CMOS Digital Integrated Circuits

Chapter 5
MOS Inverters: Static Characteristics

Y. Leblebici

Ideal Inverter

Voltage Transfer Characteristic (VTC) of the ideal inverter
Generic Inverter VTC

Voltage Transfer Characteristic (VTC) of a typical inverter

Noise Margins

Propagation of digital signals under the influence of noise

- $V_{OH} : V_{OUT,\text{MAX}}$ when the output level is logic "1"
- $V_{OL} : V_{OUT,\text{MIN}}$ when the output level is logic "0"
- $V_{IL} : V_{IN,\text{MAX}}$ which can be interpreted as logic "0"
- $V_{IH} : V_{IN,\text{MIN}}$ which can be interpreted as logic "1"
Noise Margins

Definition of noise margins

\[ NM_L = V_{IL} - V_{OL} \]
\[ NM_H = V_{OH} - V_{IH} \]
Noise Margins

Nominal output

\[ V_{\text{out}} = f(V_{\text{in}}) \]

Output under noise

\[ V'_{\text{out}} = f(V_{\text{in}} + \Delta V_{\text{noise}}) \]

\[ V'_{\text{out}} = f(V_{\text{in}}) + \frac{dV_{\text{out}}}{dV_{\text{in}}} \cdot \Delta V_{\text{noise}} + \text{higher order terms (neglected)} \]

Perturbed Output = Nominal Output + Gain \times External Perturbation

The nominal operating region is defined as the region where the gain is less than unity!
CMOS Inverter Circuit

The NMOS switch transmits the logic 0 level to the output, while the PMOS switch transmits the logic 1 level to the output, depending on the input signal polarity.
CMOS Inverter Circuit

\[ V_{IL} = \frac{2V_{out} + V_{T0,p} - V_{DD} + k_R V_{T0,n}}{1 + k_R} \] (5.62)

\[ V_{IH} = \frac{V_{DD} + V_{T0,p} + k_R \cdot (2V_{out} + V_{T0,n})}{1 + k_R} \] (5.67)

\[ V_{th} = \frac{V_{T0,n} + \sqrt{\frac{V_{DD} + V_{T0,p}}{k_R}}}{1 + \sqrt{\frac{1}{k_R}}} \] (5.71)

Determine noise margins

Inversion (switching) threshold voltage
CMOS Inverter Circuit

nMOS transistor current-voltage characteristics

CMOS Inverter Circuit

pMOS transistor current-voltage characteristics
CMOS Inverter Circuit

Intersection of current-voltage surfaces of nMOS and pMOS transistors

CMOS Inverter Circuit

Intersection of current-voltage surfaces gives the VTC in the voltage plane
CMOS Inverter Circuit

How to choose the $k_R$ ratio to achieve a desired inversion threshold voltage:

$$k_R = \frac{k_n}{k_p} = \left( \frac{V_{D,0} + V_{2,0,0} - V_{D,0}}{V_{D,0} - V_{T,0}} \right)^2$$  \hspace{1cm} (5.73)

$$k_n = \frac{\mu_n C_{ox} \cdot \left( \frac{W}{L} \right)_n}{\mu_p C_{ox} \cdot \left( \frac{W}{L} \right)_p} \quad \text{and} \quad k_p = \frac{\mu_p C_{ox} \cdot \left( \frac{W}{L} \right)_p}{\mu_n C_{ox} \cdot \left( \frac{W}{L} \right)_n}$$  \hspace{1cm} (5.77)
CMOS Inverter Circuit

Supply Voltage Scaling

VTC of a CMOS inverter for different power supply voltage values.
Dynamic Characteristics of CMOS Inverter

Switching speed determined by the time required to charge-up or charge-down the output load capacitance.

Definition of Delay Times

The propagation delay times are defined as the time delay between the 50% crossing of the input and the corresponding 50% crossing of the output.

The rise time and the fall time of the output signal are defined as the time required for the voltage to change from its 10% level to its 90% level (or vice versa).
Calculation of Propagation Delays

The falling output propagation delay is determined by the NMOS transistor, which starts pulling down the output node first in saturation, and later, in linear mode.

Calculation of Propagation Delays

The rising output propagation delay is determined by the PMOS transistor, which starts pulling up the output node first in saturation, and later, in linear mode.
Calculation of Propagation Delays

The time required by the NMOS transistor in saturation and in linear region can be calculated separately.

\[
t_{\text{PHL,sat}} = t_1 - t_0 \\
\]

\[
t_{\text{PHL,lin}} = t_2 - t_1 \\
\]

\[
t_{\text{PHL}} = t_{\text{PHL,sat}} + t_{\text{PHL,lin}} = t_2 - t_0 \\
\]

\[
c_i \frac{dV}{dt} = dt = C_i \frac{dV}{I} \\
\]

Calculation of Propagation Delays

\[
t_{\text{PHL,sat}} = \frac{2nC_i V_{\text{THS}}}{\beta_n (V_{OH} - V_{THS})^2} \\
\]

\[
t_{\text{PHL,lin}} = \frac{nC_i}{\beta_n (V_{OH} - V_{THS})} \ln \left( \frac{4(V_{OH} - V_{THS})}{V_{OL} + V_{OH}} - 1 \right) \\
\]

\[
t_{\text{PHL}} = t_{\text{PHL,sat}} + t_{\text{PHL,lin}} = \frac{nC_i}{\beta_n (V_{OH} - V_{THS})} \left[ \frac{2V_{THS}}{V_{OH} - V_{THS}} + \ln \left( \frac{4(V_{OH} - V_{THS})}{V_{OL} + V_{OH}} - 1 \right) \right] \\
\]
Calculation of Propagation Delays

The time required by the PMOS transistor in saturation and in linear region can be calculated separately.

Calculation of Propagation Delays

\[ t_{PLH,sat} = \frac{2nC_l V_{TH}}{\beta_x (V_{dd} - V_{OL} - V_{TH})^2} \]

\[ t_{PLH,lin} = \frac{nC_l}{\beta_x} \ln \left( \frac{4(V_{dd} - V_{OL} - V_{TH})}{2V_{dd} - V_{OL} - V_{OLH}} - 1 \right) \]

\[ t_{PLH} = t_{PLH,sat} + t_{PLH,lin} = \frac{nC_l}{\beta_x (V_{dd} - V_{OL} - V_{TH})} \left[ \frac{2V_{TH}}{V_{dd} - V_{OL} - V_{TH}} + \ln \left( \frac{4(V_{dd} - V_{OL} - V_{TH})}{2V_{dd} - V_{OL} - V_{OLH}} - 1 \right) \right] \]
Calculation of Propagation Delays

\[ t_{\text{pHL}} = \frac{nC_i}{\beta_p (V_{OH} - V_{TH})} \left[ \frac{2V_{TH}}{V_{OH} - V_{TH}} \ln \left( \frac{4(V_{OH} - V_{TH})}{V_{OH} + V_{TH}} - 1 \right) \right] \]

\[ t_{\text{pLH}} = \frac{nC_i}{\beta_p (V_{OH} - V_{OL} - V_{TH})} \left[ \frac{2V_{TH}}{V_{OL} - V_{TH}} \ln \left( \frac{4(V_{OL} - V_{TH})}{2V_{OL} - V_{TH} - V_{OH}} - 1 \right) \right] \]

Summary:

The propagation delay times are proportional to the load capacitance, and inversely proportional to the pull-up (or pull-down) transconductance.

Influence of Nonzero Input Rise/Fall Time

![Graph showing the influence of nonzero input rise/fall time on propagation delays.]

\[ t_{\text{pHL, real}} = \sqrt{t_{\text{pHL, ideal}}^2 + \left( \frac{t_r}{2} \right)^2} \]

\[ t_{\text{pLH, real}} = \sqrt{t_{\text{pLH, ideal}}^2 + \left( \frac{t_f}{2} \right)^2} \]
Components of the Output Load Capacitance

![Diagram of components of the output load capacitance]

Increasing the Switching Speed

- Need to increase the amount of charge-up and charge-down current that switches the output.
- Increase the transistor (W/L) ratios both in the pull-up path and in the pull-down path.
- However: increasing transistor dimensions will influence the parasitic MOSFET capacitances.

- The switching speed of a classical CMOS gate is ultimately limited by its intrinsic delay.
A closer look at the typical CMOS inverter

Drain parasitics are mainly responsible for the intrinsic gate delay.

\[ C_{\text{load}} = C_{\text{gd}}(W_n) + C_{\text{gd}}(W_p) + C_{\text{db}}(W_n) + C_{\text{db}}(W_p) + C_{\text{int}} + C_g = f(W_n, W_p) \]

Parasitic Capacitance Components

Drain parasitics:

\[
\begin{align*}
C_{\text{db,n}} &= W_n D_{\text{Drain}} C_{j0,n} K_{eq,n} + 2(W_n + D_{\text{Drain}}) C_{jsw,n} K_{eq,n} \\
C_{\text{db,p}} &= W_p D_{\text{Drain}} C_{j0,p} K_{eq,p} + 2(W_p + D_{\text{Drain}}) C_{jsw,p} K_{eq,p}
\end{align*}
\]

Total capacitive load:

\[
\begin{align*}
C_{\text{load}} &= (W_n C_{j0,n} K_{eq,n} + W_p C_{j0,p} K_{eq,p}) D_{\text{Drain}} + 2(W_n + D_{\text{Drain}}) C_{jsw,n} K_{eq,n} + 2(W_p + D_{\text{Drain}}) C_{jsw,p} K_{eq,p} + C_{\text{int}} + C_g
\end{align*}
\]
Parasitic Capacitance Components

\[ C_{load} = \alpha_0 + \alpha_n W_n + \alpha_p W_p \]

\[ \alpha_0 = 2D_{\text{drain}}(C_{jsw,n}K_{eq,n} + C_{jsw,p}K_{eq,p}) + C_{int} + C_g \]

\[ \alpha_n = K_{eq,n}(C_{j0,n}D_{\text{drain}} + 2C_{jsw,n}) \]

\[ \alpha_p = K_{eq,p}(C_{j0,p}D_{\text{drain}} + 2C_{jsw,p}) \]

Propagation Delays

\[ \tau_{PHL} = \left( \frac{\alpha_0 + \alpha_n W_n + \alpha_p W_p}{W_n} \right) \times \left( \frac{L_n}{\mu_n C_{ox}(V_{DD} - V_{T,n})} \right) \times \left[ \frac{2V_{T,n}}{V_{DD} - V_{T,n}} + \ln \left( \frac{4(V_{DD} - V_{T,n})}{V_{DD}} - 1 \right) \right] \]

\[ \tau_{PLH} = \left( \frac{\alpha_0 + \alpha_n W_n + \alpha_p W_p}{W_p} \right) \times \left( \frac{L_p}{\mu_p C_{ox}(V_{DD} - |V_{T,p}|)} \right) \times \left[ \frac{2|V_{T,p}|}{V_{DD} - |V_{T,p}|} + \ln \left( \frac{4(V_{DD} - |V_{T,p}|)}{V_{DD}} - 1 \right) \right] \]
Propagation Delays

\[ \tau_{PHL} = \Gamma_n \left( \frac{\alpha_0 + (\alpha_n + R\alpha_p) W_n}{W_n} \right) \]

\[ \tau_{PLH} = \Gamma_p \left( \frac{\alpha_0 + (\alpha_n + \alpha_p) W_p}{W_p} \right) \]

Notice that with increasing transistor width \( W_n \) and \( W_p \), both delays asymptotically approach fixed limit values, independent of external loads.

Propagation Delays

\[ \Gamma_n = \left( \frac{L_n}{\mu_n C_{\mu}(V_{DD} - V_{T,n})}\right) \times \left[ \frac{2V_{T,n}}{V_{DD} - V_{T,n}} + \ln \left( \frac{4(V_{DD} - V_{T,n})}{V_{DD}} \right) - 1 \right] \]

\[ \Gamma_p = \left( \frac{L_p}{\mu_p C_{\mu}(V_{DD} - |V_{T,p}|)}\right) \times \left[ \frac{2|V_{T,p}|}{V_{DD} - |V_{T,p}|} + \ln \left( \frac{4(V_{DD} - |V_{T,p}|)}{V_{DD}} \right) - 1 \right] \]

Intrinsic delay limits will decrease for

- Smaller gate length \( L \)
- Larger \( \mu \) and \( C_{\text{ox}} \)

But for a given technology, the limits are fixed.
Improving propagation delay times by transistor sizing

![Graph showing voltage over time with different channel widths and delays.

- Large fanout load: Gate delays ultimately converge to limit value for large W/L.
- Intrinsic delay limit value.
- Small fanout load.](image-url)

The graph illustrates how increasing the channel width improves the propagation delay times by transistor sizing. The limit value for large W/L is reached due to the reduced gate delays.
CMOS Inverter Layout

Mask layout of the inverter

Simplified cross-section