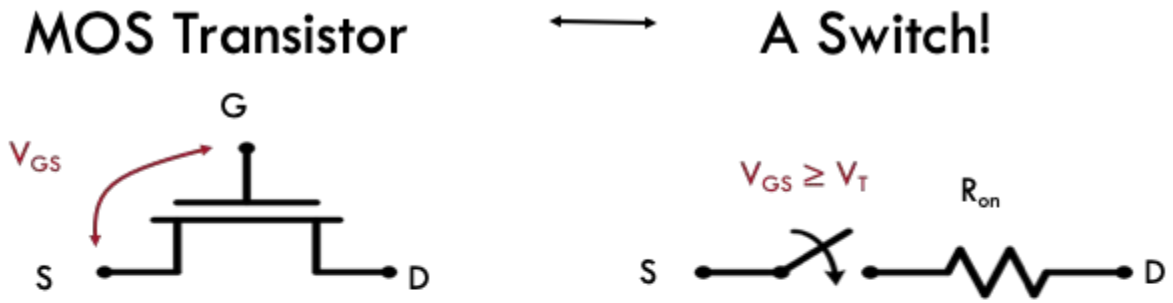
Discussion 9

MOS Switch, VTCs, CMOS Gates, Sizing, RC Delay, Logical Effort

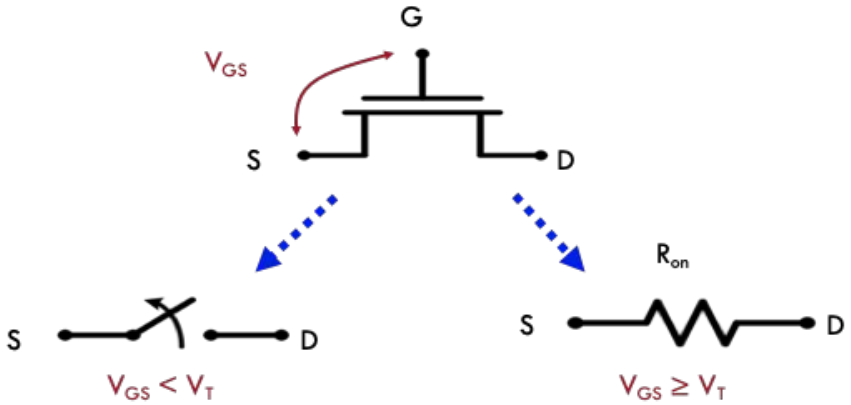
MOS Switch



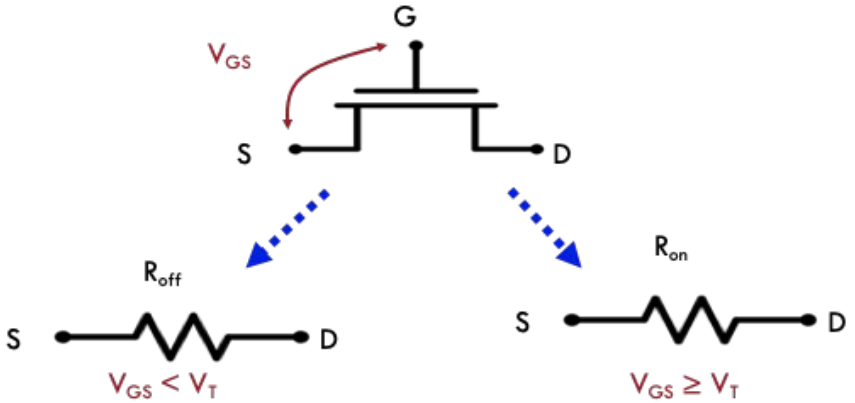
- V_{GS} controls the switch
  - (it also charges the channel capacitor)

MOS Switch

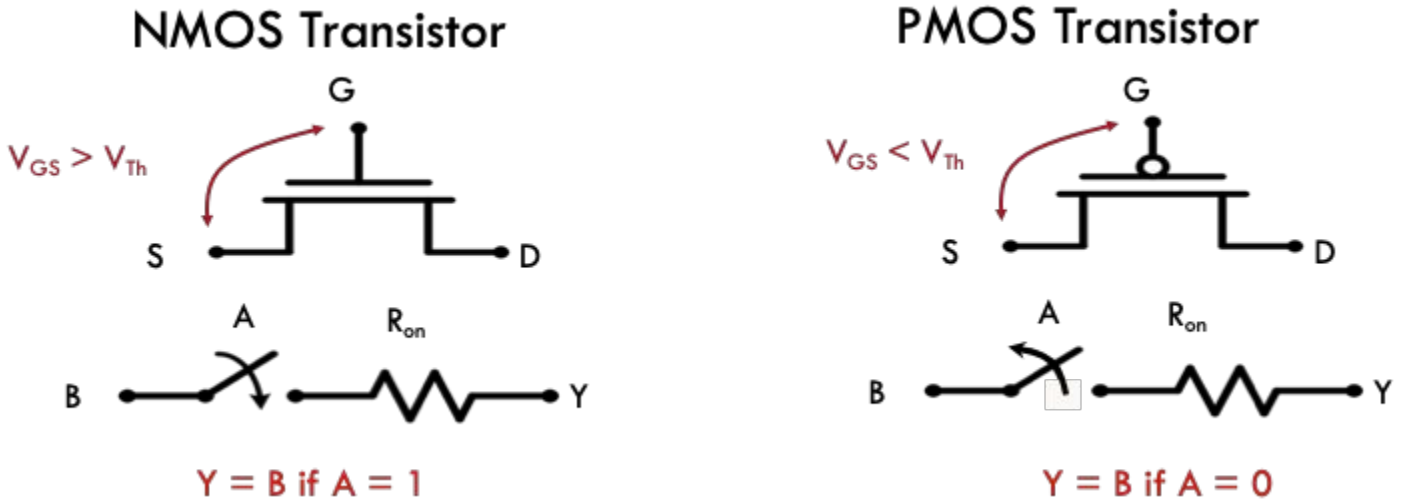
ON/OFF Model



“More Realistic” Model

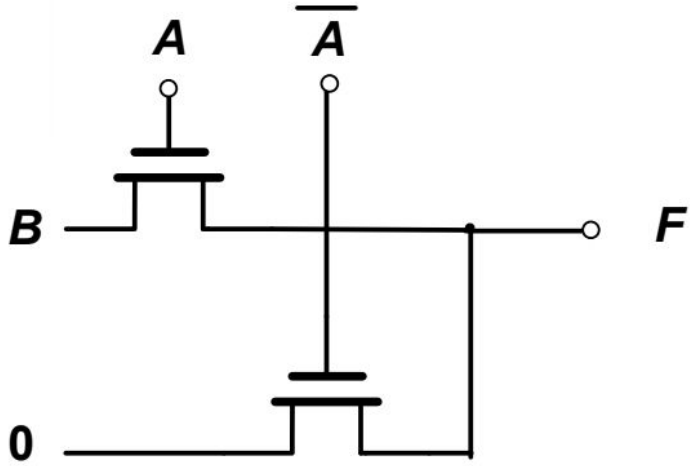


MOS Switch

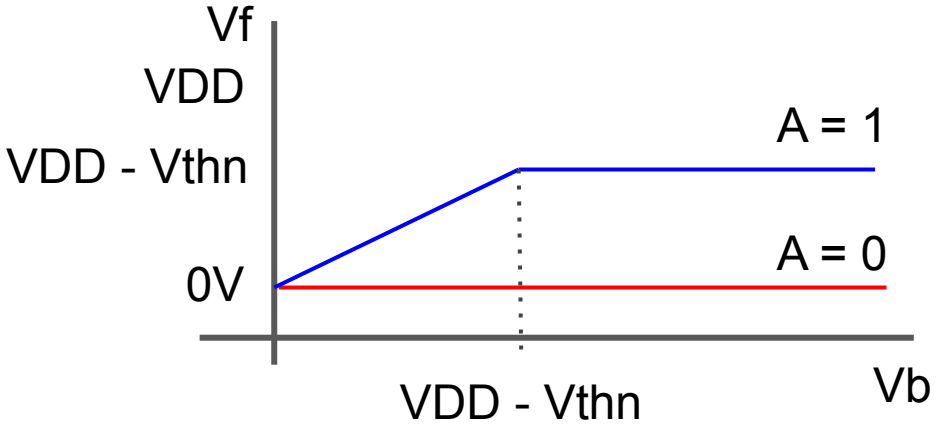


- Source of NMOS always at lower voltage
- Source of PMOS always at higher voltage
- The ‘effective’ source node can change depending on the voltage at the MOS’ terminals

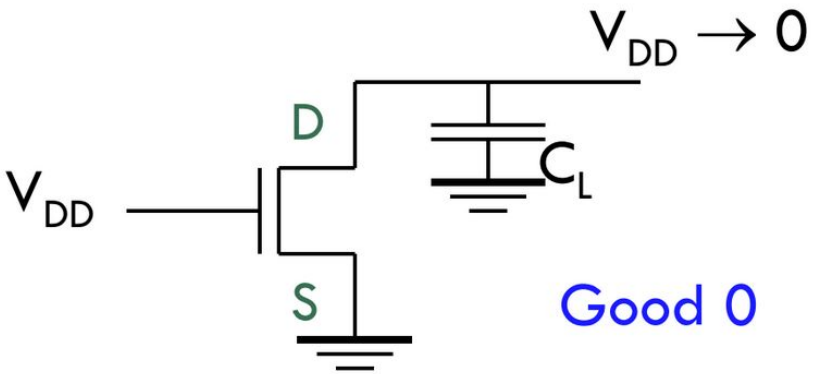
AND Switch Gate Analysis

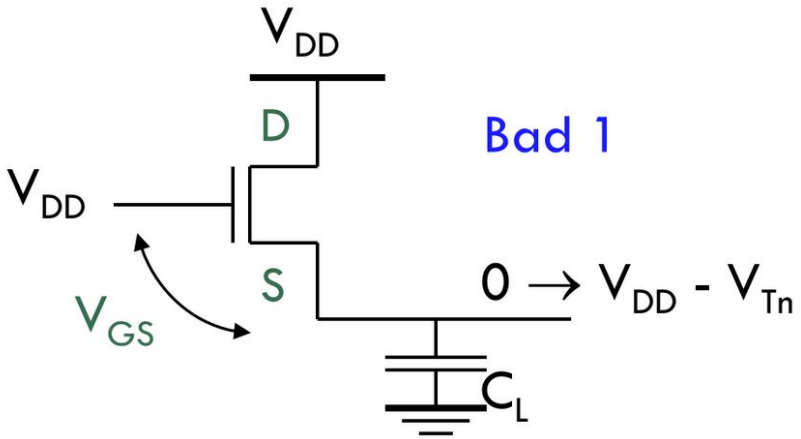


- F = AB
- Assume CL starts at 0 V
- Draw the VTC for both cases where A = 0 and A = 1



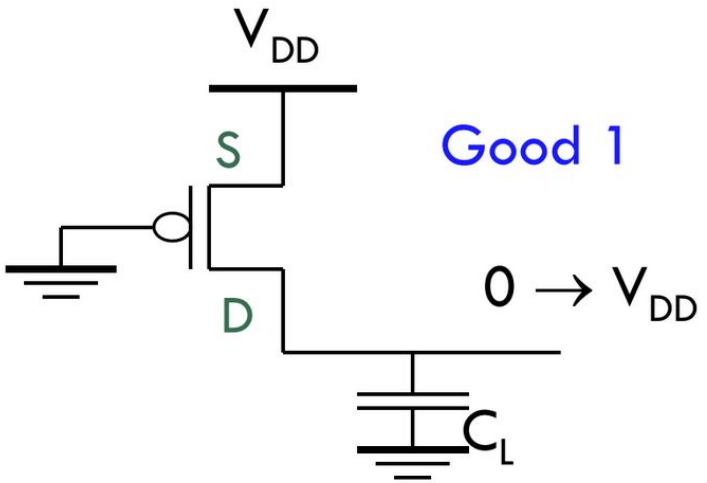
NMOS VTC

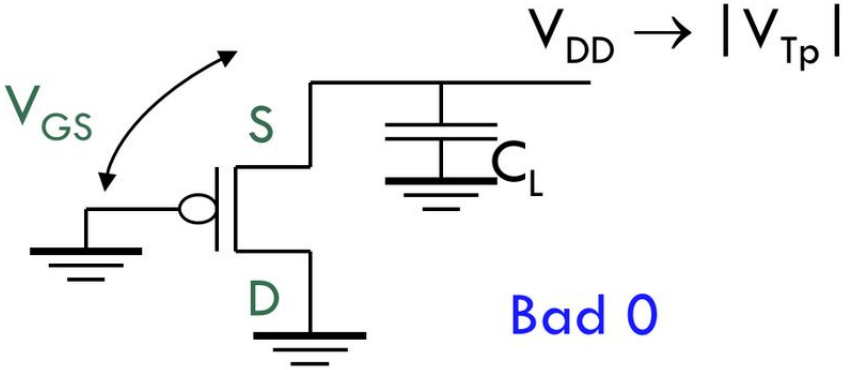




- NMOS can pull a ‘strong’ 0
- NMOS can only pull a ‘weak’ 1
  - Can only go up to Vg - Vthn
- Remember NMOS source always at lower voltage

PMOS VTC

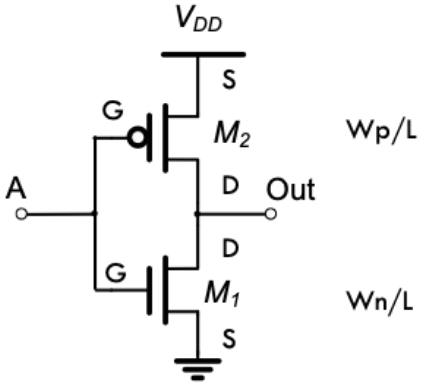


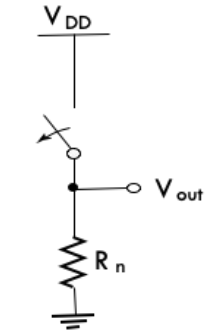


- PMOS can pull a ‘strong’ 1
- PMOS can only pull a ‘weak’ 0
  - Can only go down to Vg + |Vthp|
  - Remember PMOS source always at higher voltage

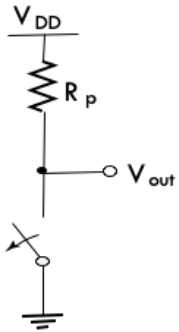
CMOS Inverter

- Schematic



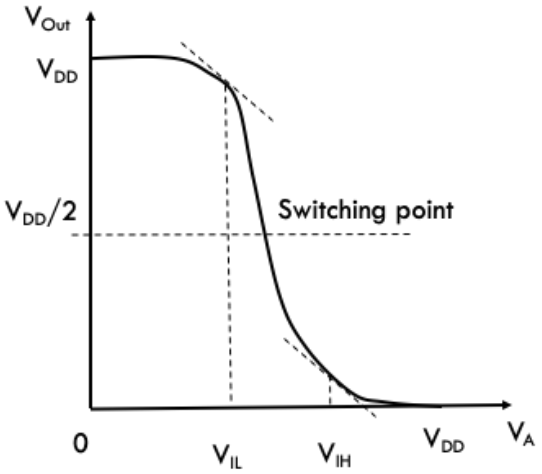


$V_{in} = V_{DD}$



$V_{in} = 0$

$V_{OL} = 0$
$V_{OH} = V_{DD}$

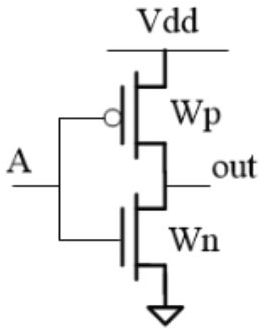


Inverter VTC

[7pts/11pts] Voltage Transfer Characteristics (VTCs).

The technology has the following parameters: $V_{th,N} = 0.2V$ and $|V_{th,P}| = 0.3V$, $R_n = 2k\Omega * \mu m$, $R_p = 3k\Omega * \mu m$ at $V_{dd} = 1V$. Draw the voltage transfer characteristic (V_{out} vs V_A) of the gates below with $W_p = W_n = 1\mu m$.

(a) [4pts] Draw the VTC and determine V_{OL}, V_{IL}, V_{OH}, V_{IH} and noise margins NM_H and NM_L.



Remember:

Noise margin high:
$NM_H = V_{OH} - V_{IH}$

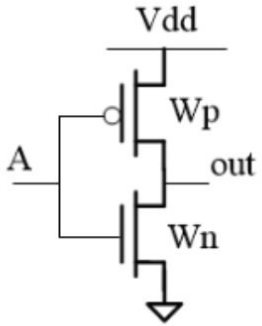
Noise margin low:
$NM_L = V_{IL} - V_{OL}$

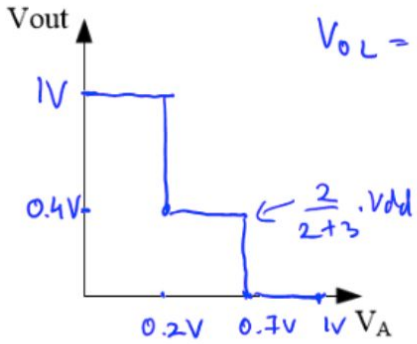
Inverter VTC

[7pts/11pts] Voltage Transfer Characteristics (VTCs).

The technology has the following parameters: $V_{th,N} = 0.2V$ and $|V_{th,P}| = 0.3V$, $R_n = 2k\Omega * \mu m$, $R_p = 3k\Omega * \mu m$ at $V_{dd} = 1V$. Draw the voltage transfer characteristic (V_{out} vs V_A) of the gates below with $W_p = W_n = 1\mu m$.

(a) [4pts] Draw the VTC and determine V_{OL}, V_{IL}, V_{OH}, V_{IH} and noise margins NM_H and NM_L.





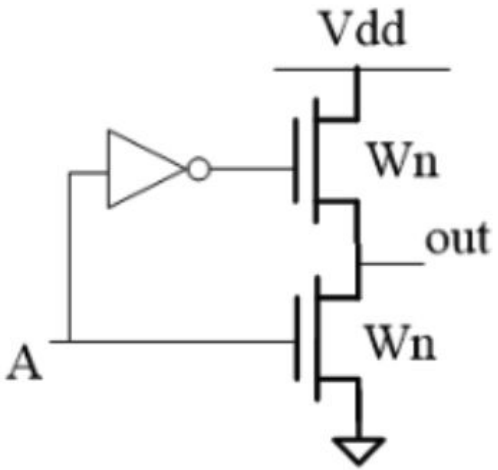
$V_{OH} = 1V$ $V_{IH} = 0.7V$

$V_{OL} = 0V$ $V_{IL} = 0.2V$

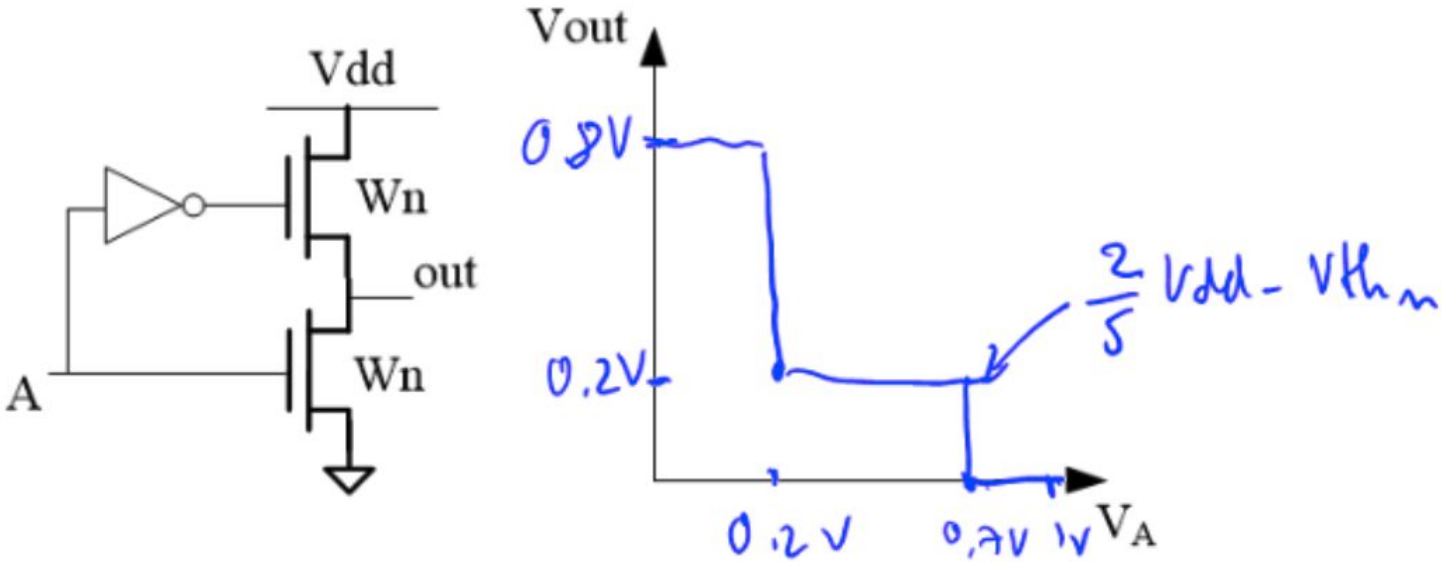
$NM_H = 0.3V$

$NM_L = 0.2V$

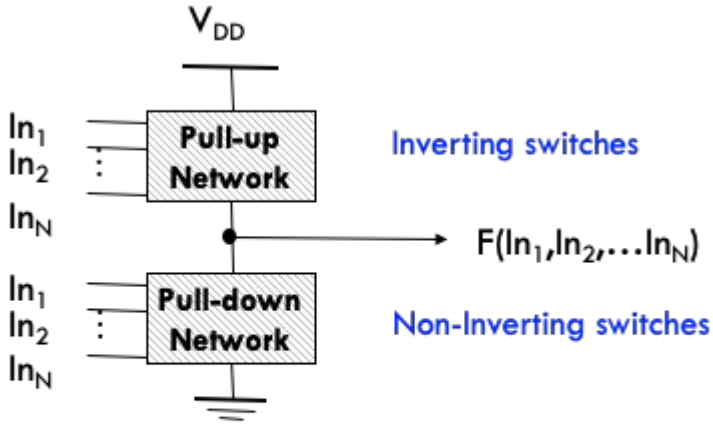
Inverter VTC



Inverter VTC

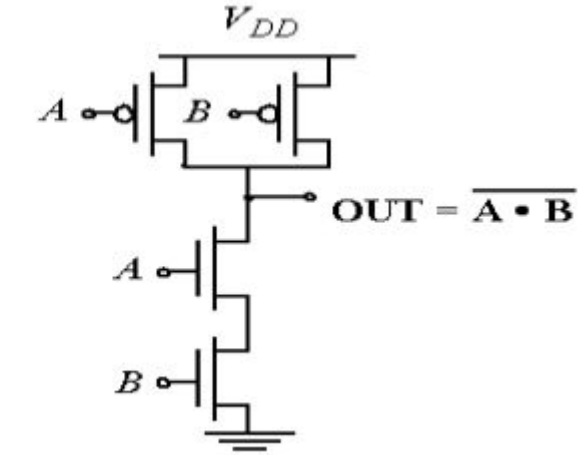


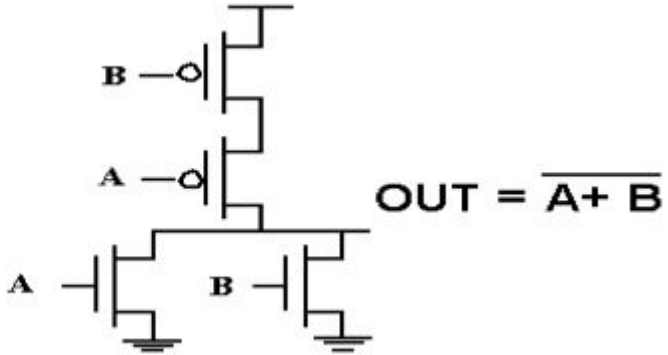
CMOS Logic



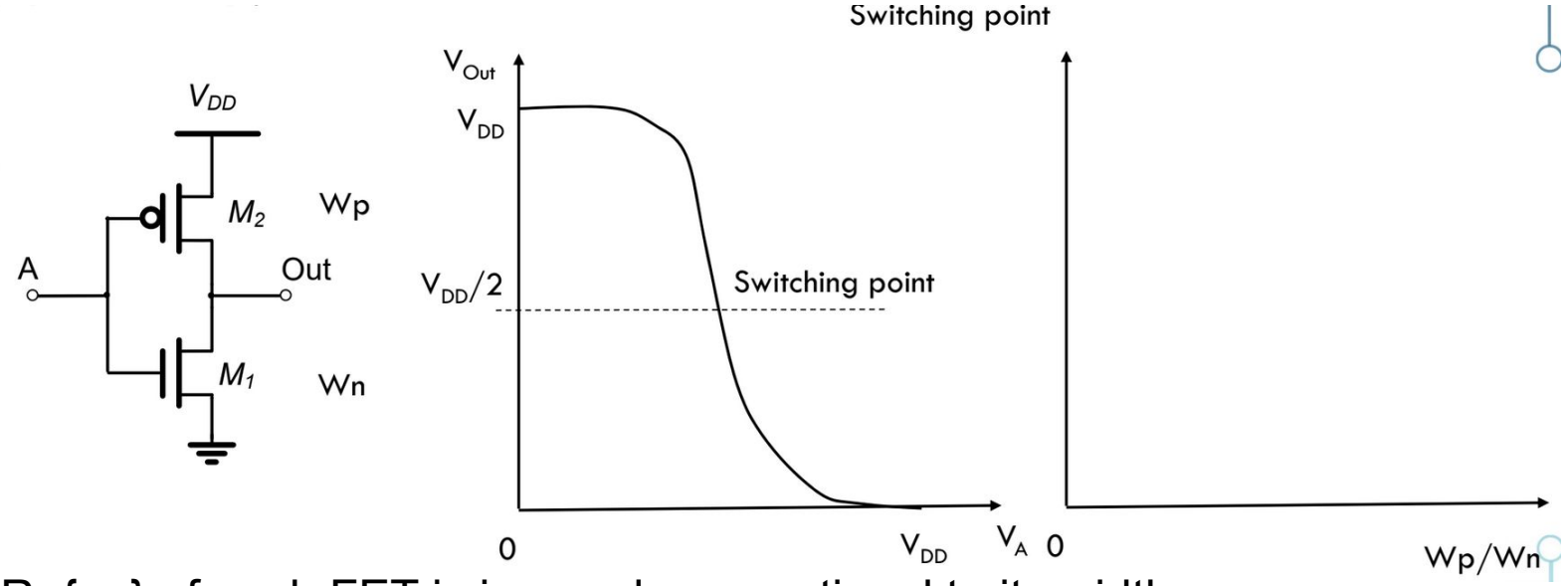
PUN and PDN are dual logic networks

PUN and PDN functions are complementary



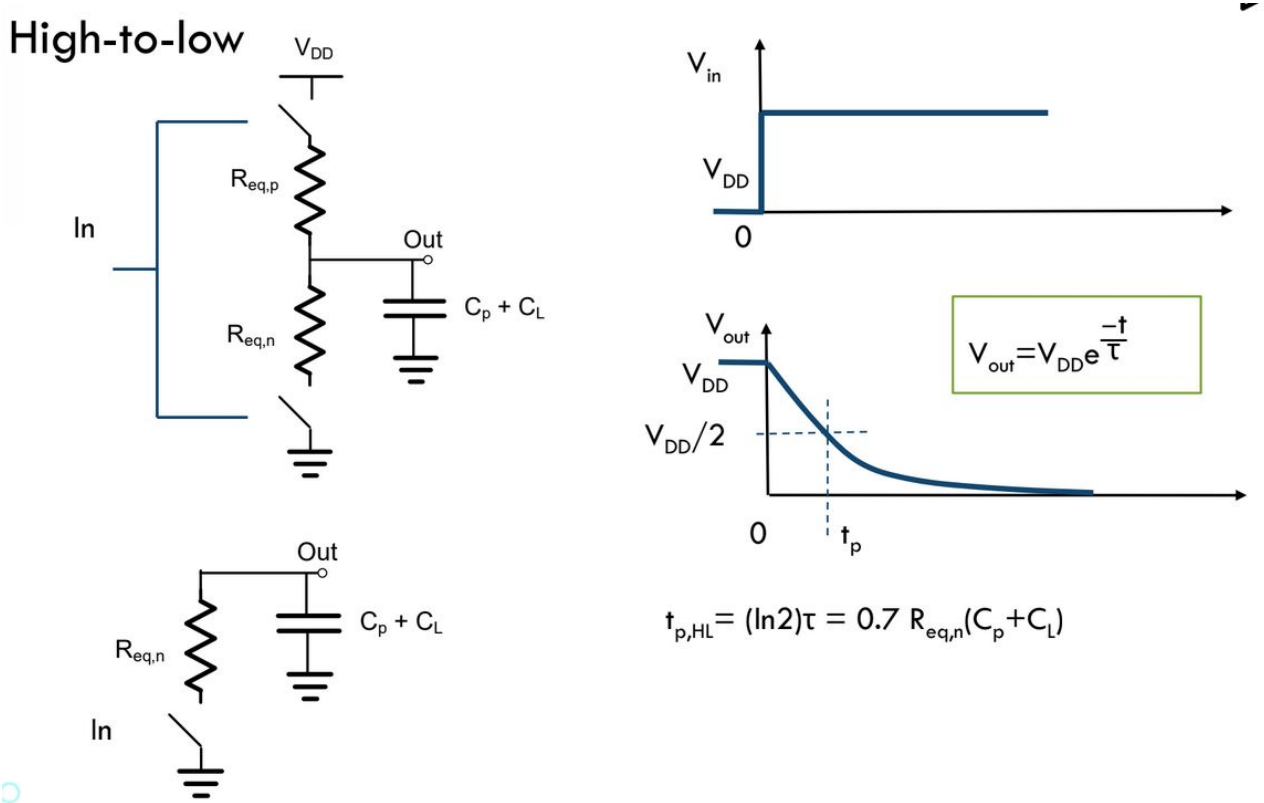


Inverter Sizing



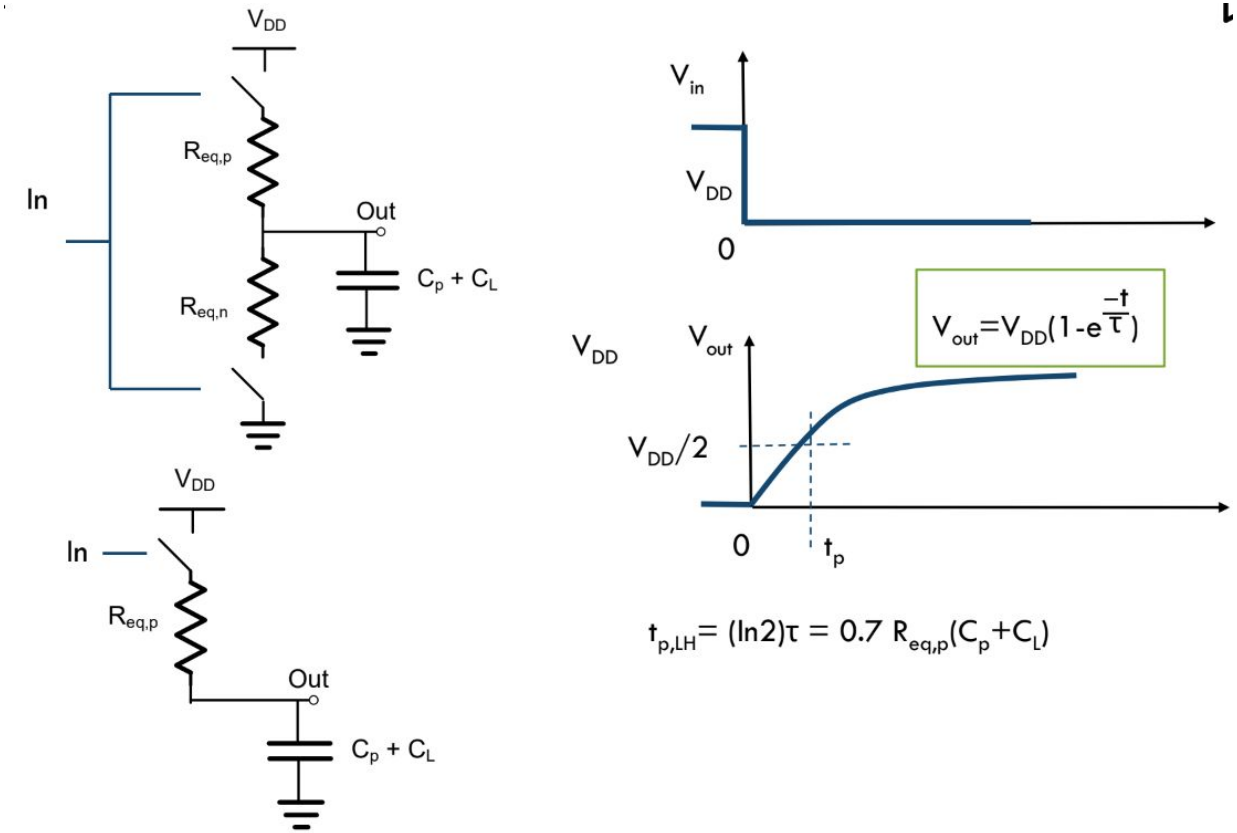
- R_{on} of each FET is inversely proportional to its width
- Assume V_{th,n} = V_{th,p}
- Assume R_{on,n} = R_{on,p} for the same width unless otherwise specified
- Wp = Wn = switching threshold of Vdd/2
- If Wp >> Wn, Vm approaches VDD | If Wn >> Wp, Vm approaches 0 (exponentially)

Inverter RC Delay (High -> Low)



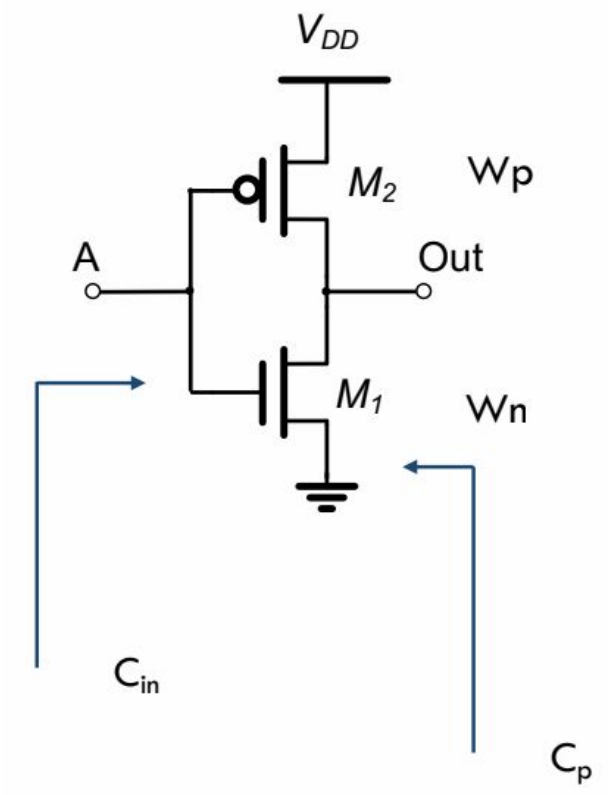
- High to low output transition time governed by strength of NMOS pulling low
- Can lower the time constant by increasing the NMOS width

Inverter RC Delay (Low -> High)



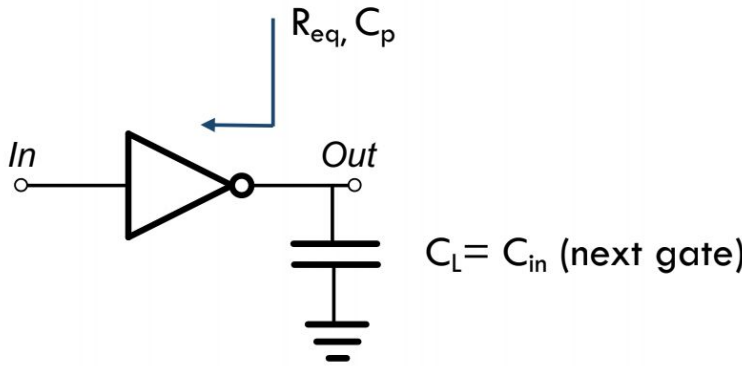
- Low to high output transition time governed by strength of PMOS pulling high
- Can lower the time constant by increasing the PMOS width

Inverter Sizing



- Cin (gate input capacitance) and Cp (intrinsic drain capacitance) are proportional to W
  - You may see Cp referred to Cd in other sources
  - Cp = γ Cin
- How does the inverter delay change if either PMOS/NMOS width is doubled?
- If both widths are doubled, does the intrinsic (unloaded) delay improve?
- Inverters are usually sized to equalize high->low and low->high delays

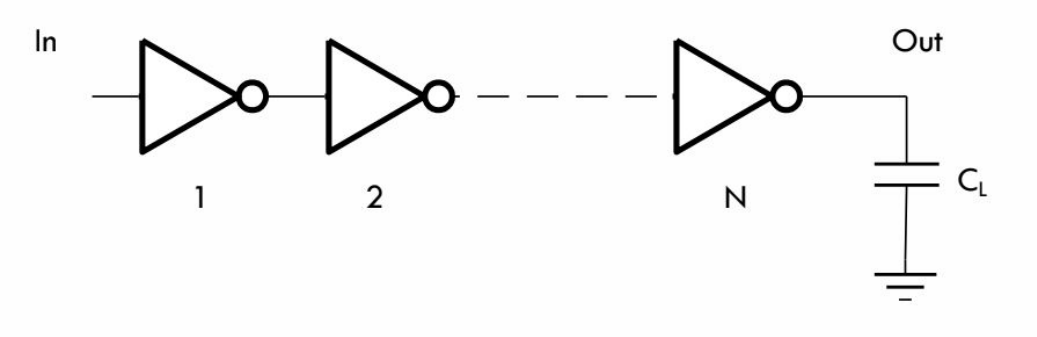
Inverter Delay



$$
\begin{aligned}
t_{p,inv} &= R_{eq} (C_{p,tot} + C_L) \\
&= R_{eq} C_{p,tot} (1 + (C_L / C_{p,tot})) \\
&= R_{eq} C_{p,tot} (1 + (C_L / C_{in} \gamma)) \\
&= R_{eq} C_{p,tot} (1 + (f / \gamma)) \\
&= \tau_{inv} (1 + (f / \gamma))
\end{aligned}
$$

- Delay can be split into 2 parts, intrinsic and extrinsic
- Fanout f = ratio between output and input cap (CL / Cin)
- This delay formula can be generalized for any CMOS gate
  - $t_{p,gate} = \tau_{inv} (p + gf/\gamma)$
  - Often assume $\gamma = 1$

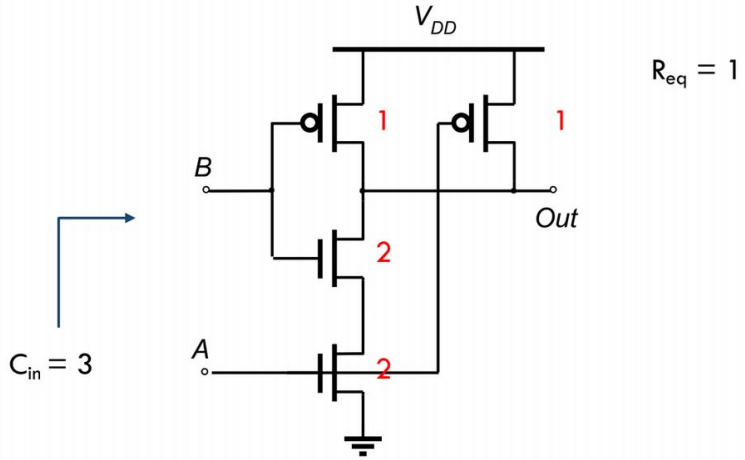
Sizing an Inverter Chain



- We want to minimize the delay of this inverter chain
- Assume the 1st inverter has a size of 1

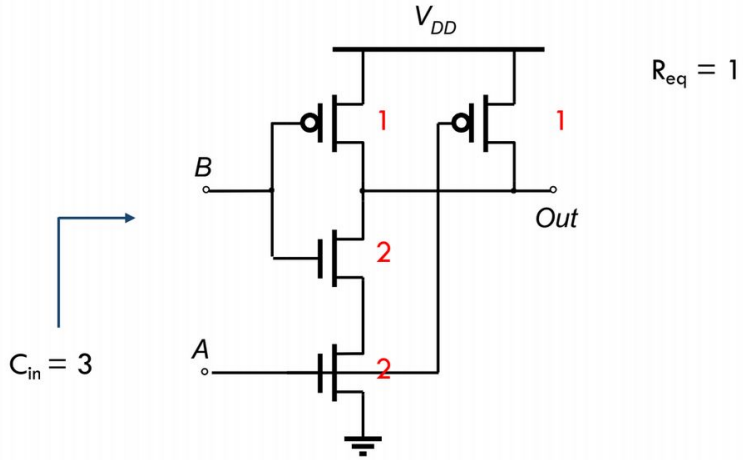
Delay = tp1 + … + tpn = (1+f1) + (1+f2) + … + (1+fn)
Take partial derivatives of this expression wrt C2, … Cn
Solution is f1 = f2 = f3 = … = fN
Each inverter has a fanout = the n-th root of F (total fanout)

Logical Effort



- g = logical effort
- To find g for a CMOS gate:
  1. Size the gate to have R_{eq} equal to that of an inverter
  2. Find the input capacitance of the gate
  3. Take the ratio C_{in,gate} / C_{in,inv}
- For a NAND2 gate: g = 3/2
- For a NOR2 gate: g = 3/2
- How do these change if the PMOS is half as strong as an nMOS?

Intrinsic Delay

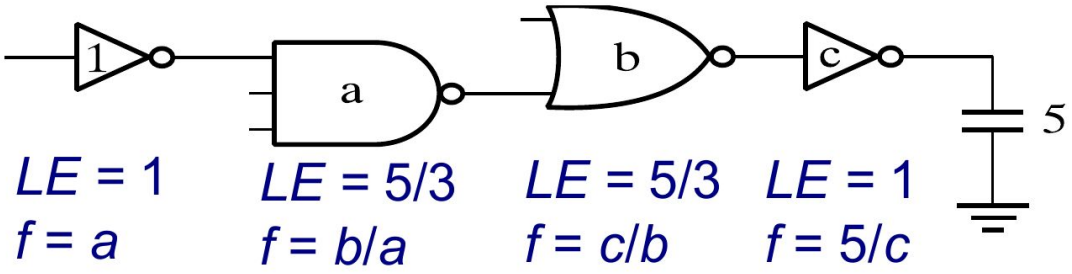


- p = intrinsic delay
- To find p for a CMOS gate:
  1. Size the gate to have R_{eq} equal to that of an inverter
  2. Find the output (intrinsic) capacitance of the gate
  3. Take the ratio C_{out,gate} / C_{out,inv}
- For a NAND2 gate: g = 4/2
- For a NOR2 gate: g = 4/2
- How do these change if the PMOS is half as strong as an nMOS?

Path Delay

- Path logical effort G = g1 * g2 * … gN
- Path fanout F = CL / C_{in,1}
- Branching factor bi = ratio of total cap seen / on-path cap
- Path total effort H = GFB
- We want to minimize the delay of a path through a series of gates
  - Each stage i in the chain has an effective fanout gi * fi
  - The solution is to make the effective fanout of each stage the same
  - EF_{opt} = H^{1/N}
  - EF = gi * fi = gi * (C_{load} / C_{in})
  - C_{in} (i.e. the gate size) = gi * C_{load} / EF
- The optimal path delay D = N * H^{1/N} + p1 + … + pN

Sizing a Logic Path for Minimum Delay



Electrical fanout, F = 5
Π *LE* = 25/9
PE = 125/9
EF/stage = 1.93
a = 1.93
b = 2.23
c = 2.59

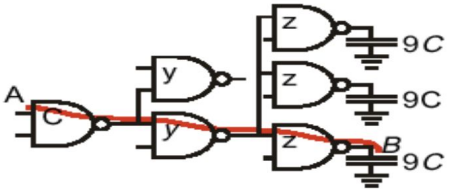
From the back

5/c = 1.93
(5/3)c/b = 1.93
(5/3)b/a = 1.93

- This example assumes the NMOS is twice as strong as the PMOS
- EF/stage is calculated as the 4th-root of G*F
- gi * fi = EF of stage i
- Easier to calculate gate sizes from the load cap backwards

Sizing a Logic Path for Minimum Delay (Branching)

Select gate sizes *y* and *z* to minimize delay from *A* to *B*



Logical Effort: $LE = (4/3)^3$

Electrical Fanout: $F = C_{out}/C_{in} = 9$

Branching Effort: $B = 2 \cdot 3 = 6$

Path Effort: $PE = \prod LE \cdot F \cdot B = 128$

Best Effective Fanout: $EF = PE^{1/3} \approx 5$

Delay: $D = 3 \cdot 5 + 3 \cdot 2 = 21$

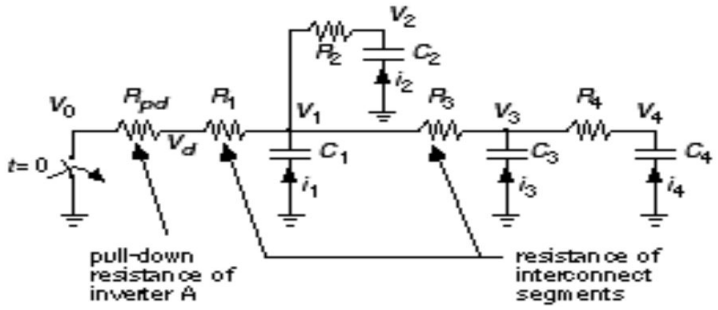
Work backward for sizes:

$$z = \frac{9C \cdot (4/3)}{5} = 2.4C$$

$$y = \frac{3z \cdot (4/3)}{5} = 1.9C$$

- This example also assumes the NMOS is twice as strong as the PMOS
- Note that when calculating the size y, you must consider the total load cap (3z) instead of just z
- Recall: D_{opt} = N * H^{1/N} + p1 + … + pN

Elmore Delay Approximation



$$\tau_{D4} = \sum_{k=0}^{4} C_k R_{4k} = C_0 R_{pd} + C_1(R_{pd} + R_1) + C_2(Rpd + R_1) + C_3(Rpd + R_1 + R_3) + C_4(Rpd + R_1 + R_3 + R_4) =$$

$$R_{pd}(C_0 + C_1 + C_2 + C_3 + C_4) + R_1(C_1 + C_2 + C_3 + C_4) + R_3(C_3 + C_4) + R_4 C_4$$

- Approximates the dominant time constant of a RC network for a given input and output node
- 2 methods (same result):
  1. Take every capacitor and multiply it by the sum of resistors on the path charging it
  2. Take every resistor on the path and multiply it by the sum of capacitances it charges

Switching Energy

- ½ CV^2 charge dumped on cap during charge cycle from VDD
- No additional charge drawn from supply when the cap discharges
- The other ½ CV^2 is burned on the charging resistance and dissipates as heat
- When the cap discharges the rest of the charge bleeds to ground via the discharging resistor
- Derive this result

Digital Power Consumption

- Dynamic switching power (f C V^2 alpha_{0 -> 1})
- Static leakage power (I_{d,leak}), what is it proportional to?
  - Model as an R_{off}

Sizing an Inverter Chain for Minimum Delay

- Don’t introduce the technique of logical effort based sizing yet, but go through writing the delay equation and take partials of every inverter size
- Then make it clear that the fanout of each stage should be the same and that determines the sizing
- A spice example would be nice here