EN1600
Design and Implementation of VLSI Systems
Fall 2016

Lecture 4 & 5: The MOS Transistor

Reading: Chapter 2, Weste & Harris

September 19, 21, 2016

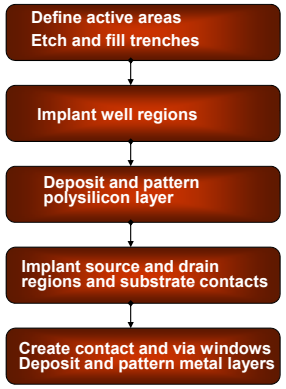
Prof. R. Iris Bahar

BROWN

Cadence tool suite tutorial

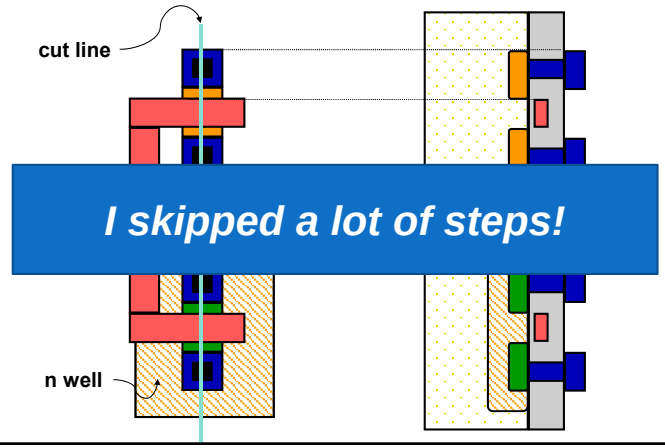
- The full tutorial can be found online on the course webpage:
  - Click on "Assignments"
- Please have your completed layout ready for Marc to view by the end of Wednesday, September 21
  - Instructions are on the webpage
- Completing the entire tutorial will put you in good shape for the first homework assignment.

CMOS process at a glance

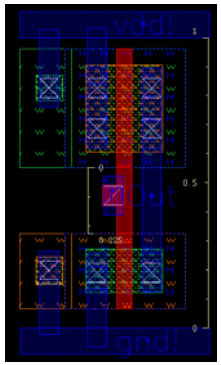


- One full photolithography sequence per layer (mask)
- Built (roughly) from the bottom up
  - 5 metal 2
  - 4 metal 1
  - 2 polysilicon
  - 3 source and drain diffusions
  - 1 tubs (aka wells, active areas)

Final Product: CMOS inverter



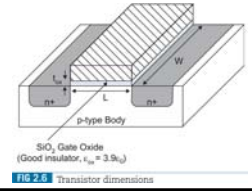
Final layout of your inverter



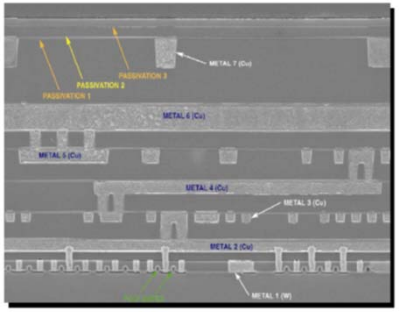
- Active region
- Pimplant, nwell
- Source/drain contacts
- VtH implant
- Nimplant, pwell
- Body contacts
- Input/output connections

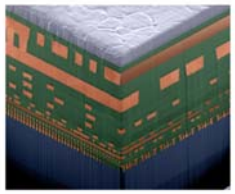
Feature size (λ vs. absolute dimensions)

- Feature size: minimum distance between source and drain of transistor
- Using λ rules:
  - If λ=45nm and feature size = 2λ, feature size 90nm
- Note that we are using absolute dimensions for our layout.
  - 45nm library
  - Minimum feature size = 50nm
    - $L_{effective} \approx 45nm$

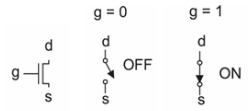


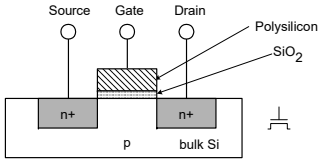
The inside of an integrated circuit





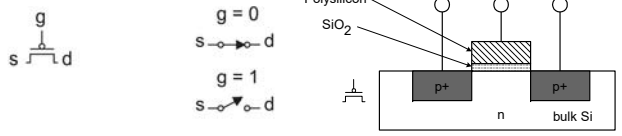
nMOS transistor





- g=0: gate is at low voltage ($V_{gs}<V_{tn}$)
  - p-type body is at low voltage
  - source and drain junction diodes are OFF
  - transistor is OFF (no current flows)
- g=1: gate is at a high voltage ($V_{gs} \geq V_{tn}$)
  - negative charge attracted to body
  - inverts a channel under gate to n-type
  - transistor is ON (current flows)
  - transistor acts as resistor

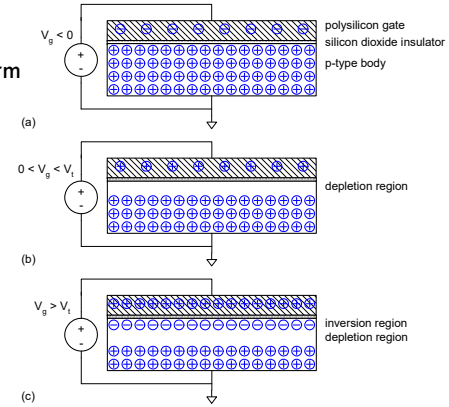
pMOS transistor



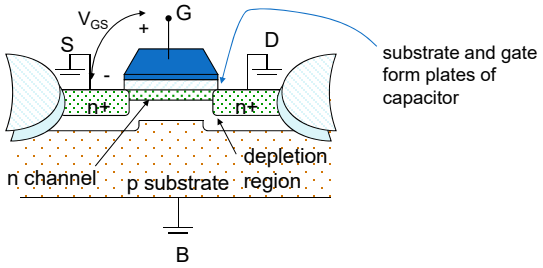
- g=0: gate is at low voltage ($|V_{gs}| > |V_{tp}|$)
  - positive charge attracted to body
  - inverts channel under gate to p-type
  - transistor is ON (current flows)
- g=1: gate is at a high voltage ($|V_{gs}| \leq |V_{tp}|$)
  - n-type body is at high voltage
  - Source and drain junctions are OFF
  - transistor is OFF (no current flows)

MOS Capacitor

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



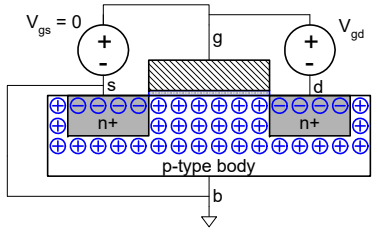
Threshold voltage concept



- Depletion region: area devoid of mobile carriers (holes)
- Inversion layer: n-channel region under oxide
- The value of V_{GS} where strong inversion occurs is called the threshold voltage, V_T

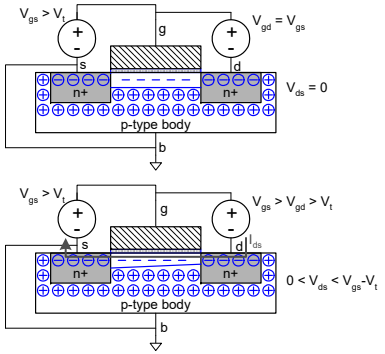
nMOS Cutoff

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



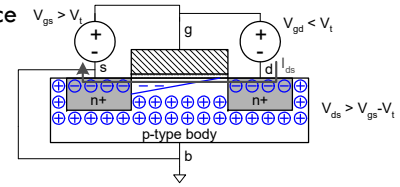
nMOS Linear

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



nMOS saturation

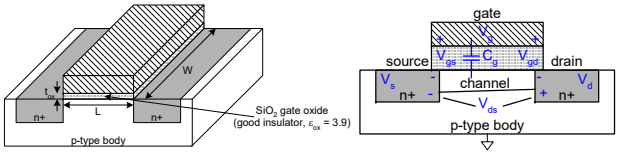
- Channel pinches off
- I_{ds} independent of V_{ds}
- We say current saturates
- Similar to current source



Channel charge

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

$C_{ox} = \varepsilon_{ox} / t_{ox}$



Crossing the channel

- Now we know
  - How much charge $Q_{channel}$ is in the channel
- How much time t does it take to cross channel?
  - $t = \frac{L}{v} = \frac{L}{\mu E}$, where μ is mobility, E is electric field across the channel, or $E = V_{ds}/L$
  - So now, $t = \frac{L}{\mu \frac{V_{ds}}{L}} = \frac{L^2}{\mu V_{ds}}$
  - for NFETs: μ_n = 500 cm^2/V-sec
  - for PFETs: μ_p = 180 cm^2/V-sec

nMOS Linear I-V

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

$$I_{ds} = \frac{Q_{channel}}{t}$$
$$= \mu C_{ox} \frac{W}{L}\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds}$$
$$= \beta\left(V_{gs} - V_t - V_{ds}/2\right)V_{ds} \qquad \beta = \mu C_{ox}\frac{W}{L}$$

nMOS Saturation I-V

- When $V_{ds} > V_{dsat} = V_{gs} - Vt$
  - channel is no longer inverted around the drain;
  - channel pinches off near drain
- Increasing V_{ds} has no further effect on 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

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

I-V relation: saturation mode

For long channel devices

- When $V_{DS} \geq V_{GS} - V_T$

Substitute (Vgs-Vt) for Vds from eqn. in linear mode since channel voltage remains fixed

$$I'_D = \frac{\mu C_{ox}}{2}\frac{W}{L}(V_{GS} - V_T)^2$$

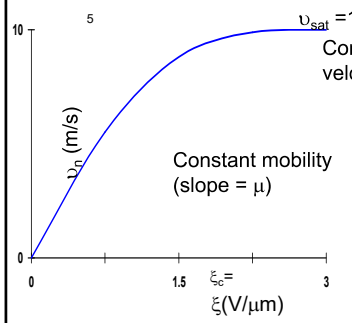
- However, the effective length of the conductive channel reduces with further increases in V_{DS}, so

$$I_D = I'_D(1 + \lambda V_{DS})$$

where λ = channel-length modulation

Short Channel Effects

- Behavior of short channel device mainly due to



- Velocity saturation
  velocity of the carriers saturates due to scattering (collisions of carriers)

- For an NMOS device with L of .25μm, only a couple of volts difference between D and S are needed to reach velocity saturation

Velocity Saturation

For short channel devices

- Linear: When $V_{DS} \le V_{GS} - V_T$

$$I_D = \kappa(V_{DS})\mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}\right]$$

where

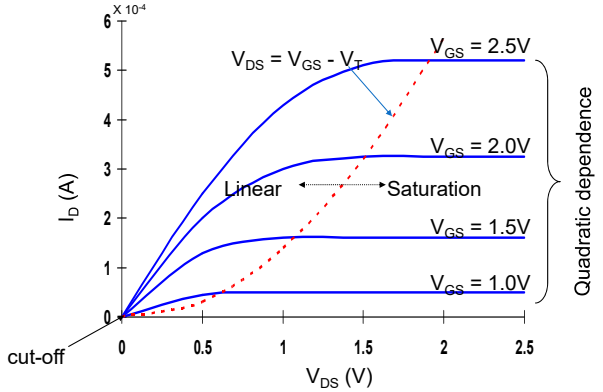
$\kappa(V) = 1/(1 + (V/\xi_c L))$, degree of velocity saturation

- Saturation: When $V_{DS} \ge V_{DSAT} = \kappa(V_{GS} - V_T)(V_{GS} - V_T)$
  That is, $V_{DSAT} < V_{GS} - V_T$ so the device reaches saturation sooner compared to long-channel devices.

Circle the correct answer…

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

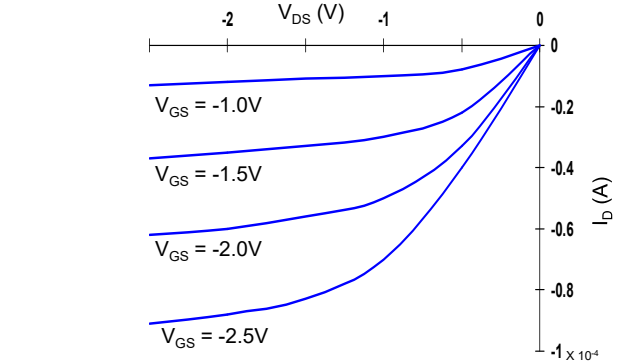
Long Channel I-V Plot (NMOS)



NMOS transistor, 0.25um, L_d = 10um, W/L = 1.5, V_{DD} = 2.5V, V_T = 0.4V

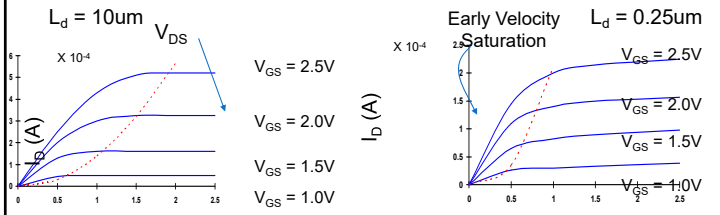
Short channel I-V plot (PMOS)

- All polarities of all voltages and currents are reversed



PMOS transistor, 0.25um, L_d = 0.25um, W/L = 1.5, V_{DD} = 2.5V, V_T = -0.4V

Short channel I-V plot (NMOS)

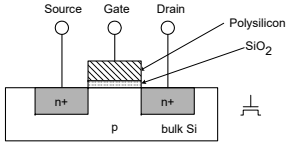


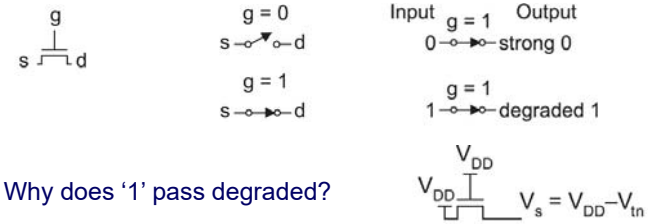
V_{DS} (V)
Quadratic dependence on VGS

V_{DS} (V)
Linear dependence on VGS

- I_{DSAT} has a linear dependence with respect to V_{GS} (as opposed to quadratic) so a reduced amount of current is delivered for a given control voltage

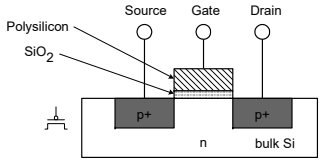
why is nMOS a poor logic "1" device?

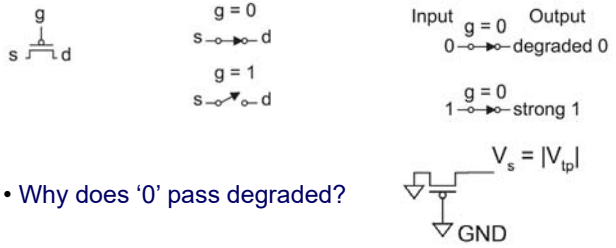




- Why does '1' pass degraded?

pMOS passes logic "0" poorly

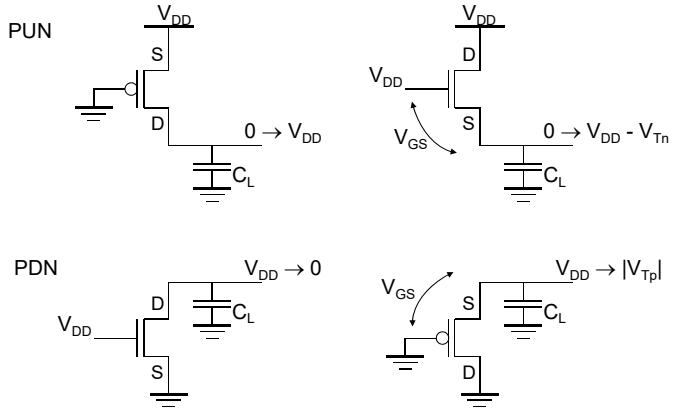




- Why does '0' pass degraded?

Threshold drops

Why PMOS for PUN and NMOS for PDN?



Lab #0 due today

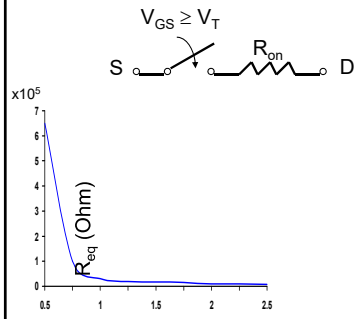
- For now, please send your files to Marc directly for him to review
- Eventually, we are trying to set up private readable directories with CCV
- Due by end of today
- Lab #1 will be available later today

Threshold voltage (body effect)

$$V_T = V_{T0} + \gamma\left(\sqrt{|(-2)\varphi_F + V_{SB}|} - \sqrt{|2\varphi_F|}\right)$$

- V_{T0} is the threshold voltage at $V_{SB} = 0$ and is mostly a function of the manufacturing process
  - V_{SB} is the source-bulk voltage
- $\varphi_F = -\varphi_T \ln(N_A / n_i)$ is the Fermi potential
- $\phi_T = kT/q = 26mV$ at 300K is the thermal voltage; N_A is the acceptor ion concentration; $n_i \approx 1.5x1010$ cm^{-3} at 300K is the intrinsic carrier concentration in pure silicon
- $\gamma = \sqrt{(2q\varepsilon_{si}N_A)/C_{ox}}$ is the body-effect coefficient
- Expresses the impact of changes in V_{SB}.
- ε_{si} is the permittivity of silicon; C_{ox} is the cap. per unit area of the gate oxide

The transistor modeled as a switch

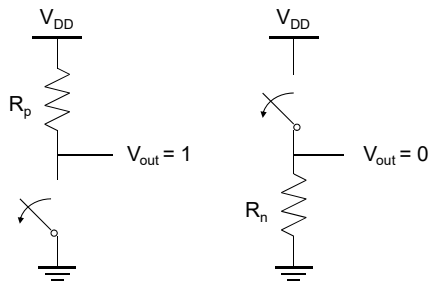


(for $V_{GS} = V_{DD}$, $V_{DS} = V_{DD} \rightarrow V_{DD}/2$)

Model as a switch with infinite off resistance and a finite on resistance, R_{on}

- Resistance inversely proportional to W/L (doubling W halves R_{on})
- For $V_{DD} >> V_T + V_{DSAT}/2$, R_{on} independent of V_{DD}
- Once V_{DD} approaches V_T, R_{on} increases dramatically

CMOS Inverter: steady state response



$V_{in} = 0$ $V_{in} = V_{DD}$

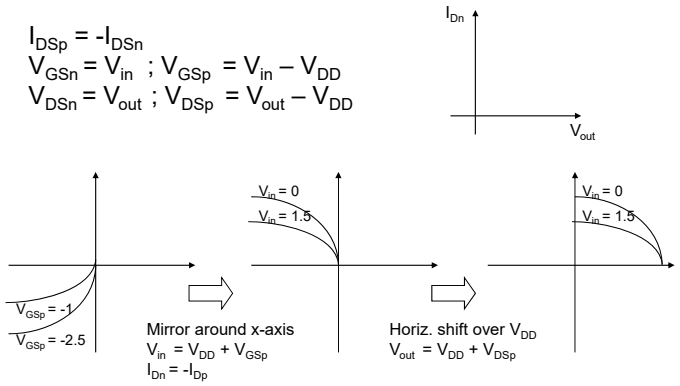
Voltage transfer characteristics

- What happens when input voltage is not "at rail"
- Vin < Vdd, or Vin > Gnd?
- If the transistor is ON, then voltage at output will change, but will not go to rail.

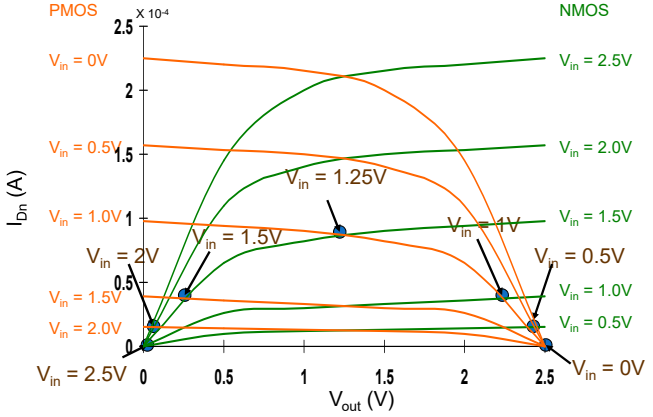
Transforming PMOS I-V lines

- Want common coordinate set V_{in}, V_{out}, and I_{Dn}

$I_{DSp} = -I_{DSn}$
$V_{GSn} = V_{in}$; $V_{GSp} = V_{in} - V_{DD}$
$V_{DSn} = V_{out}$; $V_{DSp} = V_{out} - V_{DD}$

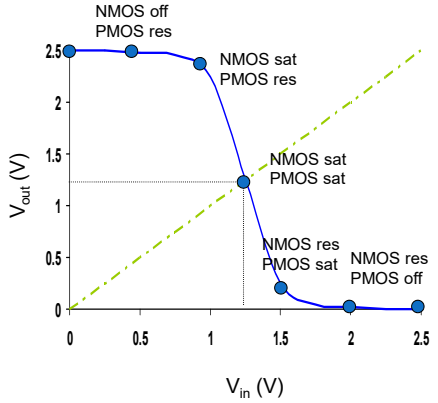


CMOS inverter I-V curves



0.25um, $W/L_n = 1.5$, $W/L_p = 4.5$, $V_{DD} = 2.5V$, $V_{Tn} = 0.4V$, $V_{Tp} = -0.4V$

CMOS inverter VTC



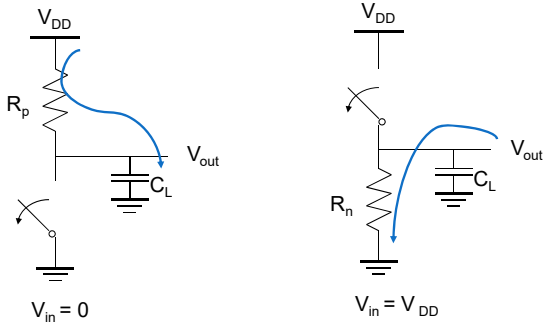
Taking those intersection points from the load curves, we obtain the voltage-transfer characteristic

CMOS properties

- Full rail-to-rail swing ➔ high noise margins
  - Logic levels independent of device sizes ➔ *ratioless*
- Always a path to V_{dd} or GND in steady state ➔ less sensitive to noise
- nearly zero steady-state input current
- No direct path steady-state between power and ground
  - no static power dissipation
- Propagation delay is a function of load capacitance and resistance of transistors

CMOS Inverter:

Switch Model of Dynamic Behavior



- Gate response time is determined by the time to charge C_L through R_p (discharge C_L through R_n)

Switching threshold

- Define V_M to be the point where $V_{in} = V_{out}$ (both PMOS and NMOS in saturation since $V_{DS} = V_{GS}$)
- If $V_M = V_{DD}/2$, then this implies *symmetric rise/fall* behavior for the CMOS gate
- Recall at saturation, $I_D=(k'/2)(W/L)\,(V_{GS}-V_T)^2$,
  - where $k'_n = \mu_n C_{ox} = \mu_n \varepsilon_{ox}/t_{ox}$
- Setting $I_{Dp} = -I_{Dn}$

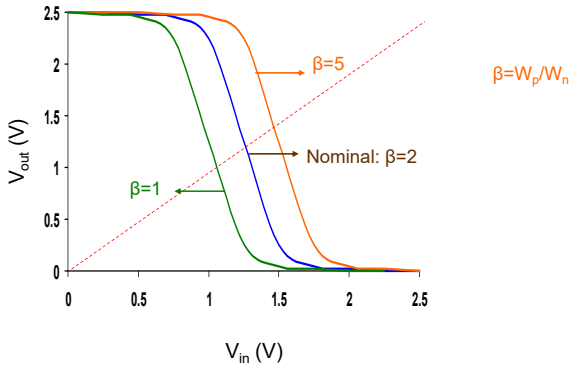
$$\frac{k'_n}{2}\frac{W_n}{L_n}(V_M - V_{Tn})^2 = \frac{k'_p}{2}\frac{W_p}{L_p}(-V_M - V_{Tp})^2$$

- Assuming $V_{Tn}=-V_{Tp}$ $\quad \dfrac{W_p/L_p}{W_n/L_n} = \dfrac{k'_n}{k'_p} = \dfrac{\mu_n}{\mu_p}$

Relative Transistor Sizing

- When designing static CMOS circuits, balance the driving strengths of the transistors by making the PMOS section wider than the NMOS section to
  - maximize the noise margins and
  - obtain symmetrical characteristics

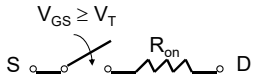
Impact of unmatched drive strengths



- Skewing the β ratio will shift the switching threshold

MOS Structure Resistance

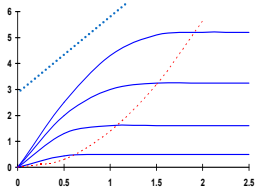
- The simplest model assumes the transistor is a switch with an infinite "off" resistance and a finite "on" resistance R_{on}



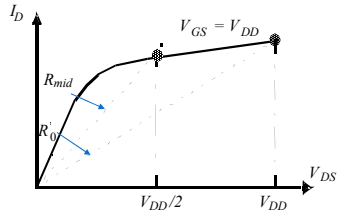
- However R_{on} is nonlinear, time-varying, and dependent on the operation point of the transistor
- *How can we determine an equivalent (constant and linear) resistance to use instead?*

MOS structure resistance

- Approximate R_{on} as the resistance found during linear operation
  - Simple to calculate but limited accuracy
- Instead use the average value of the resistances, R_{eq}, at the end-points of the transition (i.e., V_{DD} and $V_{DD}/2$)



$$R_{eq} \approx \frac{1}{2}(R_{on}(t_1) + R_{on}(t_2))$$
$$\approx \frac{1}{2}\left(\frac{V_{DD}}{I_{DSAT}} + \frac{V_{DD}/2}{I_{DSAT}}\right)$$
$$\approx \frac{3}{4}\frac{V_{DD}}{I_{DSAT}}$$



BROWN

Equivalent MOS Structure Resistance

$$R_{eq} = \frac{3}{4}\frac{V_{DD}}{I_{DSAT}}$$

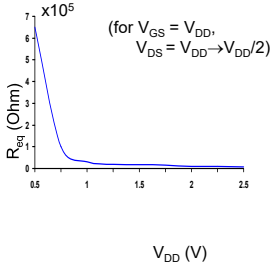
where

$$I_{DSAT} = \frac{\mu C_{ox}}{2}\frac{W}{L}\left((V_{DD} - V_T)V_{DSAT} - \frac{V_{DSAT}^2}{2}\right)$$

so,

$$R_{eq} = \frac{3}{2\mu C_{ox}}\frac{L}{W}\frac{V_{DD}}{(V_{DD} - V_T)V_{DSAT} - \frac{V_{DSAT}^2}{2}}$$

$$R_{eq_{NMOS}} \propto \frac{L}{\mu_n W}, \qquad R_{eq_{PMOS}} \propto \frac{L}{\mu_p W}$$



R_{eq} is essentially independent of V_{DD} as long as $V_{DD} >> V_T + V_{DSAT}/2$