Lecture 5:
CMOS Transistor Theory

Mark McDermott
Electrical and Computer Engineering
The University of Texas at Austin

Outline

- Introduction
- MOS Capacitor
- nMOS I-V Characteristics
- pMOS I-V Characteristics
- Gate and Diffusion Capacitance
- MOS Channel resistance
- Resistors & RC approximation

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

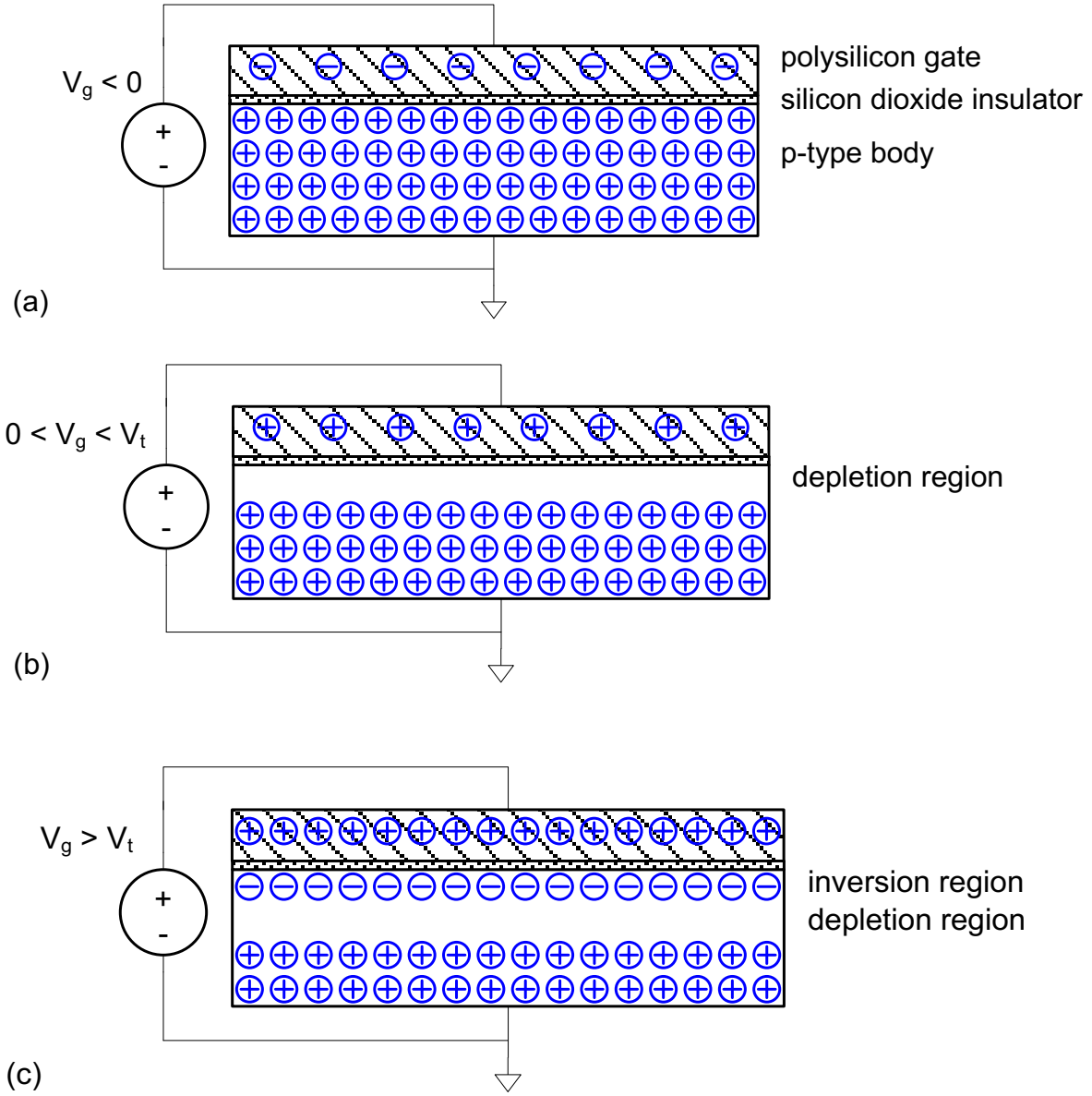
Electrical Properties of MOS Devices

- **Necessary to understand the basic electrical properties of the MOS transistor (geometry => electrical), e.g., delay/power**
  - **Ensure that the circuits are robust**
  - **Create working layouts**
  - **Predict delays and power consumption**

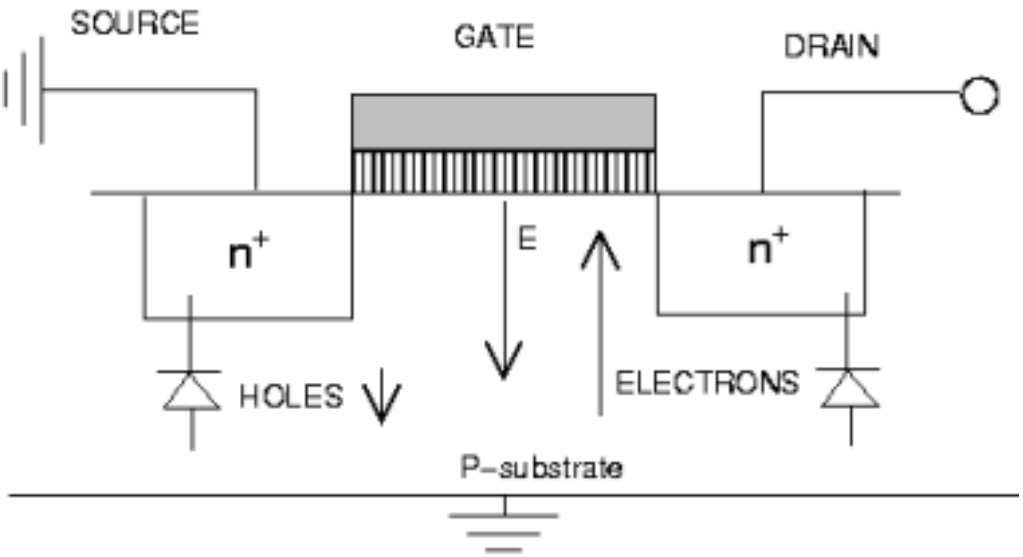
- **As technology advances and circuit dimensions scale down, electrical effects become more important**
  - **Secondary/non-ideal effects (next lecture)**

MOS Capacitor

- **Gate and body form MOS capacitor**
- **Operating modes**
  - **Accumulation**
  - **Depletion**
  - **Inversion**



The nMOS Transistor

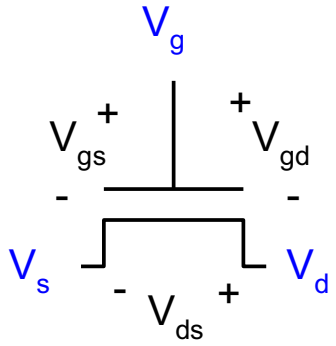


Moderately doped p- type substrate (or well) in which two heavily doped n+ regions, the Source and Drain are diffused

- **Gate is insulated from substrate by thin oxide**
  - **Resistance of oxide is > 10^{12} Ω, so current ~ 0**
- **Two types of nMOS transistor**
  - **Enhancement mode: non conducting when gate voltage $V_{gs} = V_{sb}$ (source voltage) (normally used)**
  - **Depletion mode: conducting when $V_{gs} = V_{sb}$**

Terminal Voltages

- **Mode of operation depends on V_g, V_d, V_s**
  - $V_{gs} = V_g - V_s$
  - $V_{gd} = V_g - V_d$
  - $V_{ds} = V_d - V_s = V_{gs} - V_{gd}$

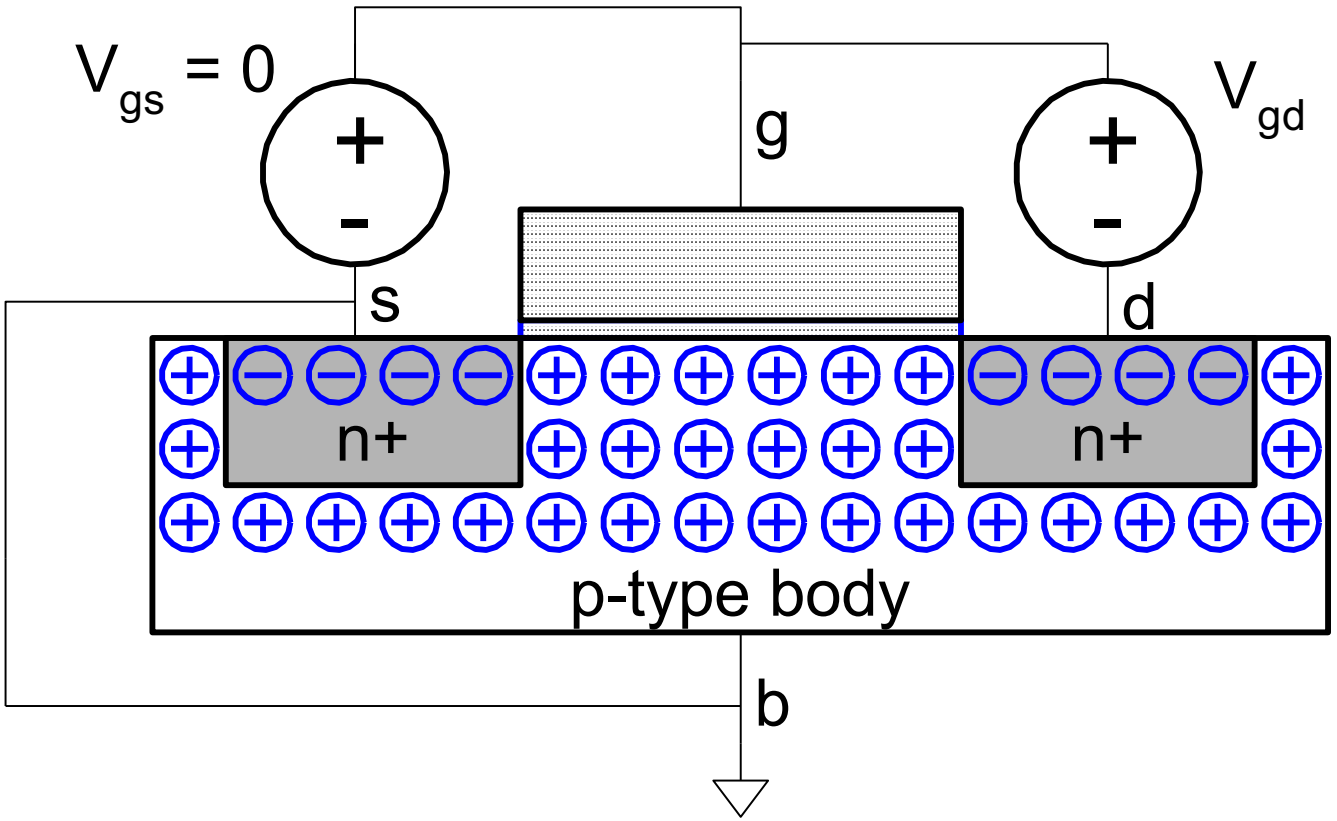


- **Source and drain are symmetric diffusion terminals**
  - By convention, source is terminal at lower voltage
  - Hence $V_{ds} \geq 0$
- **nMOS body is grounded; for simple designs, assume source is grounded too**
- **Three regions of operation**
  - *Cutoff*
  - *Linear*
  - *Saturation*

nMOS Cutoff

- **No channel**

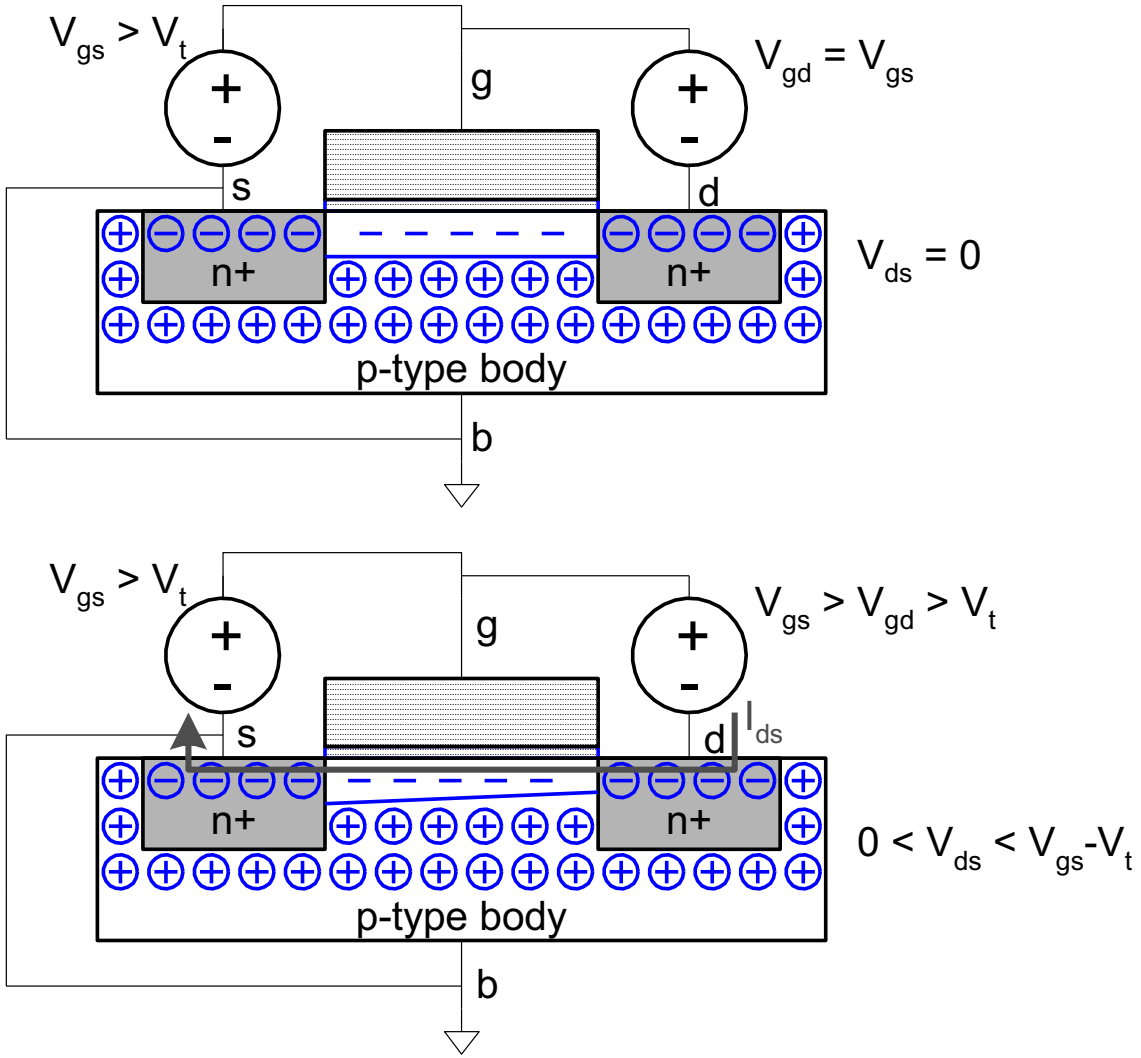
  $I_{ds} \approx 0$



nMOS Linear

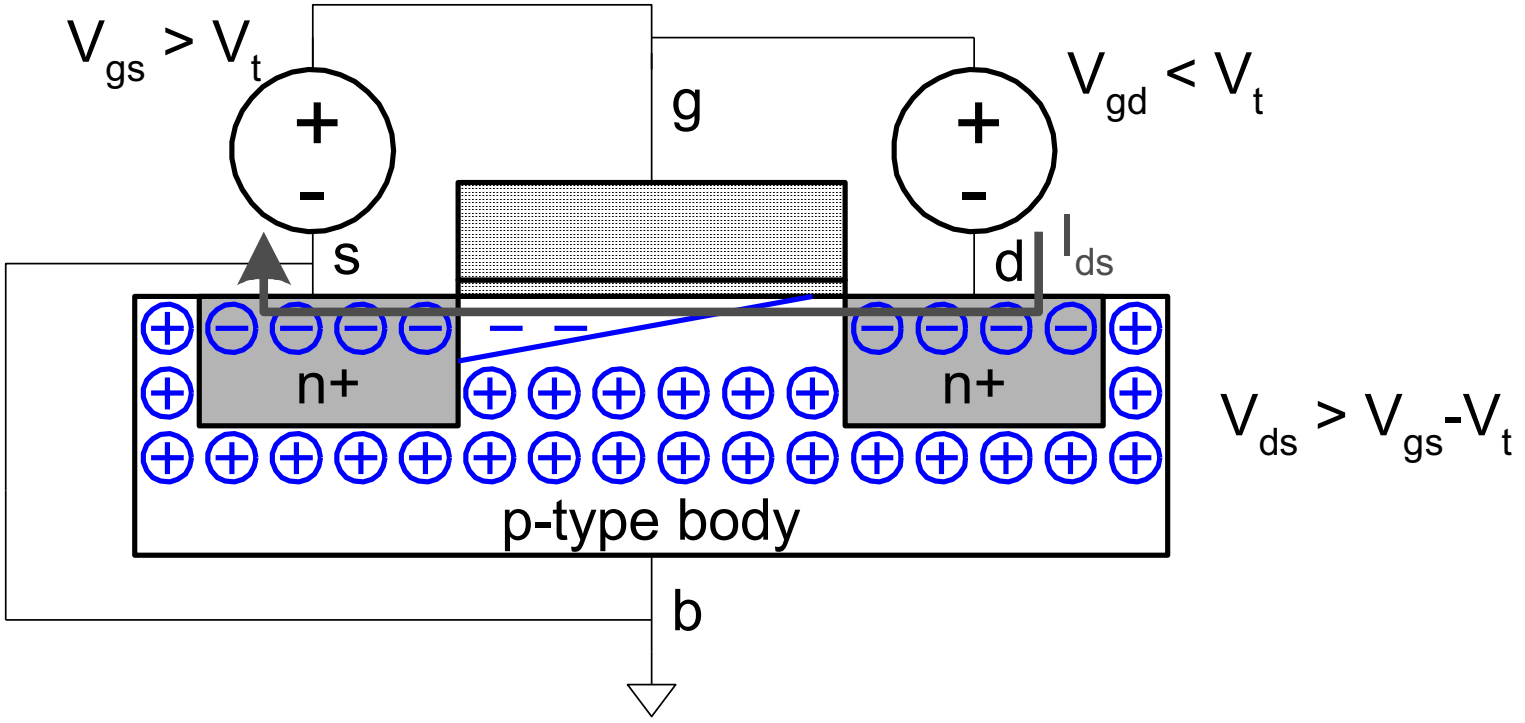
- Channel forms
- Current flows from d to s
  - e^- from s to d
- I_{ds} increases with V_{ds}
  - Similar to linear resistor
- Since there is a threshold voltage (V_t) required to invert the charge under the gate, this means that the effective gate voltage is:

$$V_g = V_{gs} - V_t$$

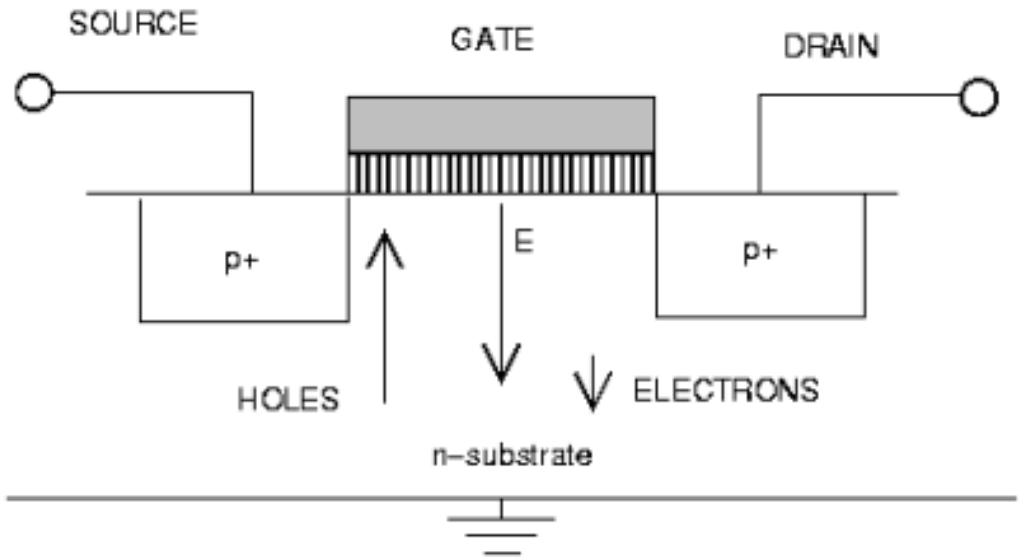


nMOS Saturation

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



The pMOS Transistor



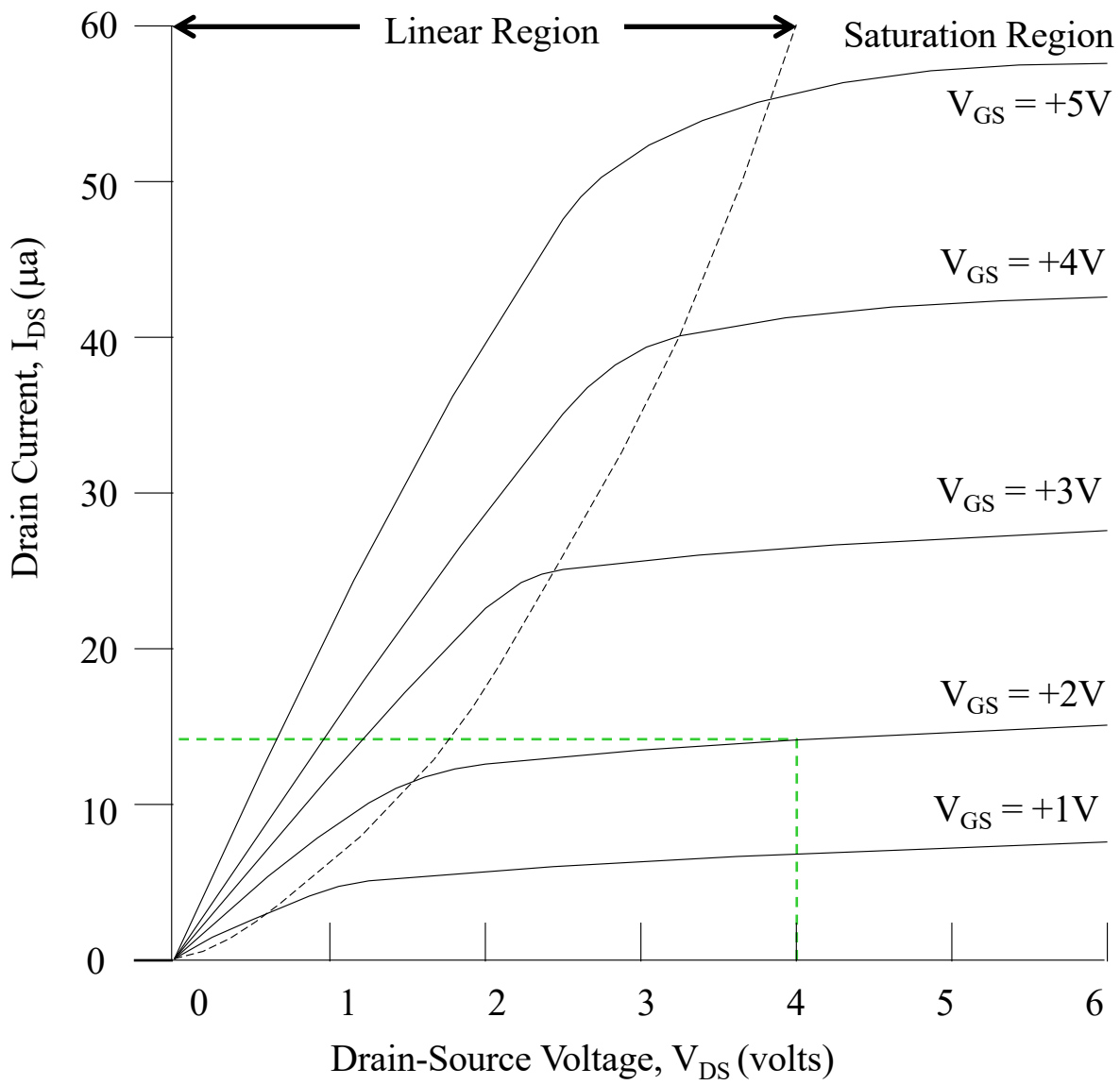
Moderately doped n- type substrate (or well) in which two heavily doped p+ regions, the Source and Drain are diffused

- **Application of a negative gate voltage (w.r.t. source) draws holes into the region below the gate; channel changes from n to p-type (source-drain conduction path)**

- **Conduction due to holes; negative V_d sweeps holes from source (through channel) to drain**

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

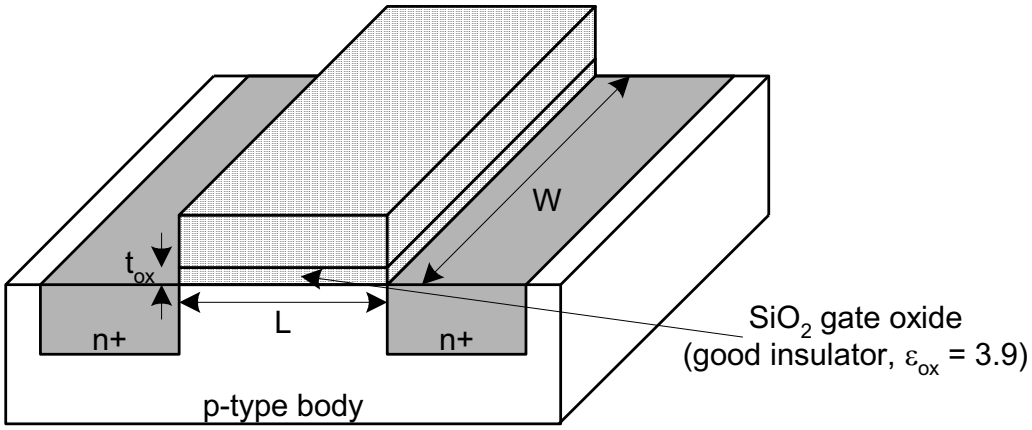
- **MOS structure looks like parallel plate capacitor while operating in inversion**

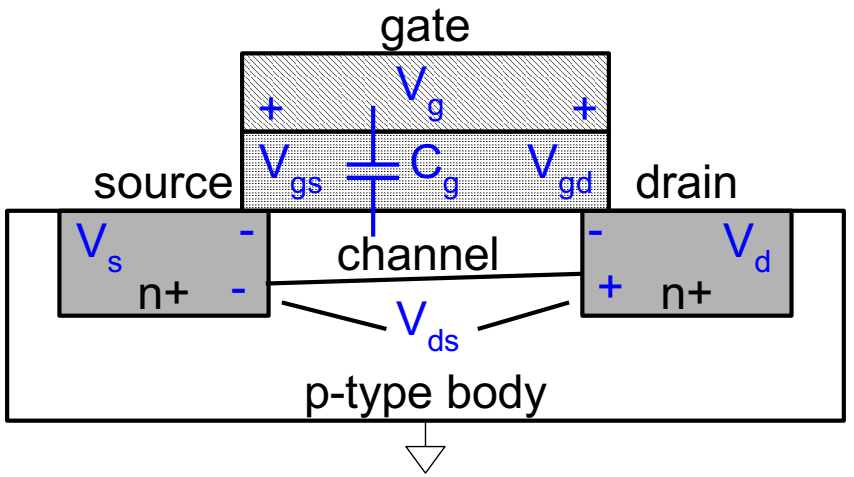
  Gate – oxide – channel

  $Q_{channel} = CV$

  $C = C_g = \varepsilon_{ox} WL/t_{ox} = C_{ox} WL$ where $C_{ox} = e_{ox} / t_{ox}$

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





Carrier Velocity

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

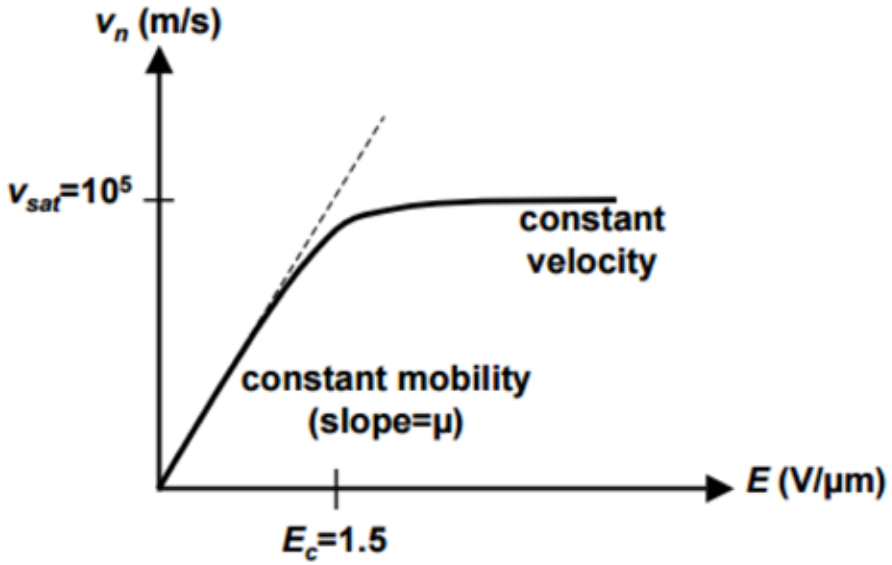
$$v = \mu E_{ds}$$ **where μ is the electron/hole mobility**

where $$E_{ds} = \frac{V_{ds}}{L}$$

- **Time for carrier to cross channel:**

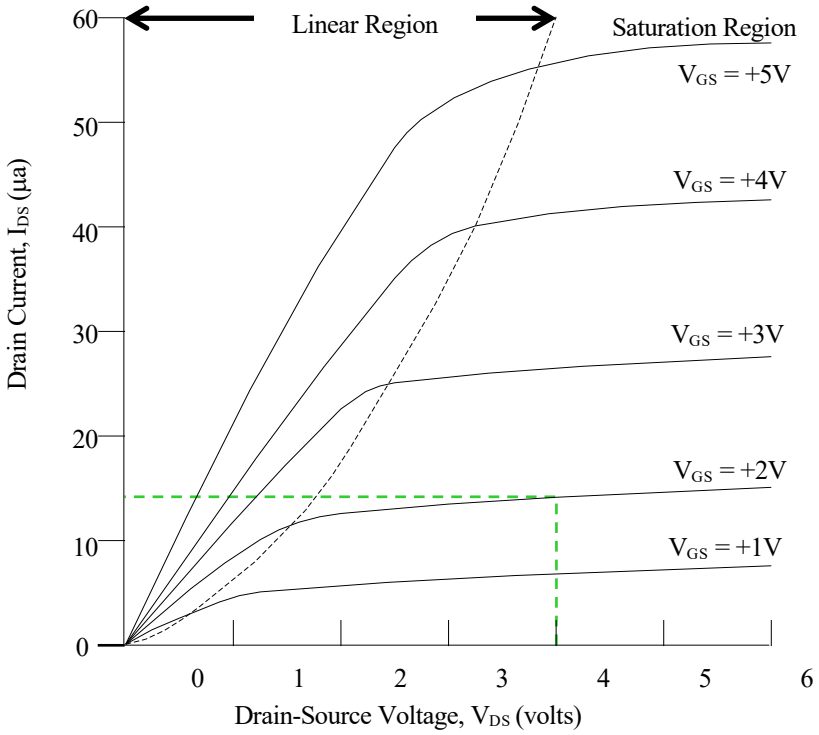
$$\tau_{sd} = \frac{Length\ of\ the\ channel\ (L)}{Velocity\ (v)}$$

or $$\tau_{sd} = \frac{L^2}{\mu V_{ds}}$$



nMOS Linear I-V

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



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

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

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

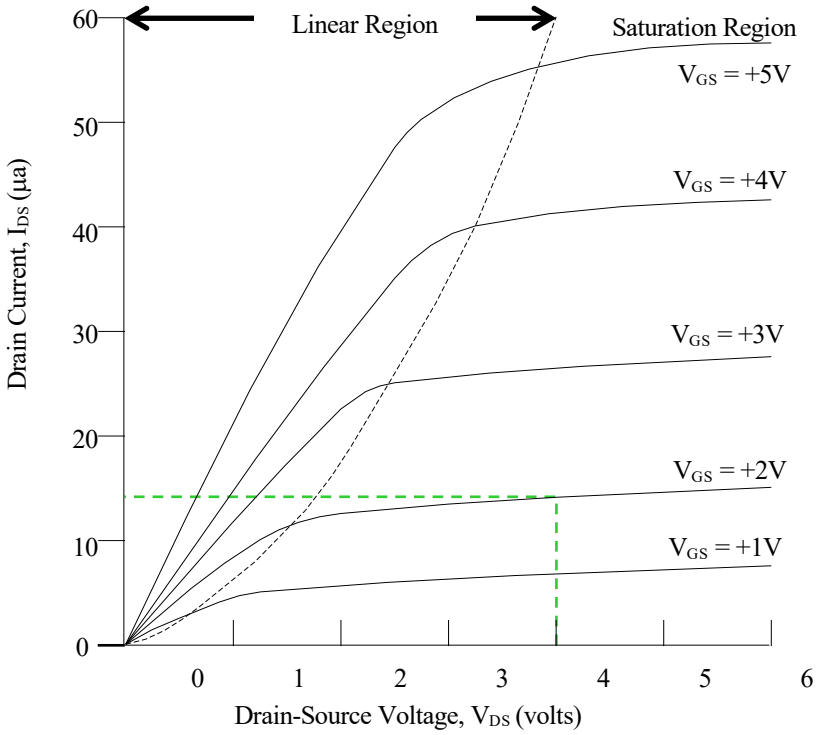
nMOS Saturation I-V

- **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**

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

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



nMOS I-V Summary

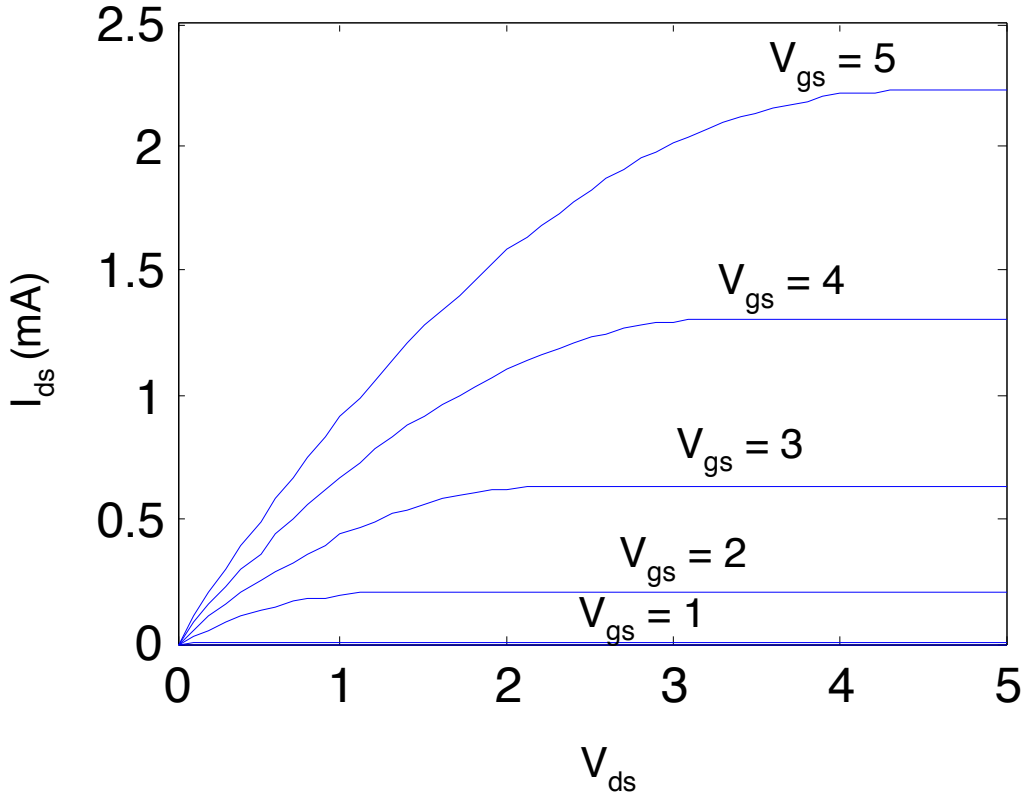
- ***Shockley* 1st order transistor models**

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

Example

- **Example: a 0.6 μm process from AMI semiconductor**
  - t_{ox} = 100 Å
  - μ = 350 cm²/V*s
  - V_t = 0.7 V

- **Plot I_{ds} vs. V_{ds}**
  - V_{gs} = 0, 1, 2, 3, 4, 5
  - Use W/L = 4/2 λ



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

pMOS I-V

- **All dopings and voltages are inverted for pMOS**

- **Mobility μ_p is determined by holes**
  - **Typically 2-3x lower than that of electrons μ_n for older technologies.**
  - **Approaching 1 for gate lengths < 20nm.**

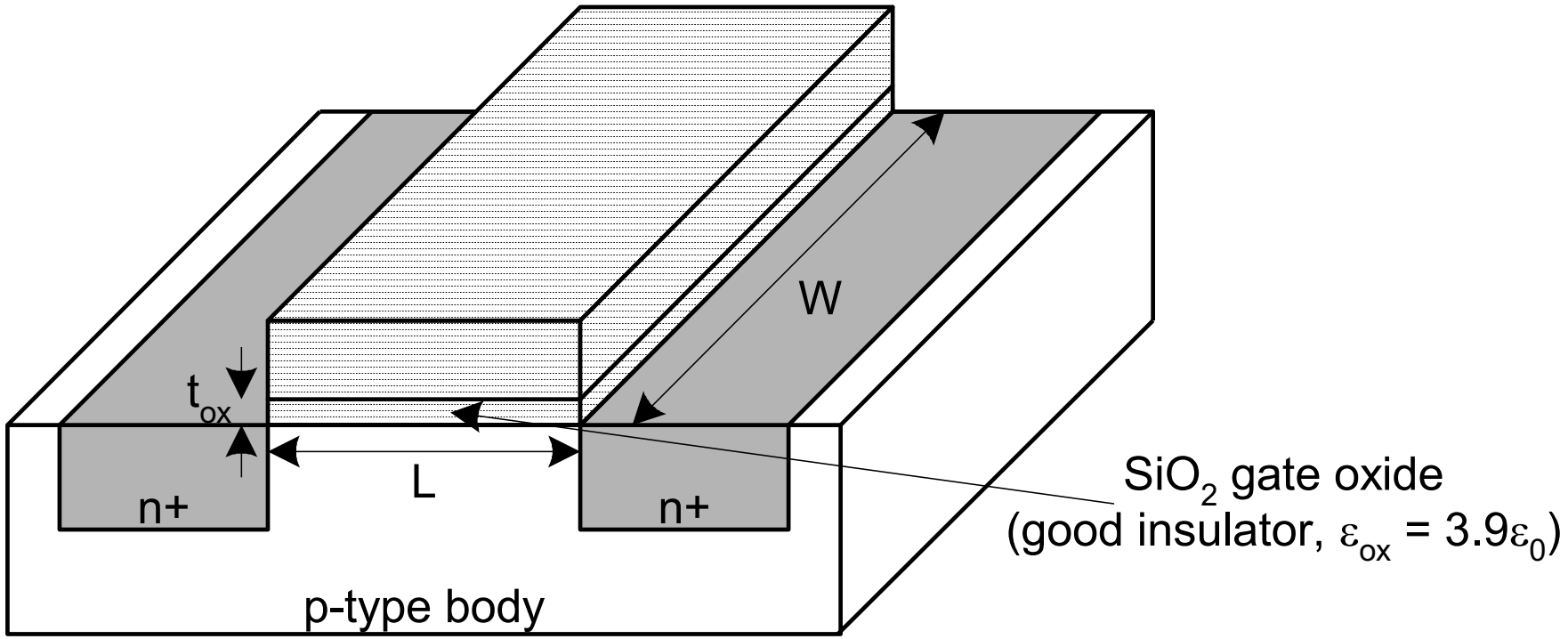
- **Thus pMOS must be wider to provide the same current**
  - **Simple assumption, μ_n / μ_p = 2 for technologies > 20nm**

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

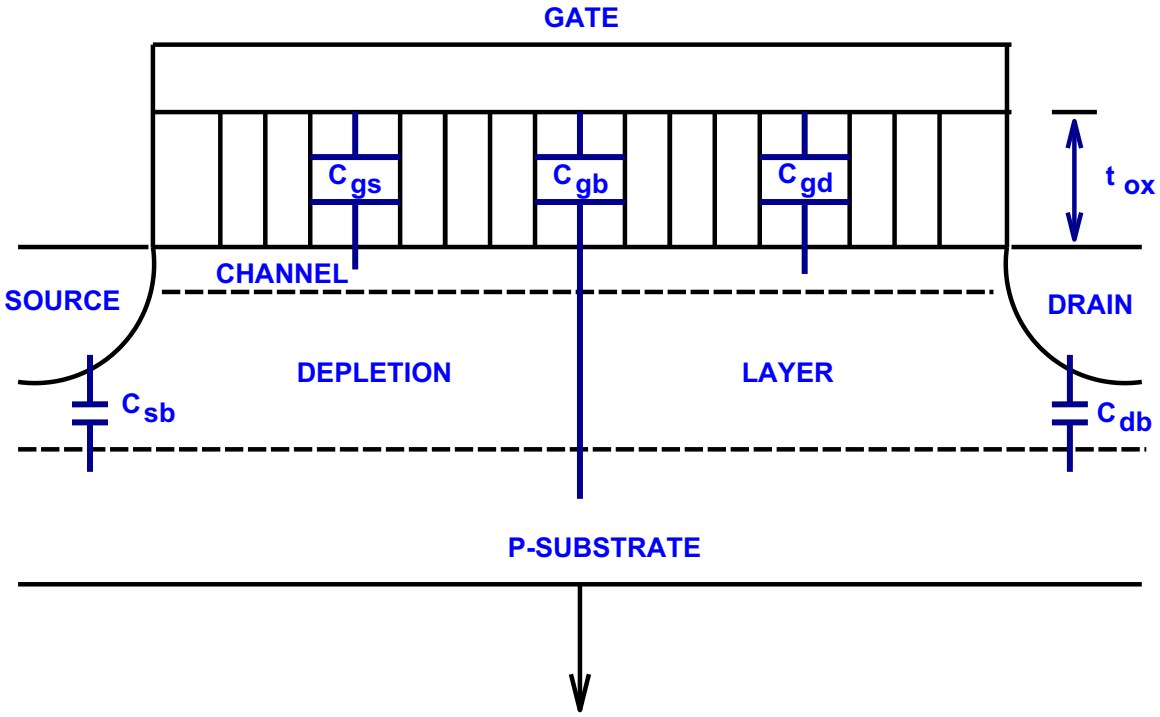
Gate Capacitance

- Approximate channel as connected to source
- $C_{gs} = \varepsilon_{ox}WL/t_{ox} = C_{ox}WL = C_{per/micron}W$
- $C_{permicron}$ is typically about 2 fF/μm for minimum channel length device



Capacitance Estimation

- **The dynamic response (switching speed) of a CMOS circuit is very dependent on parasitic capacitances associated with the circuit**



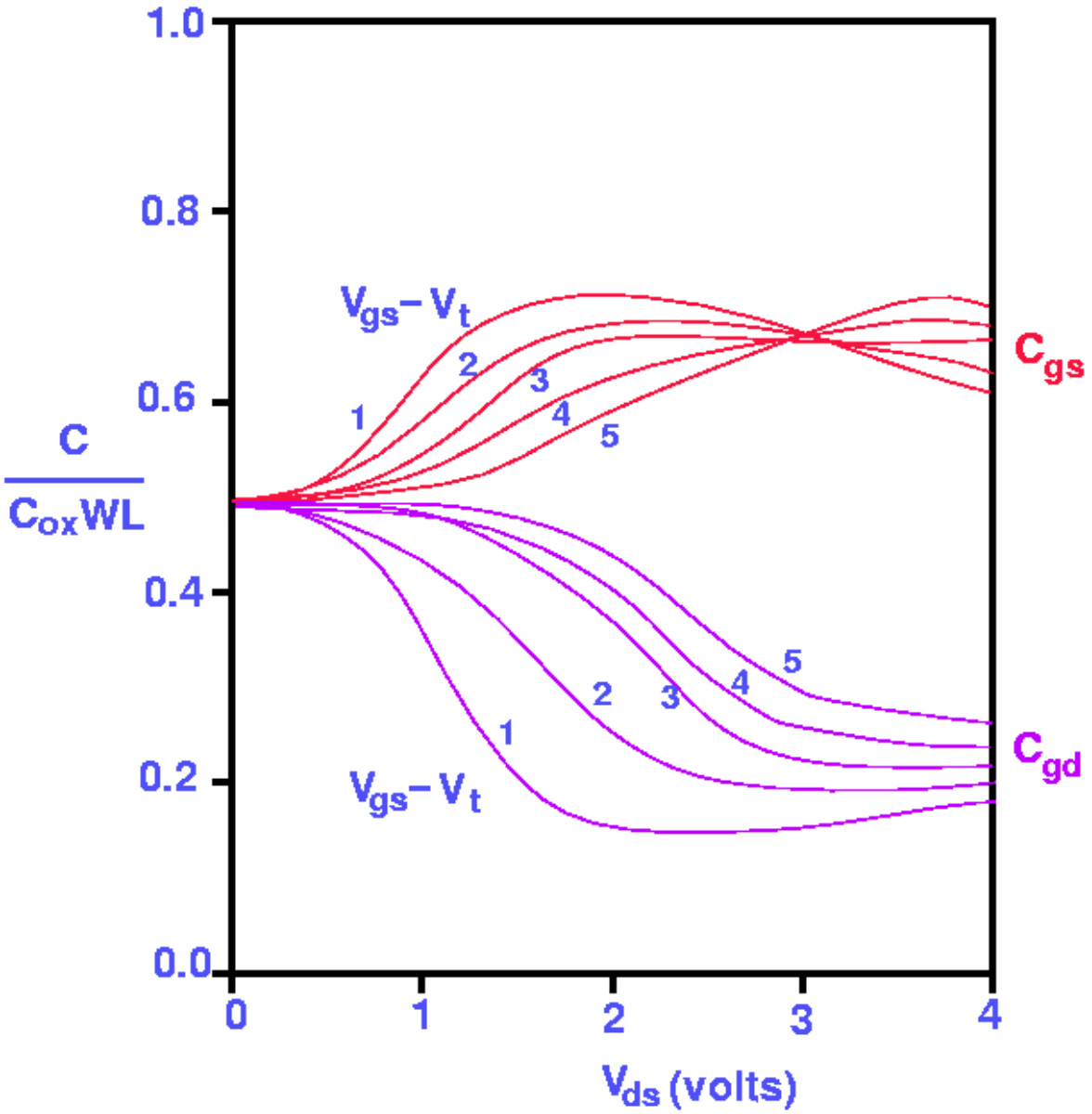
Parasitic Capacitances:

C_{gs}, C_{gd} = gate-to-channel capacitances lumped at source and drain regions

C_{sb}, C_{db} = source and drain diffusion capacitances to bulk (substrate)

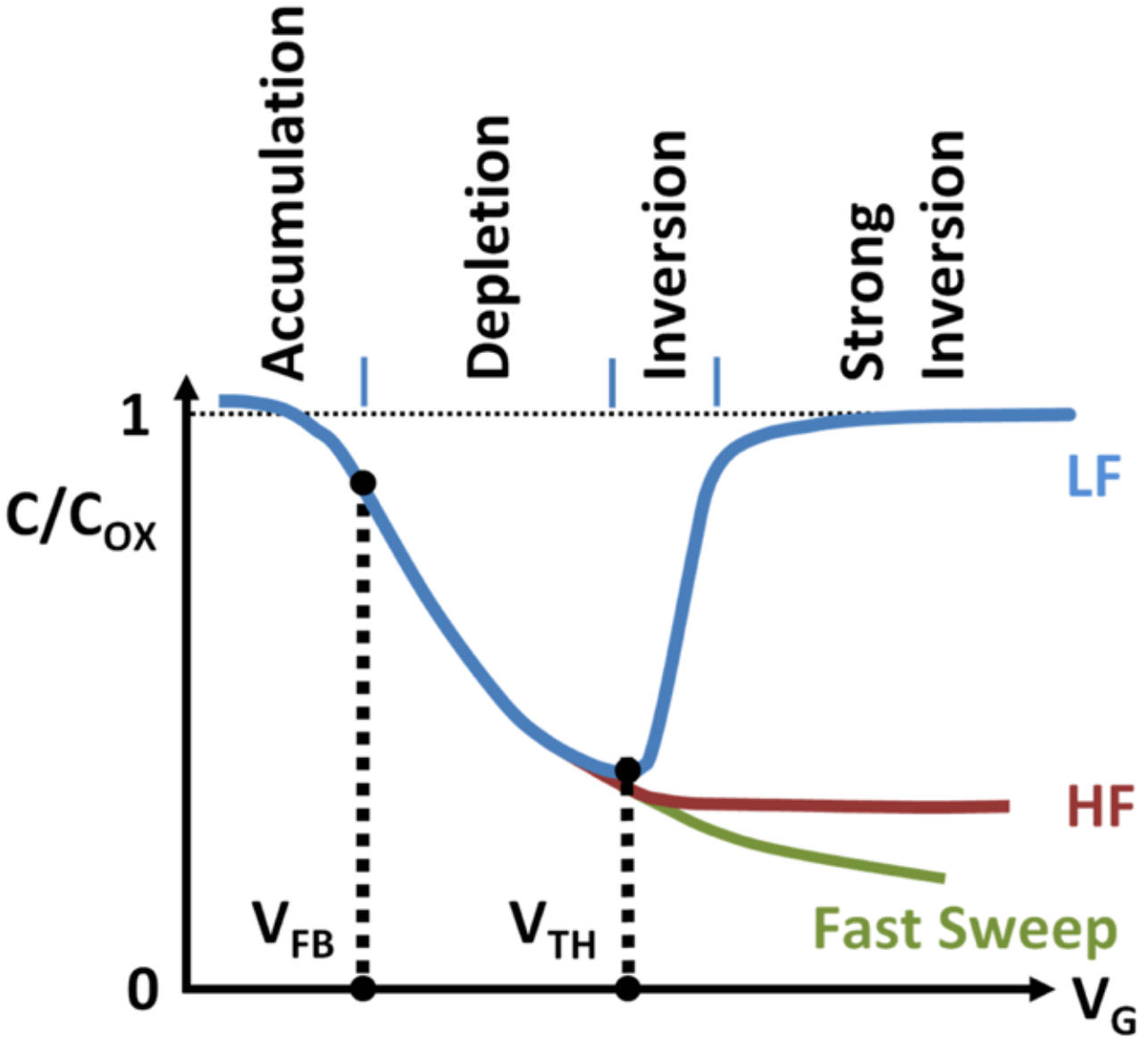
Add **routing capacitances** to get total capacitance.

Gate Capacitance of MOS Transistor

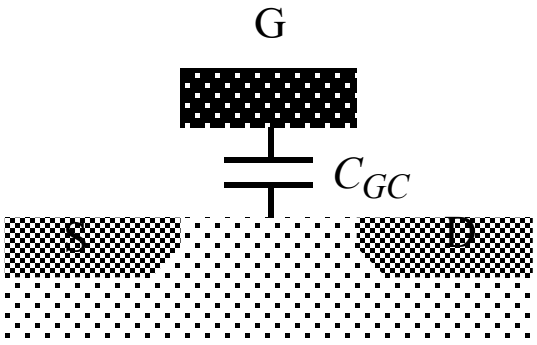


W = 49.2 μ,
L = 0.75μ

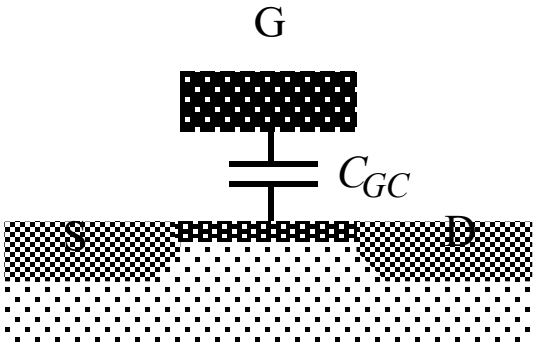
Gate Capacitance of MOS Transistor (cont.)



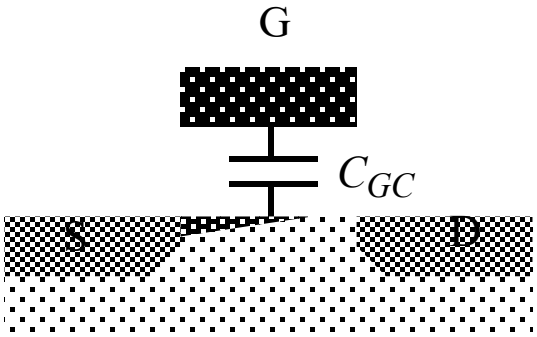
Gate Capacitance: Operation Region Dependence



Cut-off



Linear (Triode)

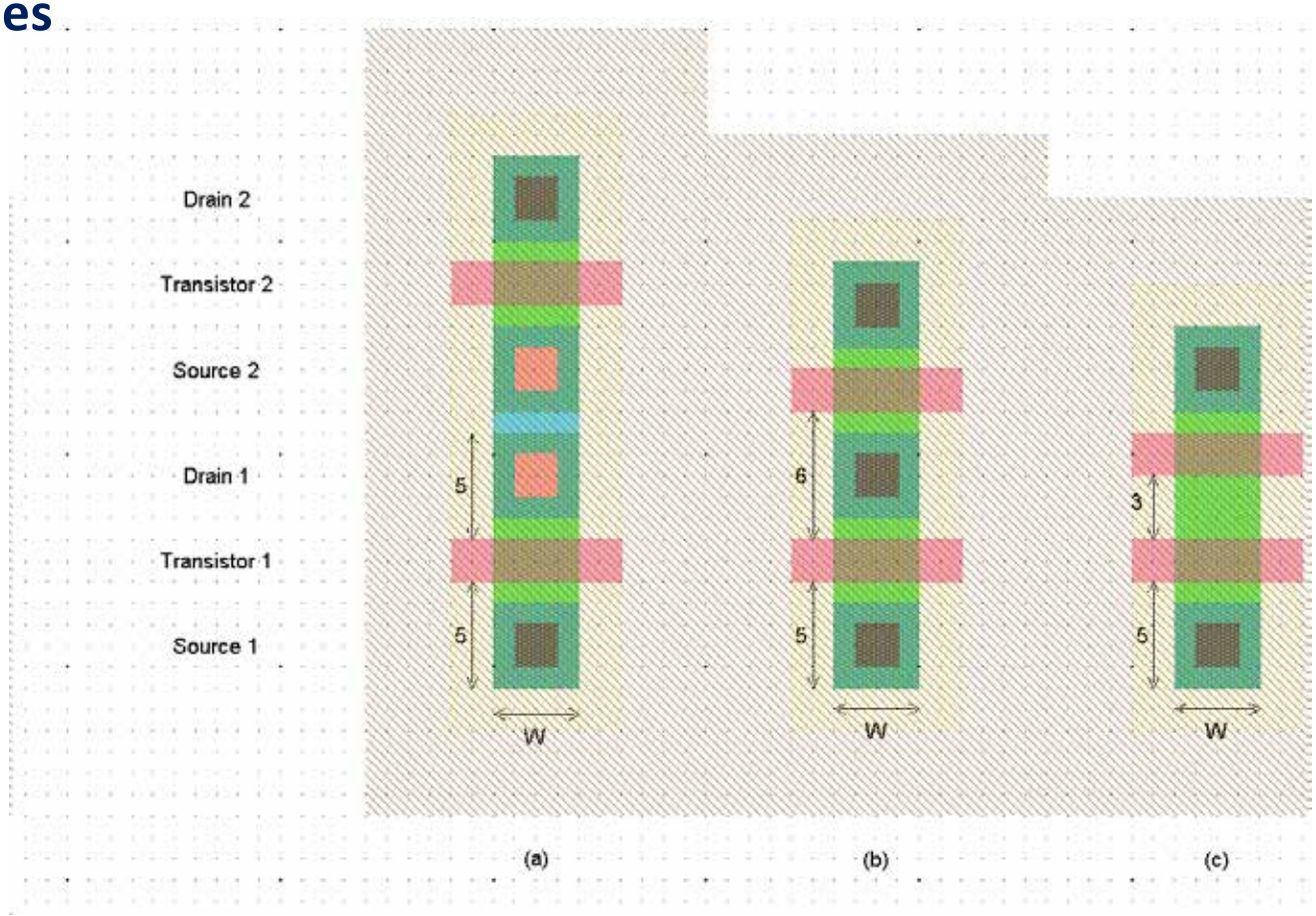


Saturation

Operation Region	C_{gb}	C_{gs}	C_{gd}
Cutoff	$C_{ox}WL_{eff}$	0	0
Triode	0	$C_{ox}WL_{eff}/2$	$C_{ox}WL_{eff}/2$
Saturation	0	$(2/3)C_{ox}WL_{eff}$	0

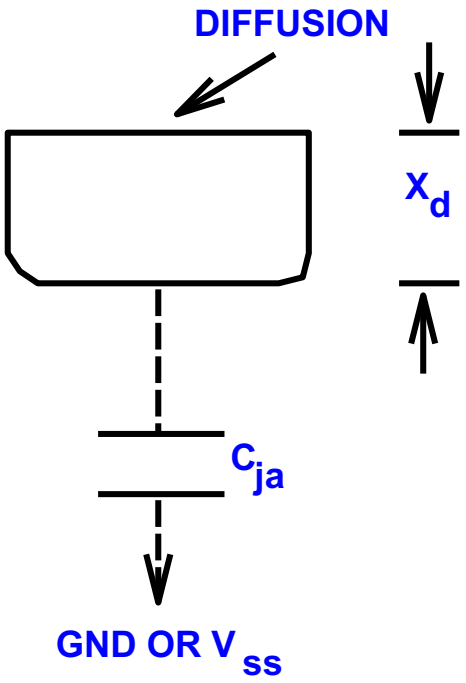
Diffusion Capacitance

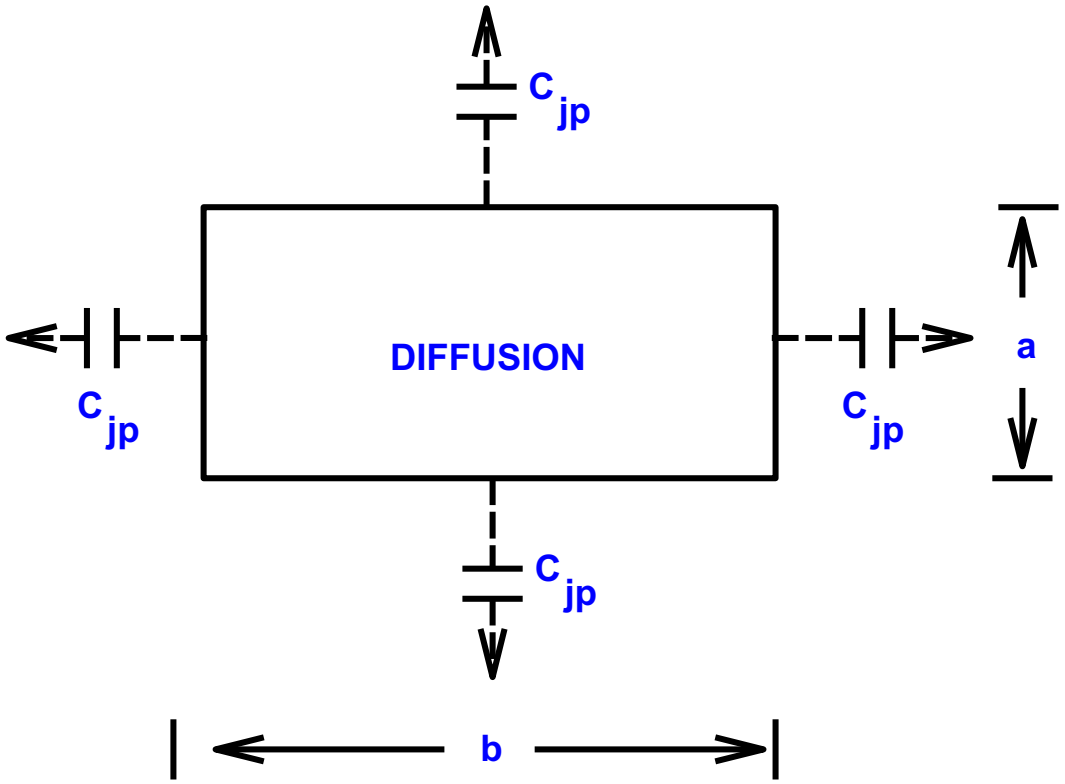
- C_{sb}, C_{db} from Source/Drain
- Undesirable, called *parasitic* capacitance
- Capacitance depends on area and perimeter
  - Use small diffusion nodes
  - Comparable to C_g for contacted diff
  - ½ C_g for uncontacted
  - Varies with process



Area and Periphery Capacitance

C_{jp} = Periphery capacitance (pf/μ)

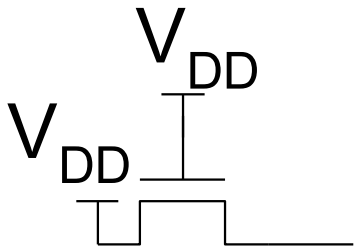




C_{ja} = Area capacitance (pf/μ^2)

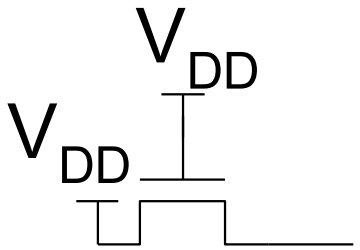
Pass Transistors

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

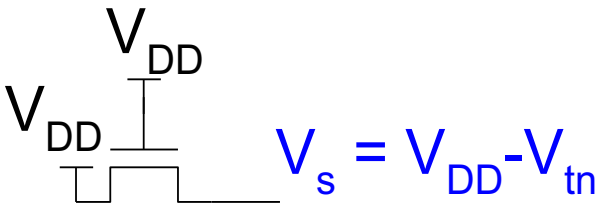


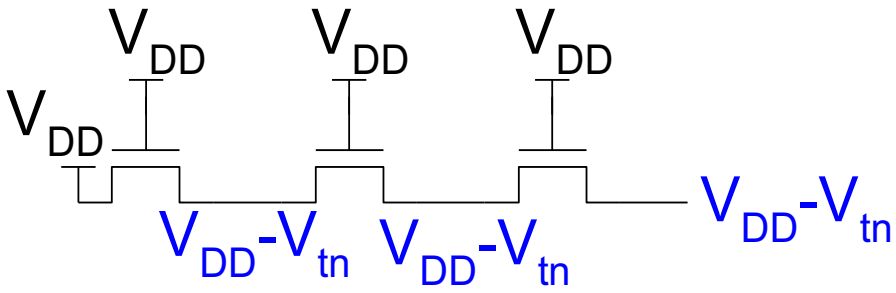
Pass Transistors

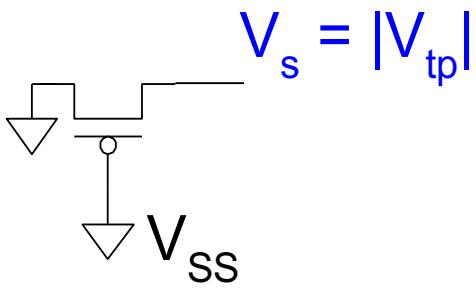
- We have assumed source is grounded
- What if source > 0?
  - e.g. pass transistor passing V_{DD}
- $V_g = V_{DD}$
  - If $V_s > V_{DD}-V_t$, $V_{gs} < V_t$
  - Hence transistor would turn itself off
- nMOS pass transistors pull no higher than $V_{DD}-V_{tn}$
  - Called a degraded "1"
  - Approach degraded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

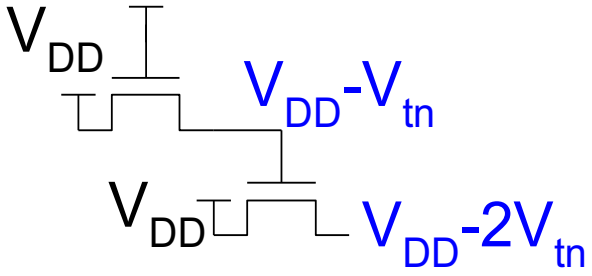


Pass Transistor Circuits





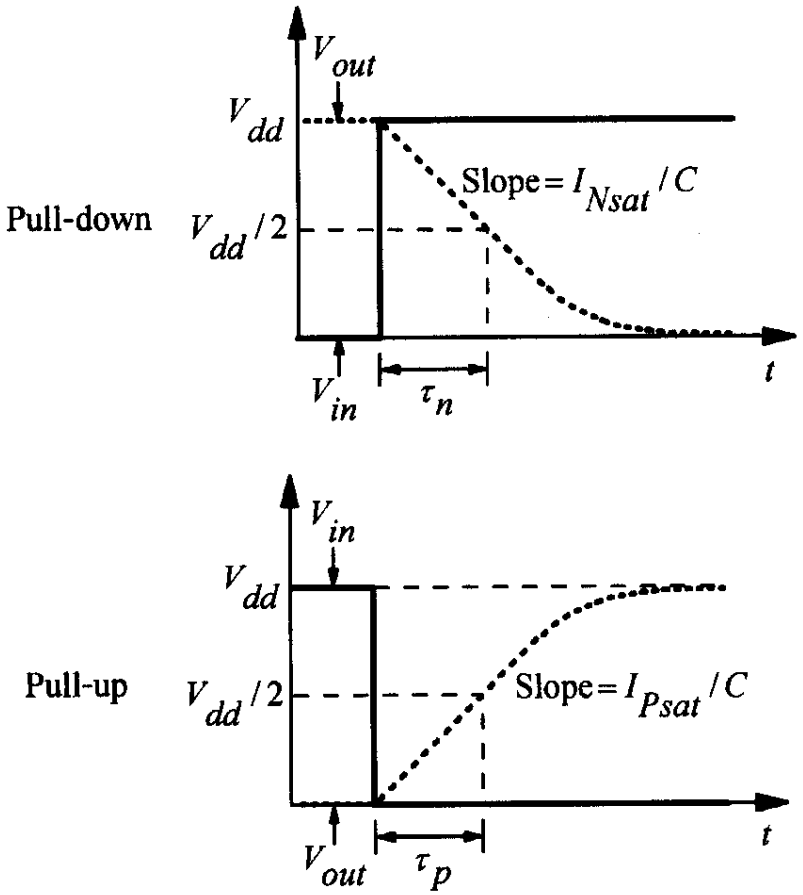




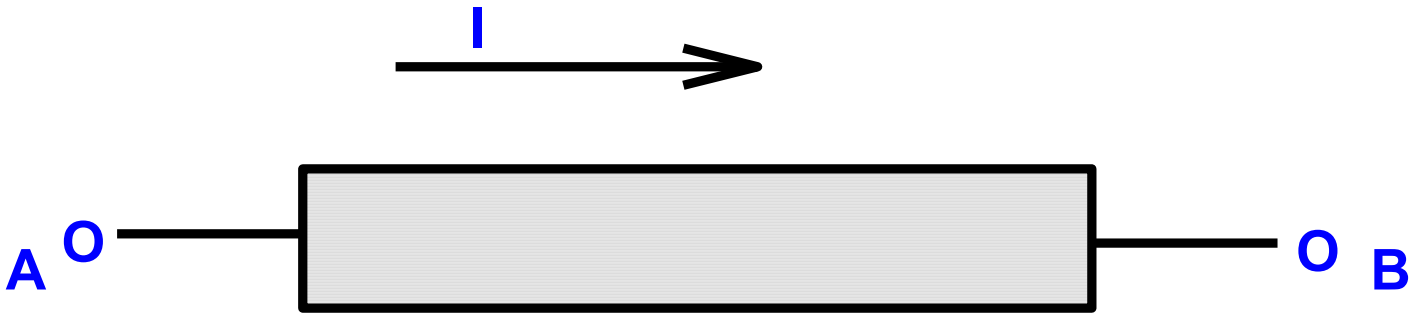
NOTE: These values are for steady state conditions only.

Effective Resistance

- **Shockley models have limited value**
  - Not accurate enough for modern transistors
  - Too complicated for 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



The MOS Resistor



- **Resistance of a bar of uniform material**

$$R = \frac{\rho \times L}{A} = (\frac{\rho}{t})(\frac{L}{W})$$

where ρ = resistivity of the material,

A = cross-section of the resistor

t, W = thickness, width of the material

- **The channel resistance of a MOS transistor in the linear region**

$$R_c = k(\frac{L}{W}),$$

where $k = \frac{1}{\mu C_{ox}(V_{gs} - V_t)}$

Resistance of Turned-On Transistor

W/L ratio defines size of N and P channel transistors

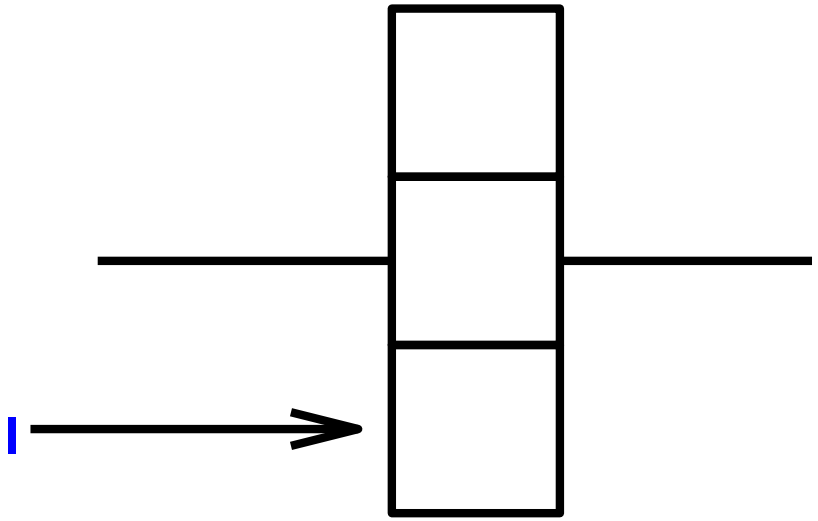
Channel resistance of turned-on transistor is:

$$R_c = k(\frac{L}{W}),\ \text{where}\ k = \frac{1}{\mu C_{ox}(V_{gs} - V_t)}$$

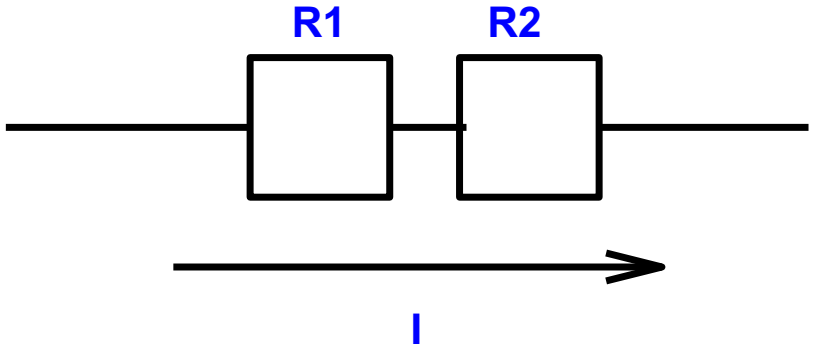
k is in the range of 1000 -- 30,000 Ω /□

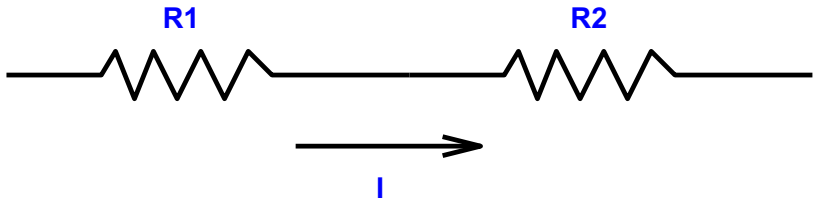
Resistance increases by about 0.25%/℃ above 25℃

R = L/W * □



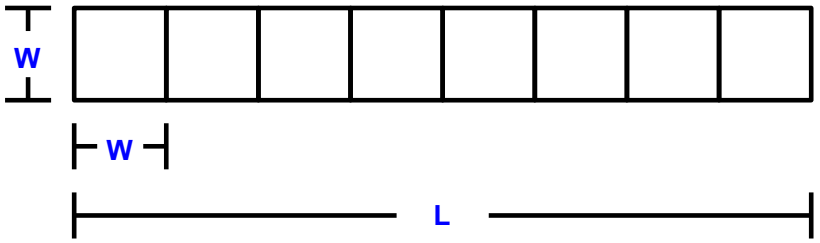
Resistors Connected in Series





$$R = R_1 + R_2 = 2\square s$$

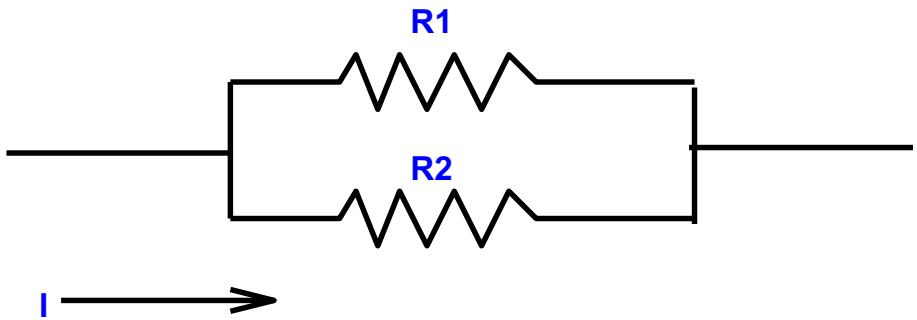
Sheet Resistance, R_s: Any material on the chip can be divided into squares W on a side with (R_s Ω/□) $R = R_s(L/W)\Omega$



Typical sheet resistances (Ω/□) for 0.25μ TSMC process:
4.7 for N+, 3.5 for P+, 4.2 for Poly, 0.06 for Metal1,
0.08 for Metal2 - Metal4, 0.03 for Metal5, and 1190 for the N-well
Increase of about 0.3%/˚ C (metal, poly), 1%/˚ C (diffusion)

Resistors Connected in Parallel

$$R = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2}}$$



For two squares in parallel, the equivalent resistance is ½ □

Expressing sheet resistance in □s simplifies the calculations

Contact resistance becomes more important as processes scale down

About 6 Ω for N+, P+, Poly, Metal 4

2 Ω for Metal2

4 Ω for Metal3

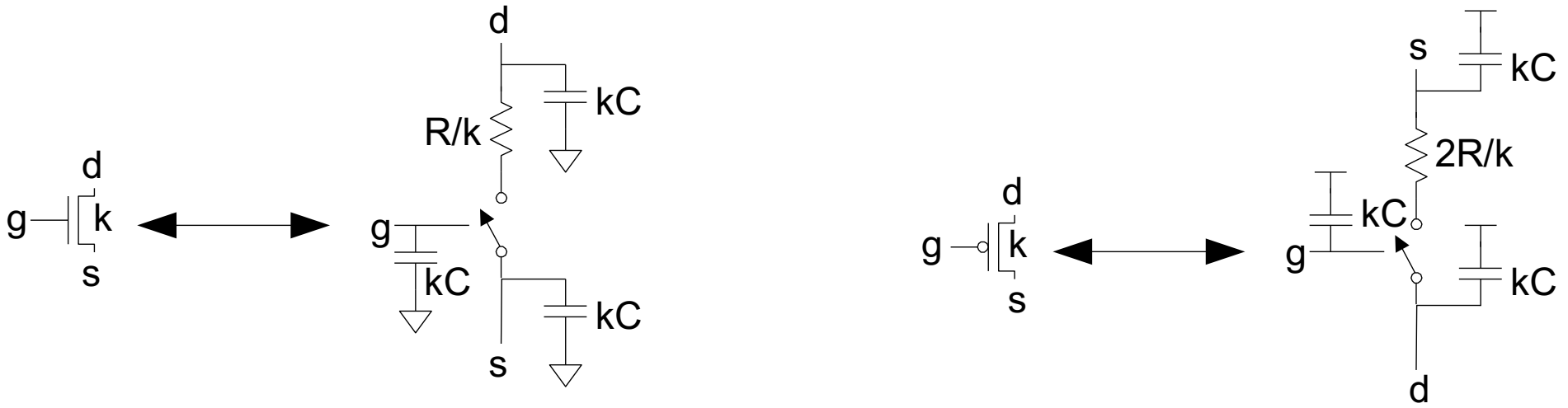
8 Ω for Metal5

in a 0.25m TSMC process

Use multiple contacts/vias for low resistance connections

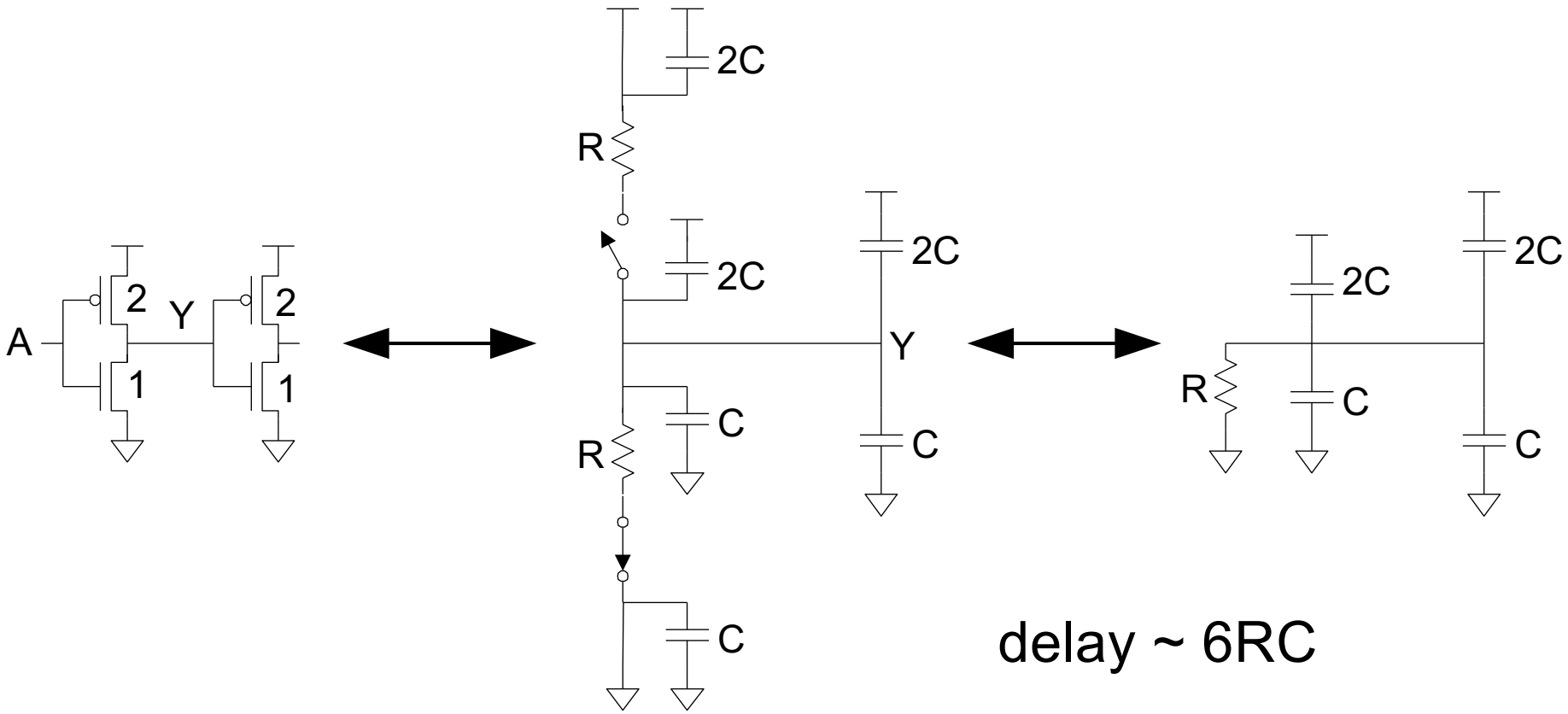
RC Delay Model

- **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**



Inverter Delay Estimate

- **Estimate the delay of a fanout-of-1 inverter**



Backup

Example of SPICE Deck

```
*file asic3.sp   test of 10 stage lumped mos model
* comments
.option scale=1e-6 post=2  nomod
vin in 0 pl 0v 0n 5v 100ps
.param rpoly=40 wt=100 lt=1.2
m1 single in 0 0 n w=wt l=lt
xm1 lumped  in 0 0 lrgtp xw=wt xl=lt
vsingle single 0 5v
vlumped lumped 0 5v
.tran 25ps 4ns
.graph tran model=time1 single=par('-i(vsingle) ‘) lumped=par('-i(vlumped)')
.model time1 plot xmin=0ps xmax=800ps
* subckt and model on next page
*.subckt lrgtp drain gate source bulk
```

Example of SPICE Deck, Cont'd

```
.subckt lrgtp drain gate source bulk
m1 drain gate source bulk n w='xw/18' l=xl
m2 drain g1 source bulk n w='xw/9' l=xl
m3 drain g2 source bulk n w='xw/9' l=xl
m4 drain g3 source bulk n w='xw/9' l=xl
m5 drain g4 source bulk n w='xw/9' l=xl
m6 drain g5 source bulk n w='xw/9' l=xl
m7 drain g6 source bulk n w='xw/9' l=xl
m8 drain g7 source bulk n w='xw/9' l=xl
m9 drain g8 source bulk n w='xw/9' l=xl
m10 drain g9 source bulk n w='xw/18' l=xl
```

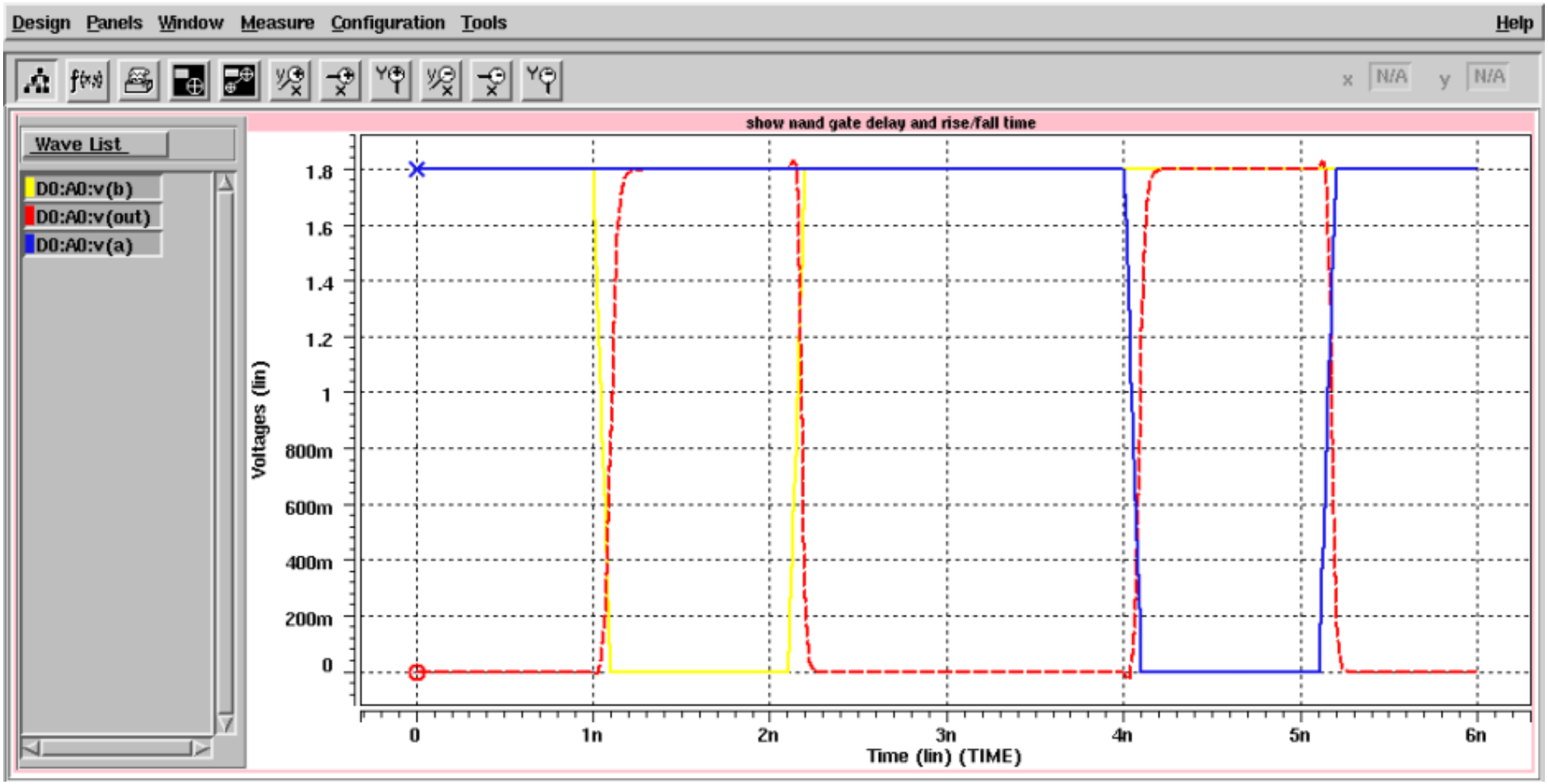
Example of SPICE Deck, Cont'd

```
r1 gate g1 'xw/xl*rpoly/9'
r2 g1 g2 'xw/xl*rpoly/9'
r3 g2 g3 'xw/xl*rpoly/9'
r4 g3 g4 'xw/xl*rpoly/9'
r5 g4 g5 'xw/xl*rpoly/9'
r6 g5 g6 'xw/xl*rpoly/9'
r7 g6 g7 'xw/xl*rpoly/9'
r8 g7 g8 'xw/xl*rpoly/9'
r9 g8 g9 'xw/xl*rpoly/9'
.ends lrgtp
* *  model section *
.model n nmos level=3 vto=0.7 uo=500 kappa=.25 kp=30u
eta=.03 theta=.04 +vmax=2e5 nsub=9e16 tox=250e-10
gamma=1.5 pb=0.6 js=.1m xj=0.5u ld=0.0
+nfs=1e11 nss=2e10 cgso=200p cgdo=200p cgbo=300p
.end
```

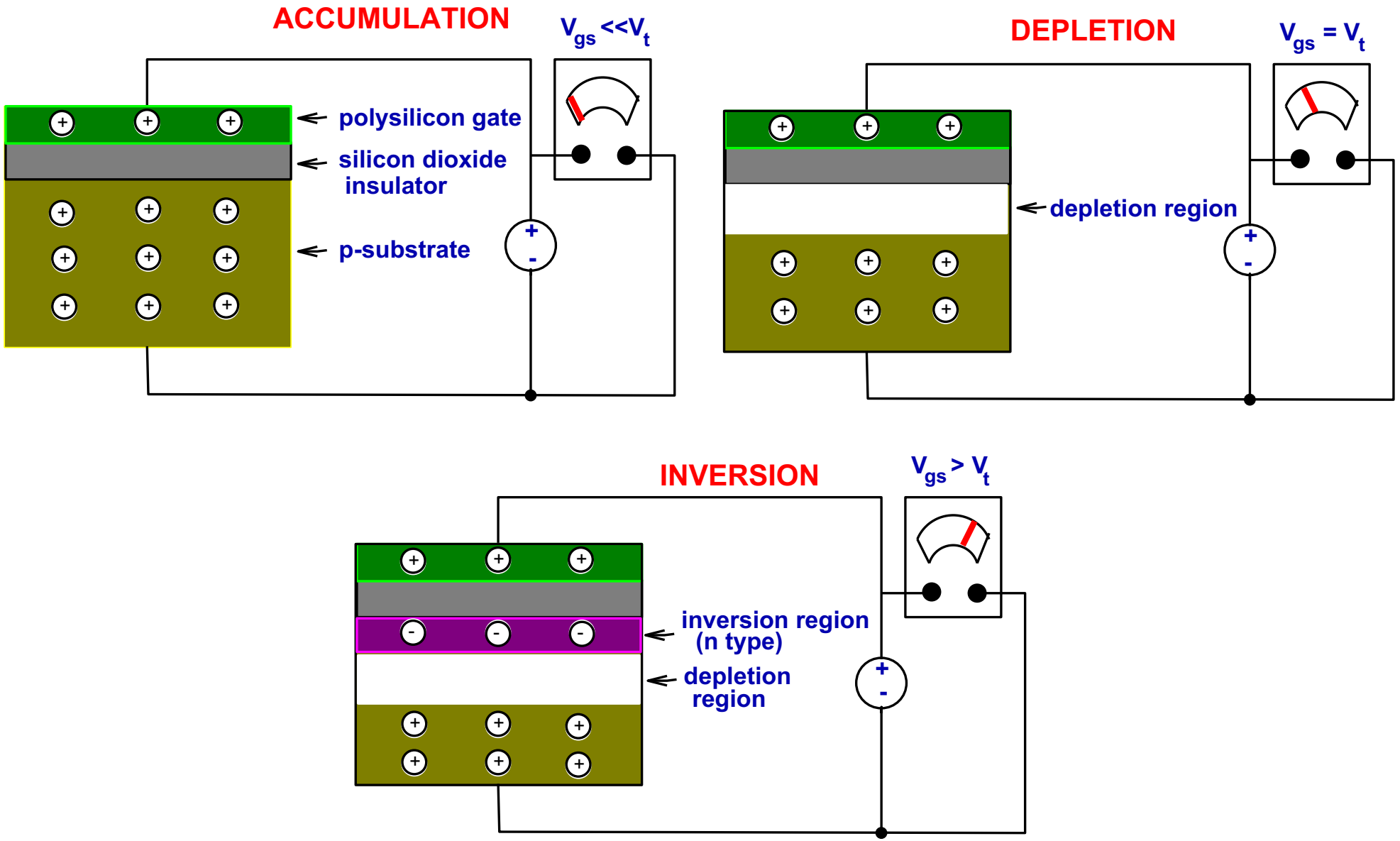
Spice Simulation: NAND Gate

SPICE "deck" has "measure", print statements; parameters, netlist

Model: 0.18 micron



Modes in MOS Structures



Greek alphabet