ECE 471/571
The CMOS Inverter
Lecture-5
Gurjeet Singh
The Static CMOS Inverter
Operation

$V_{in} = V_{DD}$

$V_{in} = 0$
Properties of CMOS Design

- High Noise Margin (Voltage swing between $V_{DD}$ and GND)
- Ratioless gates. Logic levels are not dependent on the size of gates
- Low output impedance because of finite resistance in steady state (usually in kΩ range).
- High Input Impedance as gate of CMOS act as perfect insulator and draws no dc current.
- Theoretically, infinite fan-out (Theoretically)
- No direct path between supply and ground under steady state operating conditions
Voltage Transfer Characteristics

\[ I_{DSp} = -I_{DSn} \]
\[ V_{GSn} = V_{in} ; \quad V_{GSp} = V_{in} - V_{DD} \]
\[ V_{DSn} = V_{out} ; \quad V_{DSP} = V_{out} - V_{DD} \]

- For \( V_{in} = 0 \):
  - \( V_{GS} = -1 \)
  - \( V_{GS} = -2.5 \)

- For \( V_{in} = 1.5 \):
  - \( V_{out} = V_{DD} + V_{DSP} \)
  - \( I_{Dn} = -I_{DP} \)
Load-line plot
VTC of static CMOS Inverter
Dynamic Behavior (Overview)

(a) Low-to-high

(b) High-to-low
Switching Threshold

\[ k_n V_{DSATn} \left( V_M - V_{Th} - \frac{V_{DSATn}}{2} \right) + k_p V_{DSATp} \left( V_M - V_{DD} - V_{Th} - \frac{V_{DSATp}}{2} \right) = 0 \]

Solving for \( V_M \) yields

\[ V_M = \frac{\left( V_{Th} + \frac{V_{DSATn}}{2} \right) + r \left( V_{DD} + V_{Th} + \frac{V_{DSATp}}{2} \right)}{1 + r} \]

with \( r = \frac{k_p V_{DSATp}}{k_n V_{DSATn}} = \frac{v_{satp}}{v_{satn}} \frac{W_p}{W_n} \)

For large values of VDD (compared to threshold and saturation voltages)

\[ V_M \approx \frac{r V_{DD}}{1 + r} \]
Switching Threshold (cont.)

Ideally, we want \( V_m \) in the middle of voltage swing, i.e. \( r=1 \)

For \( r = 1 \),

\[
(W/L)_p = (W/L)_n \times \left( \frac{V_{DSATn}k'_n}{V_{DSATn}k'_p} \right)
\]

Also, to find ratio PMOS to NMOS transistor sizes

\[
\frac{(W/L)_p}{(W/L)_n} = \frac{k'_n V_{DSATn}(V_M - V_{Tn} - V_{DSATn}/2)}{k'_p V_{DSATp}(V_{DD} - V_M + V_{Tp} + V_{DSATp}/2)}
\]
Switching Threshold for long channel devices

When PMOS and NMOS are long channel devices or when the supply voltage is low (velocity saturation does not occur)

\[ V_M = \frac{V_{Tn} + r(V_{DD} + V_{Tp})}{1 + r} \quad \text{with} \quad r = \left(\frac{-k_p}{k_n}\right) \]
Example

- 0.25um CMOS process
- $V_M$ is in the middle
- Process Parameters

<table>
<thead>
<tr>
<th></th>
<th>$V_{T0}$ (V)</th>
<th>$\gamma$ ($V^{0.5}$)</th>
<th>$V_{DSAT}$ (V)</th>
<th>$k'$ (A/V^2)</th>
<th>$\lambda$ (V^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMOS</td>
<td>0.43</td>
<td>0.4</td>
<td>0.63</td>
<td>$115 \times 10^{-6}$</td>
<td>0.06</td>
</tr>
<tr>
<td>PMOS</td>
<td>-0.4</td>
<td>-0.4</td>
<td>-1</td>
<td>$-30 \times 10^{-6}$</td>
<td>-0.1</td>
</tr>
</tbody>
</table>

- Supply Voltage 2.5 V
- Minimum device W/L ratio has to be 1.5
- Find ratio of PMOS to NMOS size
Solution

\[
\frac{(W/L)_p}{(W/L)_n} = \frac{115 \times 10^{-6}}{30 \times 10^{-6}} \times \frac{0.63}{1.0} \times \frac{(1.25 - 0.43 - 0.63/2)}{(1.25 - 0.4 - 1.0/2)} = 3.5
\]
PMOS-to-NMOS vs $V_M$ Plot
$V_M$ is relatively insensitive to variations in device ratio.

- Ratio of 3, 2.5 and 2 yields switching threshold of 1.22V, 1.18V and 1.13 V respectively.
- Small variations do not disturb transfer characteristics that much.
- Industry practice to keep PMOS width less than required for each symmetry.
When asymmetric inverter needed

(a) Response of standard inverter

(b) Response of inverter with modified threshold
Noise Margin

\[ V_{IH} - V_{IL} = \frac{(V_{OH} - V_{OL})}{g} = -\frac{V_{DD}}{g} \]

\[ V_{IH} = V_M - \frac{V_M}{g} \quad V_{IL} = V_M + \frac{V_{DD} - V_M}{g} \]

\[ NM_H = V_{DD} - V_{IH} \quad NM_L = V_{IL} \]
Current equation:

\[ k_n V_{DSATn} \left( V_{in} - V_{Tn} - \frac{V_{DSATn}}{2} \right) (1 + \lambda_n V_{out}) + k_p V_{DSATp} \left( V_{in} - V_{DD} - V_{Tp} - \frac{V_{DSATp}}{2} \right) (1 + \lambda_p V_{out} - \lambda_p V_{DD}) = 0 \]

Differentiating and solving for \( \frac{dV_{out}}{dV_{in}} \), yields

\[ \frac{dV_{out}}{dV_{in}} = \frac{k_n V_{DSATn} (1 + \lambda_n V_{out}) + k_p V_{DSATp} (1 + \lambda_p V_{out} - \lambda_p V_{DD})}{\lambda_n k_n V_{DSATn} (V_{in} - V_{Tn} - V_{DSATn}/2) + \lambda_p k_p V_{DSATp} (V_{in} - V_{DD} - V_{Tp} - V_{DSATp}/2)} \]
Noise Margin (Contd.)

Ignoring some second order terms and setting $V_{in} = V_M$ produces the gain expression

$$g = -\frac{1}{I_D(V_M)} \frac{k_n V_{DSATn} + k_p V_{DSATp}}{\lambda_n - \lambda_p}$$

$$\approx \frac{1 + r}{(V_M - V_{Tn} - V_{DSATn}/2)(\lambda_n - \lambda_p)}$$

Gain is mostly dependent on technology parameters, especially channel length modulation. It can be influenced in a minor way by supply voltage and transistor sizes.
Example

- 0.25um CMOS technology
- PMOS-to-NMOS ratio = 3.4
- NMOS transistor minimum size
  - $W = 0.375\text{um}$, $L=0.25\text{um}$, $W/L = 1.5$
- Compute gain at $V_M = 1.25\text{V}$
- Calculate $V_{IL}$, $V_{IH}$, $V_{NM_L}$ and $V_{NM_H}$
Solution

\[ I_D(V_M) = 1.5 \times 115 \times 10^{-6} \times 0.63 \times (1.25 - 0.43 - 0.63/2) \times (1 + 0.06 \times 1.25) = 59 \times 10^{-6} \text{ A} \]

\[ g = \frac{1}{59 \times 10^{-6}} \times \frac{1.5 \times 115 \times 10^{-6} \times 0.63 + 1.5 \times 3.4 \times 30 \times 10^{-6} \times 1.0}{0.06 + 0.1} = -27.5 \]

\[ V_{IL} = 1.2 \text{ V}, \ V_{IH} = 1.3 \text{ V}, \ NM_L = NM_H = 1.2 \]
Simulated VTC and voltage gain
Lower than expected values

- Equation overestimated the gain
- Using piecewise linear approximation of the VTC

To use CMOS inverter as an amplifier it is biased in transition region, instead of cutoff and saturation.
Device Variations

![Graph showing device variations with voltage levels. The graph compares Good PMOS, Bad NMOS, Nominal, Good NMOS, and Bad PMOS.]
Scaling of Supply Voltage

- Reducing supply voltage indiscriminately has a positive impact on the energy dissipation, but is absolutely detrimental to the delay of the gate.
- DC characteristics become sensitive, once supply voltage and intrinsic voltages become comparable.
- Scaling of supply voltage reduces voltage swing. It reduces noise in system but make the design sensitive to external noise.
Scaling of Supply Voltage (Contd.)

(a) Reducing $V_{DD}$ improves the gain...

(b) ...but it deteriorates for very-low supply voltages.
Computing the Capacitances
Gate-Drain Capacitance

\[ C_{gd} = 2 \ C_{GD0} W \]
Diffusion Capacitances, $C_{db1}$ and $C_{db2}$

- Non linear and depends heavily on supply voltage.

$$C_{eq} = K_{eq} C_{j0}$$

$C_{j0}$ is junction capacitance per unit area under zero bias

$$K_{eq} = \frac{-\phi_0^m}{(V_{high} - V_{low})(1-m)[(\phi_0 - V_{high})^{1-m} - (\phi_0 - V_{low})^{1-m}]}$$

with $\phi_0$ the built-in junction potential and $m$ the grading coefficient of the junction.
Wiring Capacitance $C_w$

Capacitance due to length and width of the connecting wires.

Gate Capacitance of Fan-Out $C_{g3}$ and $C_{g4}$

\[
C_{fanout} = C_{gate}(\text{NMOS}) + C_{gate}(\text{PMOS})
\]
\[
= (C_{GSOn} + C_{GDon} + W_n L_n C_{ox}) + (C_{GSOp} + C_{G Dop} + W_p L_p C_{ox})
\]
Gate Capacitance of Fan-Out $C_{g3}$ and $C_{g4}$

\[
C_{fanout} = C_{\text{gate}}(\text{NMOS}) + C_{\text{gate}}(\text{PMOS})
\]
\[
= (C_{GSO_n} + C_{GDO_n} + W_n L_n C_{ox}) + (C_{GSO_p} + C_{GDO_p} + W_p L_p C_{ox})
\]

This expression simplifies the actual situation in two ways:

- Assuming all gate capacitances between $V_{DD}$ and GND and ignores Miller Effect
- Assuming channel capacitance is constant over interval of interest
**Example**

<table>
<thead>
<tr>
<th></th>
<th>$W/L$</th>
<th>$AD$ (μm²)</th>
<th>$PD$ (μm)</th>
<th>$AS$ (μm²)</th>
<th>$PS$ (μm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMOS</td>
<td>0.375/0.25</td>
<td>0.3 (19 μ²)</td>
<td>1.875 (15μ)</td>
<td>0.3 (19 μ²)</td>
<td>1.875 (15μ)</td>
</tr>
<tr>
<td>PMOS</td>
<td>1.125/0.25</td>
<td>0.7 (45 μ²)</td>
<td>2.375 (19μ)</td>
<td>0.7 (45 μ²)</td>
<td>2.375 (19μ)</td>
</tr>
</tbody>
</table>

Metal-1 wire area - 42μ²
Polysilicon wire area - 72μ²

<table>
<thead>
<tr>
<th></th>
<th>$C_{ox}$ (fF/μm²)</th>
<th>$C_D$ (fF/μm)</th>
<th>$C_J$ (fF/μm²)</th>
<th>$m_J$</th>
<th>$\phi_B$ (V)</th>
<th>$C_{jsw}$ (fF/μm)</th>
<th>$m_{jsw}$</th>
<th>$\phi_{jsw}$ (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMOS</td>
<td>6</td>
<td>0.31</td>
<td>2</td>
<td>0.5</td>
<td>0.9</td>
<td>0.28</td>
<td>0.44</td>
<td>0.9</td>
</tr>
<tr>
<td>PMOS</td>
<td>6</td>
<td>0.27</td>
<td>1.9</td>
<td>0.48</td>
<td>0.9</td>
<td>0.22</td>
<td>0.32</td>
<td>0.9</td>
</tr>
</tbody>
</table>

NMOS. High-to-low -2.5V to -1.25V
Bottom plate: $K_{eq}$ ($m = 0.5, \phi_0 = 0.9$) = 0.57,
Sidewall: $K_{eq,sw}$ ($m = 0.44, \phi_0 = 0.9$) = 0.61

PMOS. High-to-low 0 to -1.25V
Bottom plate: $K_{eq}$ ($m = 0.48, \phi_0 = 0.9$) = 0.79,
Sidewall: $K_{eq,sw}$ ($m = 0.32, \phi_0 = 0.9$) = 0.86

NMOS. Low-to-high 0V to -1.25V
Bottom plate: $K_{eq}$ ($m = 0.5, \phi_0 = 0.9$) = 0.79,
Sidewall: $K_{eq,sw}$ ($m = 0.44, \phi_0 = 0.9$) = 0.81

PMOS. Low-to-high -1.25V to -2.5V
Bottom plate: $K_{eq}$ ($m = 0.48, \phi_0 = 0.9$) = 0.59,
Sidewall: $K_{eq,sw}$ ($m = 0.32, \phi_0 = 0.9$) = 0.7
Solution

<table>
<thead>
<tr>
<th>Capacitor</th>
<th>Expression</th>
<th>Value (FF) (H→L)</th>
<th>Value (FF) (L→H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_{gd1}$</td>
<td>$2 \text{ CGD}_n \text{W}_n$</td>
<td>0.23</td>
<td>0.23</td>
</tr>
<tr>
<td>$C_{gd2}$</td>
<td>$2 \text{ CGD}_p \text{W}_p$</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>$C_{db1}$</td>
<td>$K_{eqn} \text{AD}<em>n \text{CJ} + K</em>{eqw} \text{PD}_n \text{CJSW}$</td>
<td>0.66</td>
<td>0.90</td>
</tr>
<tr>
<td>$C_{db2}$</td>
<td>$K_{eqp} \text{AD}<em>p \text{CJ} + K</em>{eqswp} \text{PD}_p \text{CJSW})$</td>
<td>1.5</td>
<td>1.15</td>
</tr>
<tr>
<td>$C_{g3}$</td>
<td>$(\text{CGD}_0 + \text{CGSO}) \text{W}<em>n + C</em>{ox} \text{W}_n \text{L}_n$</td>
<td>0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>$C_{g4}$</td>
<td>$(\text{CGD}_0 + \text{CGSO}) \text{W}<em>p + C</em>{ox} \text{W}_p \text{L}_p$</td>
<td>2.28</td>
<td>2.28</td>
</tr>
<tr>
<td>$C_w$</td>
<td>From Extraction</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>$C_L$</td>
<td>$\Sigma$</td>
<td>6.1</td>
<td>6.0</td>
</tr>
</tbody>
</table>

\[C_{wire} = \frac{42}{8^2} \text{um}^2 \times 30 \text{ aF/um}^2 + \frac{72}{8^2} \text{um}^2 \times 88 \text{ aF/um}^2 = 0.12 \text{ fF}\]
Propagation Delay: First-Order Analysis

\[ t_p = \int_{v_1}^{v_2} \frac{C_L(v)}{i(v)} dv \]

\[ R_{eq} = \frac{1}{V_{DD}/2} \int_{V_{DD}/2}^{V_{DD}} \frac{V}{I_{DSAT}(1+\lambda V)} dV \approx \frac{3}{4} \frac{V_{DD}}{I_{DSAT}} \left( 1 - \frac{7}{9} \lambda V_{DD} \right) \]

with \( I_{DSAT} = k' \frac{W}{L} \left( (V_{DD} - V_T) V_{DSAT} - \frac{V_{DSAT}^2}{2} \right) \)

\[ t_{pHL} = \ln(2) R_{eqn} C_L = 0.69 R_{eqn} C_L \]

\[ t_{pLH} = 0.69 R_{eqp} C_L \]

\[ t_p = \frac{t_{pHL} + t_{pLH}}{2} = 0.69 C_L \left( \frac{R_{eqn} + R_{eqp}}{2} \right) \]
Example

- 0.25um CMOS Process

<table>
<thead>
<tr>
<th>$V_{DD}$ (V)</th>
<th>1</th>
<th>1.5</th>
<th>2</th>
<th>2.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMOS (kΩ)</td>
<td>35</td>
<td>19</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>PMOS (kΩ)</td>
<td>115</td>
<td>55</td>
<td>38</td>
<td>31</td>
</tr>
</tbody>
</table>

- VDD = 2.5 V
- W/L NMOS = 1.5
- W/L PMOS = 4.5
- Find Propagation delay
Solution

\[ t_{PHL} = 0.69 \times \left( \frac{13k\Omega}{1.5} \right) \times 6.1fF = 36 \text{ psec} \]

\[ t_{PLH} = 0.69 \times \left( \frac{31k\Omega}{4.5} \right) \times 6.0fF = 29 \text{ psec} \]

and

\[ t_p = \left( \frac{36 + 29}{2} \right) = 32.5 \text{ psec} \]

Simulated tHL and tLH are 39.9 and 31.7 ps
Propagation Delay (Contd.)

\[
t_{pHL} = 0.69 \frac{3}{4} \frac{C_L V_{DD}}{I_{DSATn}} = 0.52 \frac{C_L V_{DD}}{(W/L)_n k'_n V_{DSATn} (V_{DD} - V_Tn - V_{DSATn}/2)}
\]

\[
t_{pHL} \approx 0.52 \frac{C_L}{(W/L)_n k'_n V_{DSATn}} , V_{DD} >> V_Tn + V_{DSATn}/2
\]
Design Techniques to reduce propagation delay

- Reduce $C_L$ (Better layout)
- Increase W/L ratio for transistors
- Increase $V_{DD}$
NMOS-to-PMOS ratio

\[ C_L = (C_{dp1} + C_{dn1}) + (C_{gp2} + C_{gn2}) + C_W \]

\[ C_L = (1 + \beta)(C_{dn1} + C_{gn2}) + C_W \]

PMOS are made \( \beta \) times larger than NMOS

\[ t_p = \frac{0.69}{2}((1 + \beta)(C_{dn1} + C_{gn2}) + C_W)\left( R_{eqn} + \frac{R_{eqp}}{\beta} \right) \]

\[ = 0.345((1 + \beta)(C_{dn1} + C_{gn2}) + C_W)R_{eqn}\left( 1 + \frac{r}{\beta} \right) \]

\[ \beta_{opt} = \sqrt{r\left( 1 + \frac{C_W}{C_{dn1} + C_{gn2}} \right)} \]
Sizing Inverters for Performance

\[ t_p = 0.69 \left( \frac{R_{\text{ref}}}{S} \right) (SC_{\text{iref}}) \left( 1 + C_{\text{ext}} / (SC_{\text{iref}}) \right) \]

\[ = 0.69 R_{\text{ref}} C_{\text{iref}} \left( 1 + \frac{C_{\text{ext}}}{SC_{\text{iref}}} \right) = t_{p0} \left( 1 + \frac{C_{\text{ext}}}{SC_{\text{iref}}} \right) \]

Conclusions:

- Intrinsic delay \( t_{p0} \) is independent of sizing of gate, and is determined by technology and inverter layout.
- Making \( S \) infinitely large yields maximum obtainable performance gain, eliminating impact of any external load, and reducing delay to intrinsic one. Having \( S \gg C_{\text{ext}} / C_{\text{int}} \) works as good as infinite \( S \).
Propagation delay vs sizing factor
Sizing a Chain of Inverters

\[ C_{int} = \gamma C_g \]

\( \gamma \) is proportionality factor. It is function of technology and is close to 1 for submicron processes

\[ t_p = t_{p_0} \left( 1 + \frac{C_{ext}}{\gamma C_g} \right) = t_{p_0} (1 + f/\gamma) \]

\( f \) is effective fan out
Sizing of chain of inverters

\[ t_{p,j} = t_{p0} \left( 1 + \frac{C_{g,j+1}}{\gamma C_{g,j}} \right) = t_{p0} \left( 1 + \frac{f_j}{\gamma} \right) \]

\[ t_p = \sum_{j=1}^{N} t_{p,j} = t_{p0} \sum_{j=1}^{N} \left( 1 + \frac{C_{g,j+1}}{\gamma C_{g,j}} \right), \text{ with } C_{g,N+1} = C_L \]
Sizing of chain of inverters

\[ C_{g,j} = \sqrt[\frac{N}{\gamma}]{C_{g,j-1}C_{g,j+1}} \]

\[ f = \frac{N}{\gamma} \sqrt[\frac{N}{\gamma}]{C_{L}/C_{g,1}} = \frac{N}{\gamma} \sqrt[\frac{N}{\gamma}]{F} \]

\[ t_p = N t_{p0} \left( 1 + \frac{N}{\gamma} \sqrt[\frac{N}{\gamma}]{F/\gamma} \right) \]

F is overall effective fan-out and strong function of number of inverters N