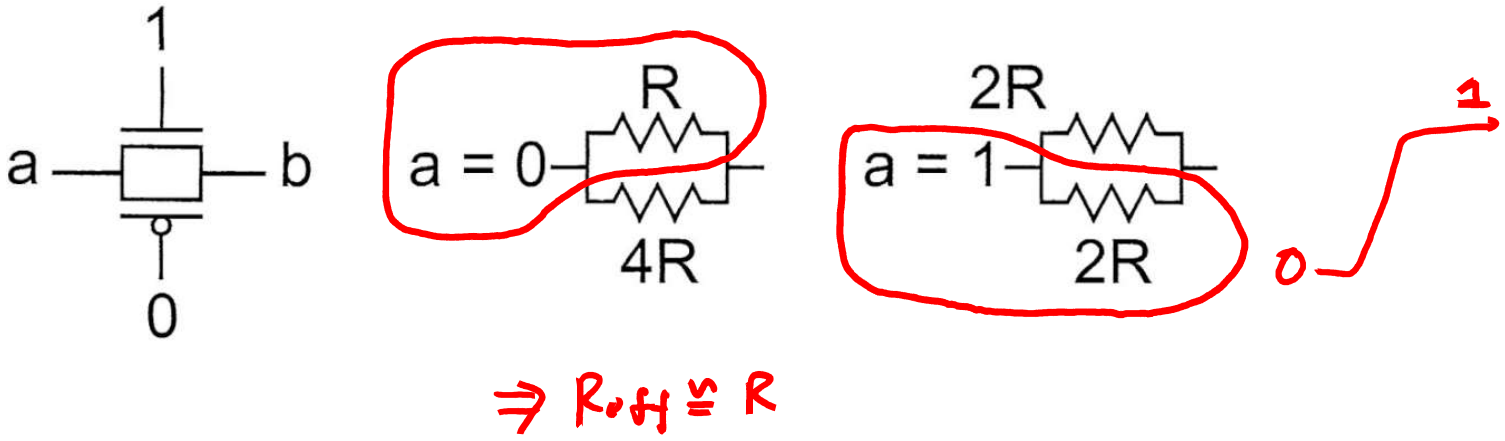




R of Transmission Gate

pass transistor

- Parallel combination of NMOS and PMOS
  - Depend on signal to pass
  - PMOS pass 0 weakly with larger resistance ***4R***
  - NMOS pass 1 weakly with larger resistance ***2R***
  - Usually the same size for NMOS and PMOS
  - Increase size → R ↓ C ↑ → need to check the trade-off



⇒ $R_{eff} \cong R$