

CHAPTER 4

MOS Inverter Circuits

CHAPTER OUTLINE

- 4.1 Introduction
- 4.2 Voltage Transfer Characteristics
- 4.3 Noise Margin Definitions
- 4.4 Resistive-Load Inverter Design
- 4.5 NMOS Transistors as Load Devices
- 4.6 Complementary MOS (CMOS) Inverters
- 4.7 Pseudo-NMOS Inverters
- 4.8 Sizing Inverters
- 4.9 Tristate Inverters
- 4.10 Summary

References

Problems

4.1 Introduction

Digital MOS circuits are broadly categorized into two groups: static and dynamic. All nodes of a static gate will have resistive paths to V_{DD} or ground. In dynamic circuits, the value of one or more nodes is based on stored charge on a capacitor. Another distinction between the two is based on whether periodic *clock signals* are necessary to achieve combinational logic functions. *Static circuits* require no clock or other periodic signals for their operation in combinational logic networks. Of course, clocks are required for sequential logic circuits, but they are usually applied only to normal logic gate inputs. *Dynamic circuits* require periodic clock signals synchronized with data signals for proper operation, even in combinational logic applications. In dynamic circuits, clock signals are applied to load elements, and so-called *transmission gates*, or *transfer gates*.

In this chapter, we study static inverters with emphasis on voltage transfer characteristics, noise margins, inverter configurations, and delay models. Analytical

expressions will be developed for the design and analysis of various inverters using the equations and models described in Chapters 2 and 3. In subsequent chapters, we will follow up with gate circuits, sequential circuits, and dynamic logic circuits.

4.2 Voltage Transfer Characteristics

As we already know, a digital logic inverter takes an input and produces its complement at the output. The two possible binary output levels are 0 and 1. If the input is a logical “0,” the output is a logical “1,” and vice versa. In an ideal implementation of an inverter circuit, the two binary output levels are typically at ground¹ (zero volts) and the supply voltage, V_{DD} . The output transition between the 1 and 0 states should occur when the input is exactly equal to one-half the supply voltage (i.e., at $V_{DD}/2$). Realistic inverters have analog voltages that are discretized into one of the two logic states. To capture the features of ideal and practical inverters more formally, a plot of the output voltage, V_{out} , as a function of the input voltage, V_{in} , can be constructed. This plot, called a dc voltage transfer characteristic (VTC), can tell us much about the inherent properties of an inverter and other gates. It is also a useful way to compare various inverter configurations. It is produced by increasing the input voltage very slowly and plotting the resulting output voltage, as shown in Figure 4.1a.

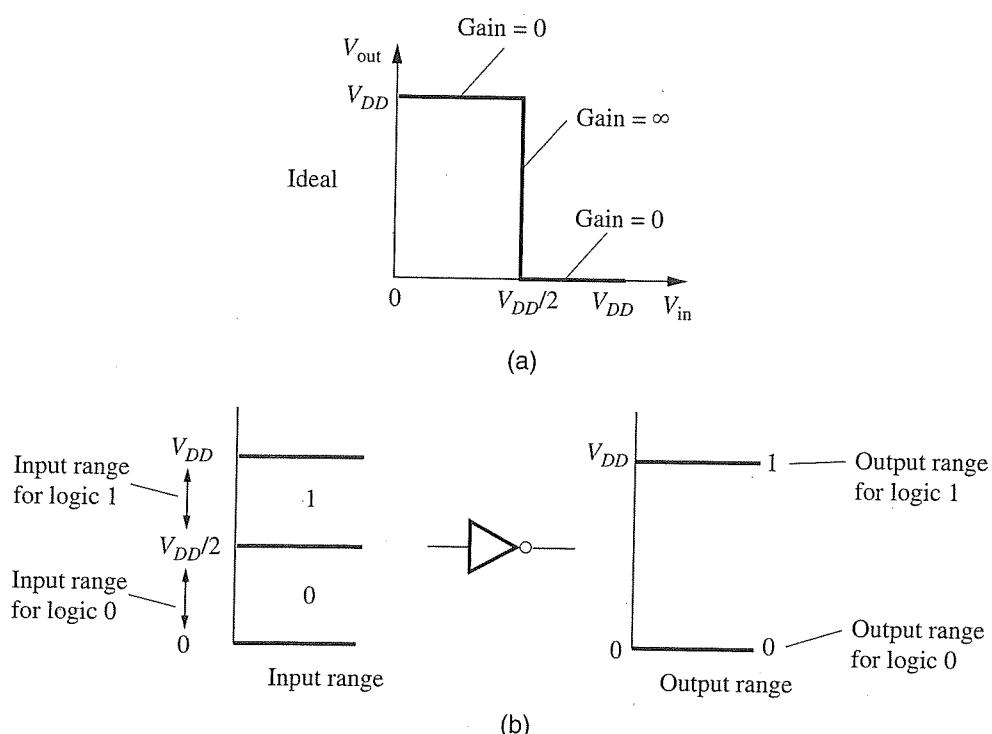


Figure 4.1

Ideal VTC of an inverter.

¹ We will also use Gnd and V_{SS} to refer to the ground terminal throughout the book.

The VTC in Figure 4.1a illustrates the desired features of an ideal inverter. The first thing to notice is that the switching point occurs at $V_{DD}/2$. At this point, the output switches from V_{DD} to Gnd. The slope of the VTC, referred to as the *gain*, is obtained by taking the derivative of V_{out} with respect to V_{in} :

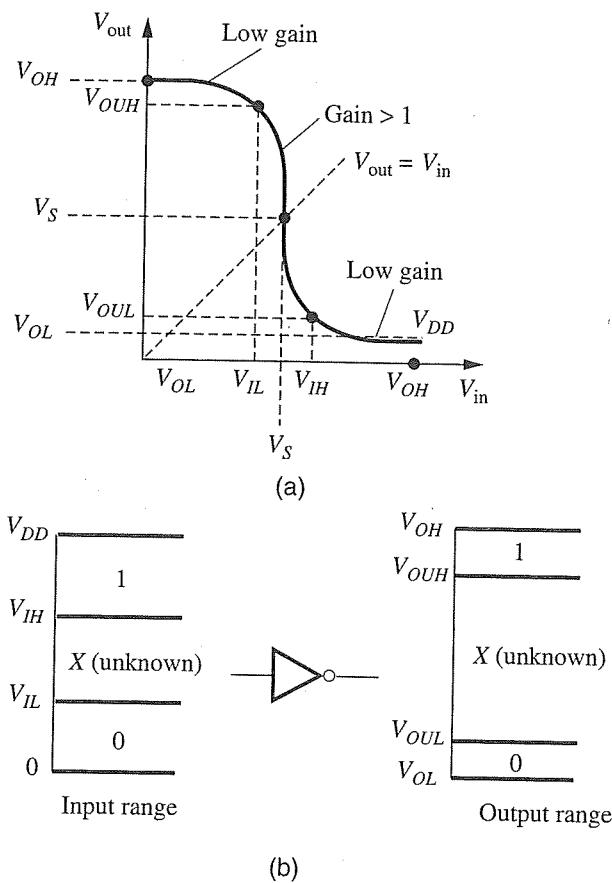
$$\text{gain} = \frac{\partial V_{out}}{\partial V_{in}} \quad (4.1)$$

The gain can be computed at any point along the VTC. By examining the characteristic in Figure 4.1a, it is clear that there are three gain regions in the ideal inverter: two zero gain regions and one infinite gain region. The high gain region, separating the high output from the low output, is a feature required by all useful logic gates to regenerate high and low logic values if there is noise in the system. We will consider this property more carefully in the next section.

Another feature of the VTC is that the input range is very large while the output range is small for the ideal inverter. The *range* refers to the voltage interval over which a signal is considered to be a logic 0 or logic 1. Having a large input range and small output range is a desirable characteristic of a logic gate for noise immunity. It implies that there are low gain regions such that the input can vary significantly with little or no effect on the output. That is, the gate is able to reject noise at the input. The input range is $V_{DD}/2$ for both logic 0 and logic 1, as shown in Figure 4.1b; the output range is essentially zero for both logic 0 and logic 1. This is the ideal condition: the input can swing anywhere between 0 and $V_{DD}/2$ and the output will stay at 1; or, it can swing between $V_{DD}/2$ to V_{DD} and the output will stay at 0.

All practical logic gates fall short of the ideal performance, but depending upon the application, some are better than others. A more realistic VTC for the inverter is shown in Figure 4.2a. In practical inverters, the low output voltage, V_{OL} , may not reach Gnd, and the high output voltage, V_{OH} , may not reach V_{DD} . The output does not abruptly switch from V_{DD} to Gnd at $V_{DD}/2$. Instead, the concept of a switching point, V_S , is defined as the point where $V_{out} = V_{in}$. Rather than two zero gain regions and an infinite gain region, there is a region of low gain, followed by high gain, followed again by low gain as the input increases from 0 to V_{DD} . For a valid gate, the gain must be larger than one in the high gain region and less than one in the low gain region.

The input and output ranges for the nonideal inverter are provided in Figure 4.2b. The two input ranges that define 0 and 1 are smaller than the ideal case, while the two output ranges are larger than the ideal case. The input range for logic 0 is from 0 V to a point called V_{IL} where the input is still considered to be low. The input range for logic 1 is the interval from V_{OH} to V_{IH} , where the input is still considered to be high. Since $V_{IL} \neq V_{IH}$, a new interval exists between 0 and 1, defined as the *unknown (X)* or *uncertain region*. The output ranges are from V_{OL} to V_{OUL} for logic 0 and V_{OH} to V_{OUH} for logic 1. The values V_{OUL} and V_{OUH} define the edges of the two output ranges, and since they are not equal, an uncertain region (X) also exists in the output. Perhaps more importantly, the nonideal inverter retains the desirable noise rejection property if the input range is larger than the output range. This

**Figure 4.2**

Practical VTC of an inverter.

implies that output fluctuations are small even with large input fluctuations due to noise.

To study this concept further, consider the effect of noise at the input of the inverter shown in Figure 4.3. In this example, the input varies considerably but remains within the valid logic 0 range for the input. The output varies by a smaller amount but remains in the range considered to be a valid logic 1. In terms of the voltage transfer characteristics, the input is varying in the range where the logic gate acts as a low-gain amplifier, from V_{OL} to V_{IL} . This attenuates the noise since the gain is less than one in this range. As a result, the output remains in the range V_{OEH} to V_{OH} . The next few inverters in the chain will attenuate the noise even further and eventually the noise is damped out of the system. If the output range were larger than the input range, a small fluctuation of the input voltage would produce a larger fluctuation at the output voltage. As the signal propagated through subsequent inverter stages, it would be corrupted and eventually the wrong answer would be produced. However, if the noise takes the input signal outside the proper range, the high gain of the inverter can still force subsequent stages to their correct values,

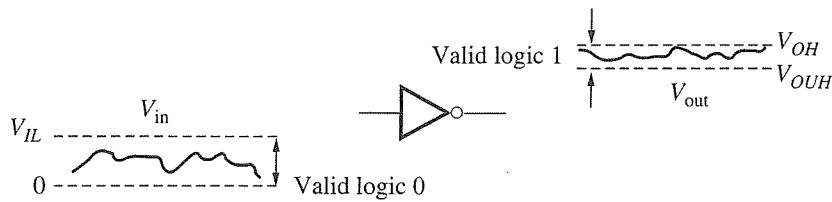


Figure 4.3

Effect of noise at logic gate input on the output.

assuming that the noise does not exceed a predefined threshold. This is the key (desirable) property of logic gates called the *regenerative* property.

4.3 Noise Margin Definitions

The word *noise* in the context of digital circuits and systems means unwanted variations of voltages or currents at logic nodes. If the magnitude of noise is too great, it will cause logic errors. However, if the noise amplitude at the input of any logic circuit is smaller than a specified critical magnitude known as the *noise margin* of that circuit, the noise will be attenuated as it passes from input to output. In properly functioning digital systems, the desired logic signals are restored to full amplitude without error, and noise does not accumulate from one logic stage to the next (as it does in analog systems).

Digital circuits typically exhibit variations in logic levels due to circuit manufacturing tolerances, temperature changes, power supply variations, and electrical loading at the output node. For example, noise may be transferred to logic nodes or interconnecting lines by unwanted capacitive or inductive coupling. Series inductance and resistance in ground and power supply lines shared by many logic elements are also a common source of noise problems. The worst combinations of all these effects are used to define the worst-case output voltage range. It is desirable to minimize these variations so that output logic levels are held within two narrow voltage ranges around V_{OH} and V_{OL} .

Noise margins are used to specify the range over which the circuits will function properly. A variety of noise metrics have been defined to characterize the noise performance of logic gates using the VTC. We will examine the single-source noise margin (SSNM) and the multiple-source noise margin (MSNM).² The robustness of a gate, that is, its ability to operate properly in the presence of noise, depends on how much noise can be applied before the gate fails and how much noise actually couples into the gate. The first factor is a function of the gate itself while the second factor is a function of the environment of the gate. We study the first factor in the next section and the second factor in Chapters 10 and 11. Both are important in the performance of a logic gate in an actual circuit.

² In the literature, these are also referred to as single-stage and multi-stage noise margins, respectively.

4.3.1 Single-Source Noise Margin (SSNM)

The SSNM is associated with a *single noise source* and its effect on downstream logic gates. The noiseless inverter chain of Figure 4.4a produces a sequence of inverted values starting with V_{OL} and ending with V_{OL} . Without noise, the output of the first inverter is V_{OH} and this serves as input to the second inverter, and so on. As long as there is no noise in the system, the output of each inverter will remain at the values shown in the figure. When noise events occur, the outputs deviate from their steady-state values. We seek to find how much noise the inverter chain can tolerate before it fails.

The SSNM metric is defined as the largest noise level in a single stage that allows subsequent stages to recover to their proper value (the regenerative property). To illustrate this metric using Figure 4.4b, let us arbitrarily insert a noise source of magnitude v_n at the output of the first inverter. This noise will change the input to the second inverter to $V_{OH} - v_n$ and this, in turn, will affect its output value. The question is, how much noise can be tolerated at the input of the second inverter before its output and all subsequent outputs flip their logic values?

Figure 4.5 provides some insight into how the VTC can be used to determine the SSNM. Note that we have plotted the second inverter VTC in the normal orientation, but the third inverter has the input and output axes swapped for convenience. If V_{OH} is the output of the first inverter, we know that the output of the second inverter should be at V_{OL} . However, if a noise level of v_n is applied, the input of the second inverter will drop to $V_{OH} - v_n$ and its output will rise to a value well above V_{OL} , as shown in the figure (second inverter). This new output is the input to the third stage. To determine the value at the output of the third inverter, we transfer the second stage output value to the adjacent VTC (mirror image of the first one), and then read the new value directly from this VTC. This is depicted on the graph using arrows. Of course, to obtain the output of the fourth inverter, we now have to transfer this new value back to the first graph. This is a very cumbersome process given that we have so many stages in the inverter chain.

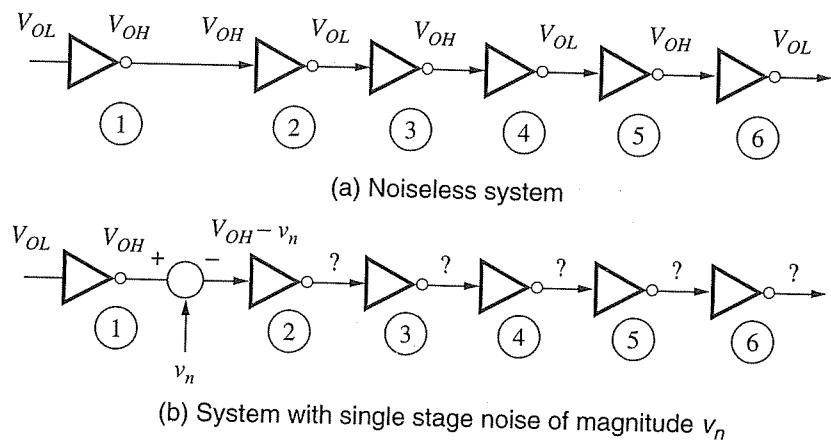


Figure 4.4

Single-source noise analysis for inverter chain.

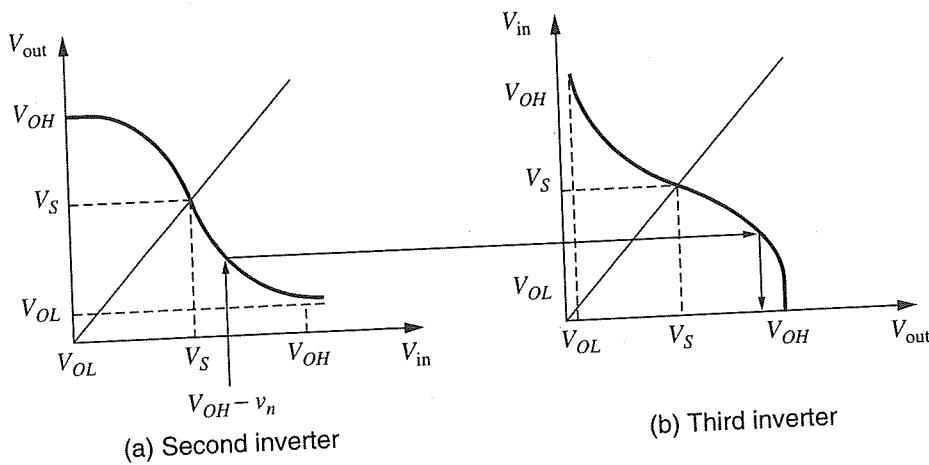


Figure 4.5
Noise propagation in an inverter chain.

A more convenient way of obtaining the output values is shown in Figure 4.6. Since we are interested in the input and output conditions of the second inverter and beyond, we can plot the VTC of the even and odd inverters on the same graph. The normal orientation is used for all even inverters, while the swapped axes (V_{in} versus V_{out}) are used for odd inverters. This configuration allows us to switch back and forth between the two curves very quickly to determine the outputs of the subsequent stages. With this efficient method in hand, we can now go back to the question of determining the SSNM.

Consider Figure 4.6 in the context of the maximum possible noise that can be tolerated. When the input with noise v_n is applied to the second inverter, the output is V_{O2} . This is the input to the next stage which produces an output of V_{O3} . This

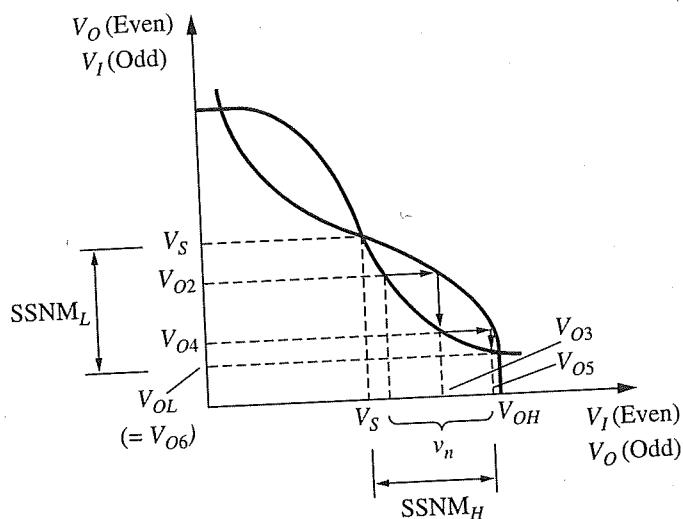


Figure 4.6
Graphical approach to SSNM calculation.

output feeds the next stage and produces V_{O4} , and so on. The sequence of output values can be traced on the two curves by following alternate horizontal and vertical paths. An important observation can be made from this graph: a noise level that does not push the voltage past V_S can be tolerated by the downstream inverters (inverters that are two or three stages away) due to the regenerative effect. That is, the inverter would be able to tolerate noise until the input exceeds V_S in one direction or the other.

From this analysis, we can develop a noise metric for the single-source scenario for both high and low output cases:

$$\begin{aligned} \text{SSNM}_H &= V_{OH} - V_S \\ \text{SSNM}_L &= V_S - V_{OL} \end{aligned} \quad (4.2)$$

Any noise above these levels will cause the output to flip to the opposite value and recovery to the original value will not be possible. The importance of the switching threshold in the noise metric is clear. If an input is high, then noise will not adversely affect the chain until it reaches the switching threshold. Similarly, if the input is low, noise will not greatly affect the outputs until it reaches the switching threshold. An interesting side note is that if we want the inverter to switch, the input must be taken past V_S ; hence, the name *switching threshold*.

Exercise 4.1

Using the graph of Figure 4.6, plot the trajectory of the outputs of a six-stage inverter chain when the noise level is greater than V_S at the input of the second inverter.

4.3.2 Multiple-Source Noise Margin (MSNM)

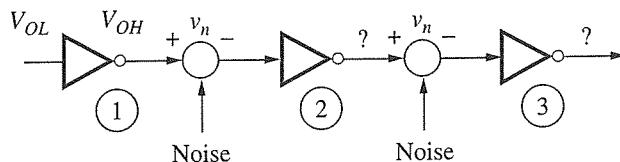
While the single-source noise margin is instructive on the importance of the switching threshold, V_S , it does not fully capture the true nature of noise. Noise tends to be added to all nodes in the circuit simultaneously rather than just to one node. We need to develop a multi-source noise margin metric.

In Figure 4.7, a three-stage inverter chain is shown with two noise sources. The question we need to ask is, how much noise can the inverters tolerate in this new situation? One way to answer this question is to determine the point at which the noise is amplified from input to output. This is a sensible definition since we want noise to be attenuated as it moves through inverter stages. In order to develop a metric suitable for the multi-source situation, we first assume that the noise is small. For a noiseless system, we can write the equation for an inverter as

$$V_{\text{out}} = f(V_{\text{in}})$$

With noise ν_n added, a new output is produced:

$$V'_{\text{out}} = f(V_{\text{in}} + \nu_n)$$

**Figure 4.7**

Multistage noise margin.

A Taylor series expansion³ of the output function allows us to examine the important factors determining V_{out} in the presence of noise:

$$V'_{\text{out}} = f(V_{\text{in}}) + \nu_n \frac{\partial V_{\text{out}}}{\partial V_{\text{in}}} + (\nu_n)^2 \frac{\partial^2 V_{\text{out}}}{\partial V_{\text{in}}^2} + (\nu_n)^3 \frac{\partial^3 V_{\text{out}}}{\partial V_{\text{in}}^3} + \dots$$

A simple interpretation of this equation is that the

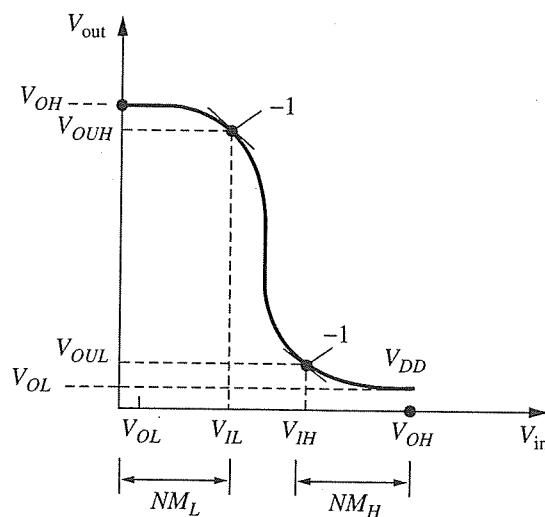
$$\text{noisy_output_voltage} = \text{noiseless_output} + \text{noise} \times \text{gain} + \text{higher_order_terms}$$

If we ignore the higher-order terms, then the output is simply the noiseless value plus the noise multiplied by the gain of the inverter. Given this result, we can develop a metric that is based on the gain of the inverter since it controls how the noise is amplified as it passes through the inverter.

Specifically, if the inverter is operating in a region where the gain $|\partial V_{\text{out}} / \partial V_{\text{in}}| > 1$, then the noise is amplified and added to the output. This is not a desirable situation. However, if we operate in the region where $|\partial V_{\text{out}} / \partial V_{\text{in}}| < 1$, then the circuit will attenuate the noise and faithfully try to hold the output in the desired range.

Based on the above analysis, we can now define more useful noise margins by using the VTC points where the gain is 1 to establish important transition points. In Figure 4.8, the VTC of the inverter is shown with the unity gain points identified. There are two unity gain points, with the slope of the line equal to -1 in both cases. The first unity gain point occurs where $V_{\text{in}} = V_{IL}$ and $V_{\text{out}} = V_{OUL}$. When $V_{\text{in}} < V_{IL}$, the output is still considered to be high. If $V_{\text{in}} > V_{IL}$, the gain exceeds unity and the output begins to drop significantly. Therefore, the input can safely swing from V_{OL} to V_{IL} without a significant drop in V_{out} or a significant gain in the noise. Likewise, the second unity gain point occurs at $V_{\text{in}} = V_{IH}$ and $V_{\text{out}} = V_{OUL}$. When $V_{\text{in}} > V_{IH}$, the output is considered to be a valid low. However, if $V_{\text{in}} < V_{IH}$, then the output passes the unity gain point and begins to rise in value. Therefore, the input can safely swing from V_{OH} to V_{IH} without a significant rise in V_{out} .

³ A Taylor series expansion of a function $f(x)$ is a polynomial expansion of the function about a given operating point. We can write $f(x) = f(x_0) + (x - x_0)(\partial f / \partial x) + (x - x_0)^2(\partial^2 f / \partial x^2) + \dots$. In our case, $x = V_{\text{in}} + \nu_n$, $x_0 = V_{\text{in}}$ and therefore $(x - x_0)$ is ν_n .

**Figure 4.8**

Unity gain noise margin definitions.

These two unity gain points and the output high and low values can be used to define the noise margins as follows:

$$\begin{aligned} NM_H &= V_{OH} - V_{IH} \\ NM_L &= V_{IL} - V_{OL} \end{aligned} \quad (4.3)$$

As long as the inputs stay within these ranges, the gate will operate properly. In general, noise margins are different in high and low logic states so both must be calculated. There are a number of tradeoffs between the noise margins and other characteristics of circuit performance, as we will discover in the following sections. We will use the equations in (4.3) to compare a number of different inverter configurations in the rest of this chapter.

Example 4.1 VTC for a Buffer

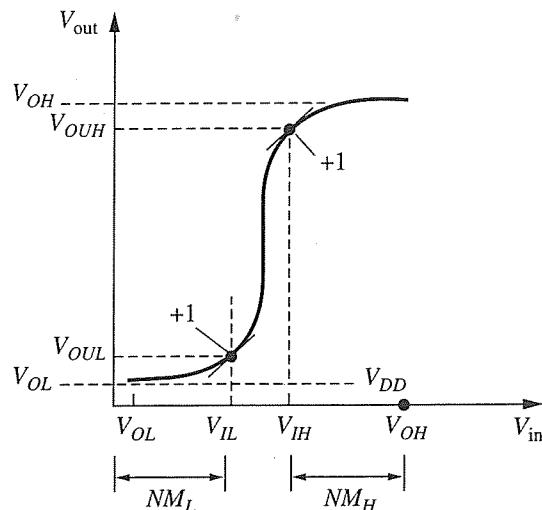
Problem:

Sketch the VTC for a noninverting (buffer) gate. Identify the noise margins and any other important features on the VTC.

Solution:

The VTC for buffer is shown in the following diagram. In the noninverting case, the logic output value is equal to the logic input value so the buffer VTC is the mirror image of the inverter VTC. The two unity gain points have a positive slope and are given by V_{IL} and V_{IH} . The output high and low values are V_{OH} and V_{OL} , respectively. If we now place these values along the V_{in} axis, we can see from the figure that the input range of V_{OL} to V_{IL} and V_{IH} to V_{OH} are the intervals over which the output

remains approximately constant. Beyond these ranges, the output enters the high gain region and amplifies the noise. It is clear that the noise margin definitions given above are meaningful metrics for inverters and buffers.

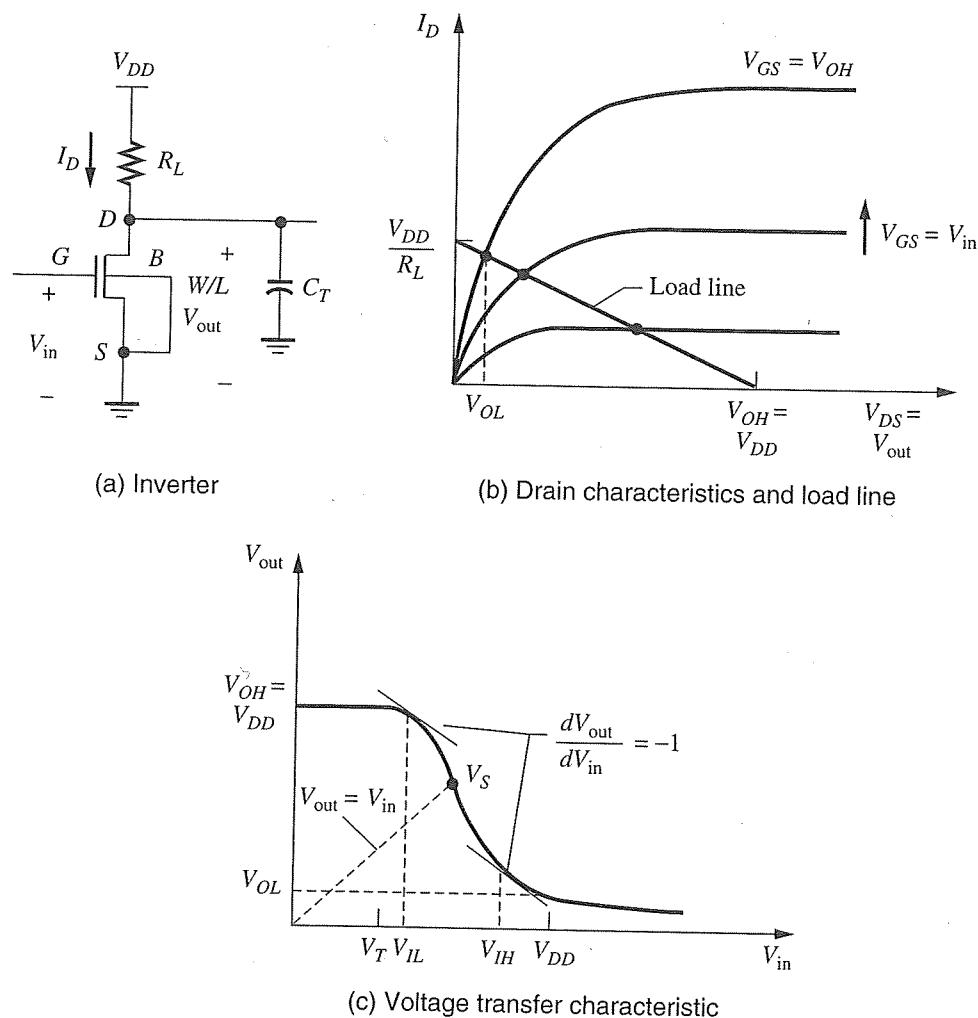


4.4 Resistive-Load Inverter Design

We begin our study of inverters with a configuration that uses a resistor attached to the drain of a MOS transistor, as in Figure 4.9. The resistive-load MOS inverter exhibits all of the essential features of MOS logic gates and is therefore a useful starting point. Subsequent extension of these concepts to CMOS inverter, NOR and NAND gates is very simple. In the rest of this chapter we will spend a good deal of time establishing analysis techniques to obtain the dc voltage transfer characteristics, noise margins, and propagation delays for inverters. Both NMOS and CMOS circuits are considered. Alternative load elements are compared, considering noise margin, power consumption, circuit density, transfer characteristics, and transient performance.

Figure 4.9a shows a single NMOS transistor connected with a resistor load to form an inverter. When the NMOS device is off, the resistor pulls the output high to V_{OH} . When the NMOS device is on, it forms a resistive divider with the pull-up device and produces a low output, V_{OL} . The design parameters are the values of R_L and W/L of the NMOS device. The supply voltage is also a design parameter, but it is usually specified by the technology chosen by the designer, or by system-level specifications. The quantitative design of such an inverter is guided by several considerations such as timing, power, and area, in addition to specifications for voltage levels and noise margins. By choosing the appropriate tradeoffs, values of R_L and W/L are selected that satisfy the design specifications across a wide range of process variations and operating conditions.

A new perspective on the VTC can be gained by plotting I versus V for the resistor and transistor on the same graph. Such a plot for the resistive-load inverter is

**Figure 4.9**

NMOS inverter, resistor load.

called a load line characteristic, as shown in Figure 4.9b. It shows the current through the resistor superimposed with the current through the MOS device. At one extreme, the current through the resistor is V_{DD}/R_L and at the other extreme it is zero. Since the current is linearly related to voltage, the load characteristic is linear from $V_{DS} = 0$ to V_{DD} . From a dc perspective, the inverter current will always lie on the load line. When the output is V_{OH} , there is no current flowing through the two devices. When the output is V_{OL} , the current through both devices is equal

$$I_{DS} = I_R = \frac{V_{DD} - V_{OL}}{R_L} \quad (4.4)$$

The VTC can be found from the load line construction by noting that each value of $V_{in} = V_{GS}$ for the transistor gives a different drain current curve in Figure 4.9b. The intersection of the transistor drain characteristic curve with the resistor load line

gives a value of $V_{\text{out}} = V_{DS}$. The plot of V_{out} versus V_{in} is the desired VTC as seen in Figure 4.9c.

While this graphical load line representation is instructive, we will generally not use this approach to determine the voltage transfer characteristics. Instead, we will formulate explicit solutions for each of the five critical points on the VTC of Figure 4.9c. Specifically, we will seek to find V_{OH} , V_{OL} , V_{IH} , V_{IL} , and V_S . This information is all that is required to sketch the VTC and determine the noise margins. To simplify the equations in the following sections, we assume that channel-length modulation factor, λ , is too small to have a significant effect on voltage transfer characteristics, although strictly speaking this is not true.

The value of V_{OH} can be obtained by setting the input voltage below the transistor threshold voltage V_T . Then no current flows, and the inverter output voltage V_{out} remains at V_{DD} . The nominal voltage representing a logic high level is

$$V_{OH} = V_{DD} \quad (4.5)$$

When a logic value of 1, represented by V_{OH} , is applied at the input of this inverter, the transistor is driven into the linear region of operation.⁴ The output low level, V_{OL} , is found by equating the currents in the transistor and load, assuming the input of this inverter is driven by V_{OH} :

$$I_R = I_{DS} (\text{lin}) \quad (4.6a)$$

Substituting in the expressions for the resistor current and NMOS current:

$$\frac{V_{DD} - V_{OL}}{R_L} = \frac{W_N}{L_N} \frac{\mu_n C_{ox}}{\left(1 + \frac{V_{OL}}{E_C L}\right)} [2(V_{OH} - V_T)V_{OL} - V_{OL}^2] \quad (4.6b)$$

The term in the denominator $V_{OL}/E_C L$ is small since V_{OL} is small. It can be neglected to first order to avoid iteration.⁵ Then, setting $k = (W/L) \mu_n C_{ox}$:

$$\frac{V_{DD} - V_{OL}}{R_L} = \frac{k}{2} [2(V_{OH} - V_T)V_{OL} - V_{OL}^2]$$

Restating this in terms of V_{OL} ,

$$V_{OL}^2 - 2 \left(\frac{1}{kR_L} + V_{DD} - V_T \right) V_{OL} + \frac{2V_{DD}}{kR_L} = 0 \quad (4.6c)$$

⁴ When such an assumption is made about the operating region of a device, it must always be checked at the end of the calculation to make sure that the assumption is correct.

⁵ It is important to be able to apply this type of engineering judgment to simplify potentially complicated equations.

This quadratic equation may be solved for V_{OL} . Only the positive root has physical significance. If V_{OL} is sufficiently small, the squared term may also be neglected to further simplify the solution. Using this approximation,

$$V_{OL} \approx \frac{V_{DD}}{1 + kR_L(V_{DD} - V_T)} \quad (4.6d)$$

Equation (4.6d) is instructive from a design perspective. Since we want to make V_{OL} small, we can accomplish this by increasing k (i.e., W/L), or increasing the resistance, R_L . In both cases, the decrease in V_{OL} is at the expense of area. Also, increasing R_L will make the rise time slower, while increasing k will make the fall time faster. As V_{OL} decreases with increasing k , we would also find that the power increases. The power, when the output is low, is given by

$$P = IV = \left(\frac{V_{DD} - V_{OL}}{R_L} \right) V_{DD} \quad (4.7)$$

The equation indicates that power will increase but only marginally. However, an increase in R_L will act to reduce the power consumption. The complex tradeoffs just described appear throughout the design process with timing, power, area, noise margins, etc. being traded-off against one another to achieve a given set of design specifications. Ultimately, the noise margins will dictate the choice of k and R_L for this type of inverter.

The noise margin voltages V_{IL} and V_{IH} are defined as the points at which

$$\frac{\partial V_{out}}{\partial V_{in}} = -1.0$$

The minus sign in this equation arises because V_{out} decreases as V_{in} increases for inverters. The VTC requires equal current in both devices. At $V_{in} = V_{IL}$, the output voltage is near V_{DD} and the transistor is operating in the saturation region. Thus,

$$I_R = I_{DS}(\text{sat}) \quad (4.8a)$$

By substituting the saturation current equation, we obtain

$$\frac{V_{DD} - V_{out}}{R_L} = \frac{Wv_{sat}C_{ox}(V_{in} - V_T)^2}{(V_{in} - V_T) + E_C L}$$

Since V_{in} is typically close to V_T when computing V_{IL} , we can simplify the saturation current expression by removing this term in the denominator:

$$\frac{V_{DD} - V_{out}}{R_L} \approx \frac{Wv_{sat}C_{ox}(V_{in} - V_T)^2}{E_C L}$$

Recalling that $v_{\text{sat}} = \mu E_C / 2$, a substitution into the above results in

$$\frac{V_{DD} - V_{\text{out}}}{R_L} = \frac{W\mu_n C_{ox}(V_{\text{in}} - V_T)^2}{2L}$$

Setting $k = (W/L)\mu_n C_{\text{OX}}$ and differentiating with respect to V_{in} ,

$$-\frac{1}{R_L} \left(\frac{\partial V_{\text{out}}}{\partial V_{\text{in}}} \right) = k(V_{\text{in}} - V_T)$$

Setting $V_{\text{in}} = V_{IL}$ and the gain term to -1 , we arrive at

$$k(V_{IL} - V_T)R_L = 1$$

This is easily solved for V_{IL} :

$$V_{IL} = V_T + \frac{1}{kR_L} \quad (4.8b)$$

It is interesting to note that V_{IL} is slightly above the threshold voltage, which makes sense from an intuitive standpoint. The n -channel device turns on at V_T and, at a slightly higher V_{GS} , the gain is equal to 1 in magnitude. To increase V_{IL} thereby increasing NM_L , we need to decrease k and R_L . This also increases V_{OL} , which is not desirable. Since both V_{IL} and V_{OL} shift in the same direction, it is difficult to affect NM_L significantly. Again, the two design parameters, R_L and W/L , will most likely be controlled by V_{OL} .

The value of V_{OUL} can be determined by substituting $V_{\text{in}} = V_{IL}$ into Equation (4.6b) and solving for V_{out} , if needed.

At the other unity gain point, where $V_{\text{in}} = V_{IH}$, the output voltage is near 0 V and the transistor is operating in the linear region. Therefore,

$$\frac{W}{L} \frac{\mu_n C_{ox}}{\left(1 + \frac{V_{\text{out}}}{E_C L}\right)} \left[(V_{\text{in}} - V_T)V_{\text{out}} - \frac{V_{\text{out}}^2}{2} \right] = \frac{V_{DD} - V_{\text{out}}}{R_L}$$

The denominator of the current expression has a term $V_{\text{out}}/E_C L$ that will be small in the computation of V_{IH} so it can be neglected. We use this simplification and write

$$k \left[(V_{\text{in}} - V_T)V_{\text{out}} - \frac{V_{\text{out}}^2}{2} \right] \approx \frac{V_{DD} - V_{\text{out}}}{R_L} \quad (4.8c)$$

Taking the partial derivative of this equation:

$$k \left[(V_{\text{in}} - V_T) \frac{\partial V_{\text{out}}}{\partial V_{\text{in}}} + V_{\text{out}} - V_{\text{out}} \frac{\partial V_{\text{out}}}{\partial V_{\text{in}}} \right] = -\frac{1}{R_L} \frac{\partial V_{\text{out}}}{\partial V_{\text{in}}}$$

Since $V_{in} = V_{IH}$ and the gain term is -1 , we obtain

$$V_{IH} = V_T + 2V_{out} - \frac{1}{kR_L} \quad (4.8d)$$

The explicit value of $V_{in} = V_{IH}$ and the corresponding value of V_{out} are found by substituting Equation (4.8d) into Equation (4.8c), namely,

$$\frac{k}{2}[2(V_{IH} - V_T)V_{out} - V_{out}^2] = \frac{V_{DD} - V_{out}}{R_L} \quad (4.8e)$$

Solving these two equations together produces:

$$V_{IH} = V_T + \sqrt{\frac{8V_{DD}}{3kR_L}} - \frac{1}{kR_L} \quad (4.8f)$$

By increasing kR_L , we can reduce V_{IH} and increase NM_H . However, this increases the rising delay and increases power slightly. Note that changes in k and R_L have no effect on V_{OH} , so it is possible to increase NM_H , if desired.

The value of V_{OUL} can be determined by substituting V_{IH} into Equation (4.8e) and solving for V_{out} , if needed.

An additional point on the VTC that is important is where $V_{in} = V_{out} = V_S$. At this operating point, the NMOS transistor is in saturation since $V_{DS} = V_{GS}$, which implies that $V_{DS} > V_{Dsat}$. The switching voltage is found by equating currents

$$\frac{Wv_{sat}C_{ox}(V_S - V_T)^2}{(V_S - V_T) + E_C L} = \frac{V_{DD} - V_S}{R_L} \quad (4.9)$$

Since V_S is typically in the midpoint of the voltage swing, it is possible to set up an iterative equation with an initial guess of $V_{DD}/2$ to obtain V_S . Alternatively, Equation (4.9) can be written as a quadratic equation and solved directly.

Once all five points have been computed, the VTC of the inverter can be quickly sketched.

Example 4.2 Noise Margins for R-Load Inverter

Problem:

Given the following data for an inverter like the one shown in Figure 4.9a, find the multi-source noise margins.

$$k' = \mu_n C_{ox} = 430 \text{ } \mu\text{A/V}^2, \quad V_T = 0.4 \text{ V}, \quad \frac{W}{L} = 2.0$$

$$V_{DD} = 1.2 \text{ V}, \quad R_L = 20 \text{ k}\Omega$$

Solution:

$$V_{OH} = V_{DD} = 1.2 \text{ V} \quad \text{and} \quad k = k' \left(\frac{W}{L} \right) = 430 \frac{\mu\text{A}}{\text{V}^2} \times 2.$$

Find V_{OL} by solving (4.6d) with the given data (assuming transistor is linear):

$$V_{OL} \approx \frac{V_{DD}}{1 + kR_L(V_{DD} - V_T)} = \frac{1.2}{1 + (430 \times 10^{-6})(2)(20 \times 10^3)(1.2 - 0.4)} \approx 0.08 \text{ V}$$

Note that V_{OL} is safely below V_T . Also, the assumption that the transistor is in the linear region is valid.

Find V_{IL} using (4.8b):

$$V_{IL} = \frac{1}{kR_L} + V_T = 0.46 \text{ V}$$

A check confirms that the transistor is in the saturation region, as assumed.

Find V_{IH} from (4.8f):

$$V_{IH} = V_T + \sqrt{\frac{8V_{DD}}{3kR_L}} - \frac{1}{kR_L} = 0.4 + \sqrt{\frac{8}{3} \frac{1.2}{(2)430(10^{-6})20(10^3)}} - \frac{1}{(2)430(10^{-6})20(10^3)} \approx 0.77 \text{ V}$$

Again, the transistor is in the linear region, as assumed. The noise margins are then

$$NM_L = V_{IL} - V_{OL} = 0.46 - 0.08 = 0.38 \text{ V}$$

$$NM_H = V_{OH} - V_{IH} = 1.2 - 0.77 = 0.43 \text{ V}$$

SPICE Comparison to Hand Analysis in 0.18 μm CMOS

Example 4.3

Problem:

Compare SPICE VTC results to hand calculations for a resistive-load NMOS inverter with $R_L = 30 \text{ k}\Omega$, $W = 400 \text{ nm}$, $L = 200 \text{ nm}$ (i.e., $W/L = 2$). That is, find V_{OH} , V_{OL} , V_{IH} , V_{IL} , and V_S . Useful parameters for 0.18 μm CMOS are as follows:

$$k' = \mu_n C_{ox} = 270 \text{ } \mu\text{A/V}^2, \quad C_{ox} = 1 \text{ } \mu\text{F/cm}^2, \quad V_T = 0.5 \text{ V}$$

$$E_C L = 1.2 \text{ V}, \quad V_{DD} = 1.8 \text{ V}, \quad v_{sat} = 8 \times 10^6 \text{ cm/s}$$

Solution:

The SPICE file would be as follows:

```
*Resistive-load inverter example
*Set supply and library
.param Supply=1.8                      *for 0.18 technology
.include 'bsim3v3.cmosp18.lib'
.opt scale=0.1u                            *Set lambda
*Power supply
.global Vdd Gnd
```

```

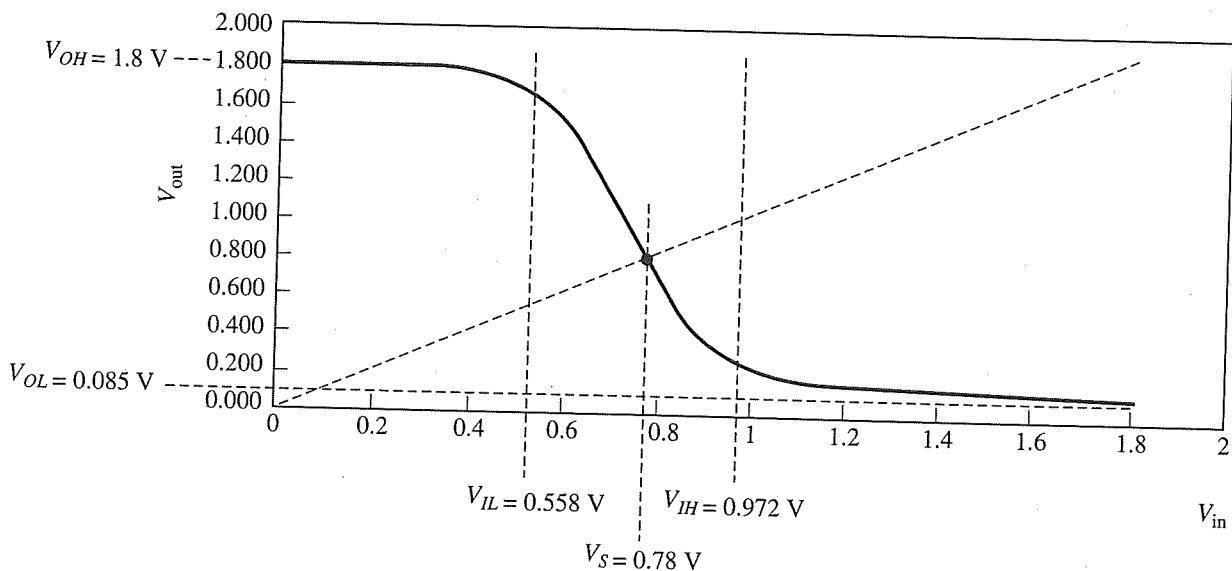
Vdd      Vdd Gnd 'Supply'          *Supply is set by .lib call
*Top level simulation netlist
Rup      Vdd out 30k
mn       out in  Gnd   Gnd NMOS1 l=2 w=4 ad=0 pd=0 as=0 ps=0
Vin      in  Gnd 'Supply'
*Simulation
.dc      Vin 0   'Supply'        'Supply/50'
.plot    dc   V(out)
.end
    
```

The SPICE file produces the following VTC. The graph is annotated with the key points. Using these values, the noise margins and switching threshold are

$$NM_L = V_{IL} - V_{OL} = 0.558 - 0.085 = 0.473 \text{ V}$$

$$NM_H = V_{OH} - V_{IH} = 1.8 - 0.972 = 0.828 \text{ V}$$

$$V_S = 0.78 \text{ V}$$



To compare against hand calculation, we first find $V_{OH} = V_{DD} = 1.8 \text{ V}$.
Next find V_{OL} :

$$V_{OL} \approx \frac{V_{DD}}{1 + kR_L(V_{DD} - V_T)} = \frac{1.8}{1 + (2)\left(270 \frac{\mu\text{A}}{\text{V}^2}\right)(30 \text{ k}\Omega)(1.8 - 0.5)} \approx 0.082 \text{ V}$$

Then find V_{IL} :

$$V_{IL} = \frac{1}{kR_L} + V_T = 0.06 + 0.5 = 0.56 \text{ V}$$

Find V_{IH} as follows:

$$V_{IH} = V_T + \sqrt{\frac{8V_{DD}}{3kR_L}} - \frac{1}{kR_L} = 0.5 + 0.54 - 0.06 = 0.98 \text{ V}$$

The noise margins are

$$N_{ML} = V_{IL} - V_{OL} = 0.478 \text{ V}$$

$$N_{MH} = V_{OH} - V_{IH} = 0.817 \text{ V}$$

These are quite close to the SPICE results.

The last step is to compute the switching threshold using Equation (4.9):

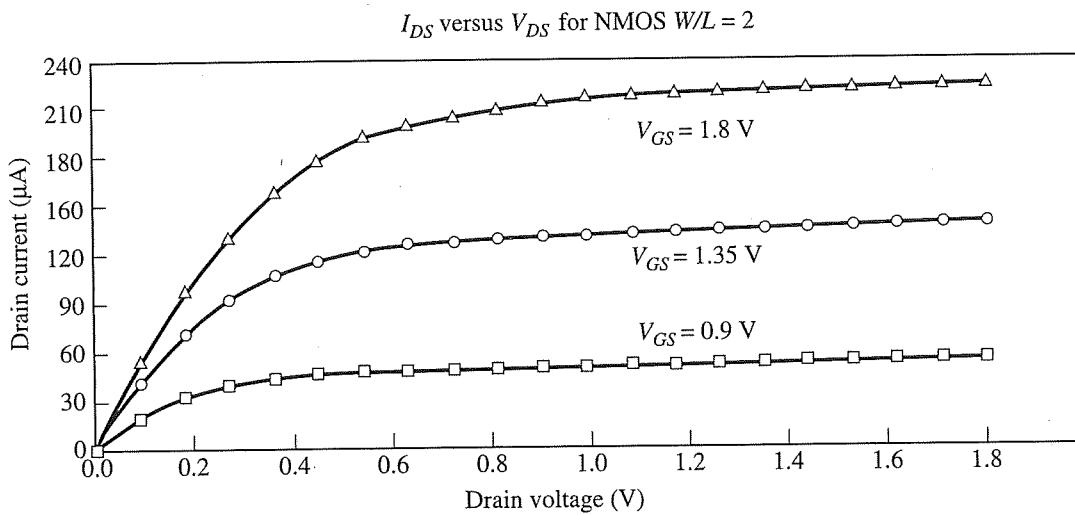
$$\frac{Wv_{sat}C_{ox}(V_S - V_T)^2}{(V_S - V_T) + E_C L} = \frac{V_{DD} - V_S}{R_L}$$

$$\frac{0.4(10^{-4})(8 \times 10^6)(10^{-6})(V_S - 0.5)^2}{(V_S - 0.5) + 1.2} = \frac{1.8 - V_S}{30 \text{ k}\Omega}$$

This equation can be rearranged and solved to produce

$$V_S = 0.89 \text{ V}$$

However, this value is about 15% higher than the SPICE result. Part of the reason is that the channel-length modulation parameter has been neglected in this calculation. Recall that the transistor is in saturation when computing V_S . By examining the SPICE curve shown below for $V_{GS} = 0.9 \text{ V}$, there is a noticeable increase in the current in saturation due to channel-length modulation and DIBL. If we included a term $(1 + \lambda V_S)$ in the current equation and decreased V_T due to DIBL, the actual saturation current would be larger. This would reduce the switching threshold to 0.78 V rather than the value of 0.89 V computed by hand.



4.5 NMOS Transistors as Load Devices

The resistor loads require a large amount of chip area if realized in a standard MOS process. In fact, the area would be more than 100 times that of the transistor. It would also produce a large rise time. For these reasons, conventional resistors are rarely used as loads in MOS digital circuits. The following section describes several alternative ways of using NMOS transistors to perform the function of a pull-up resistor.

4.5.1 Saturated Enhancement Load

A single NMOS transistor with the gate connected to the drain can be used as a load device, as shown in Figure 4.10a. The purpose of this device is to pull up the output node to V_{OH} , so it is referred to as the pull-up (or load) device. This configuration is called a saturated enhancement load inverter since $V_{GS} = V_{DS}$. As a result, this load transistor can operate only in saturation or cutoff. Note that the body is

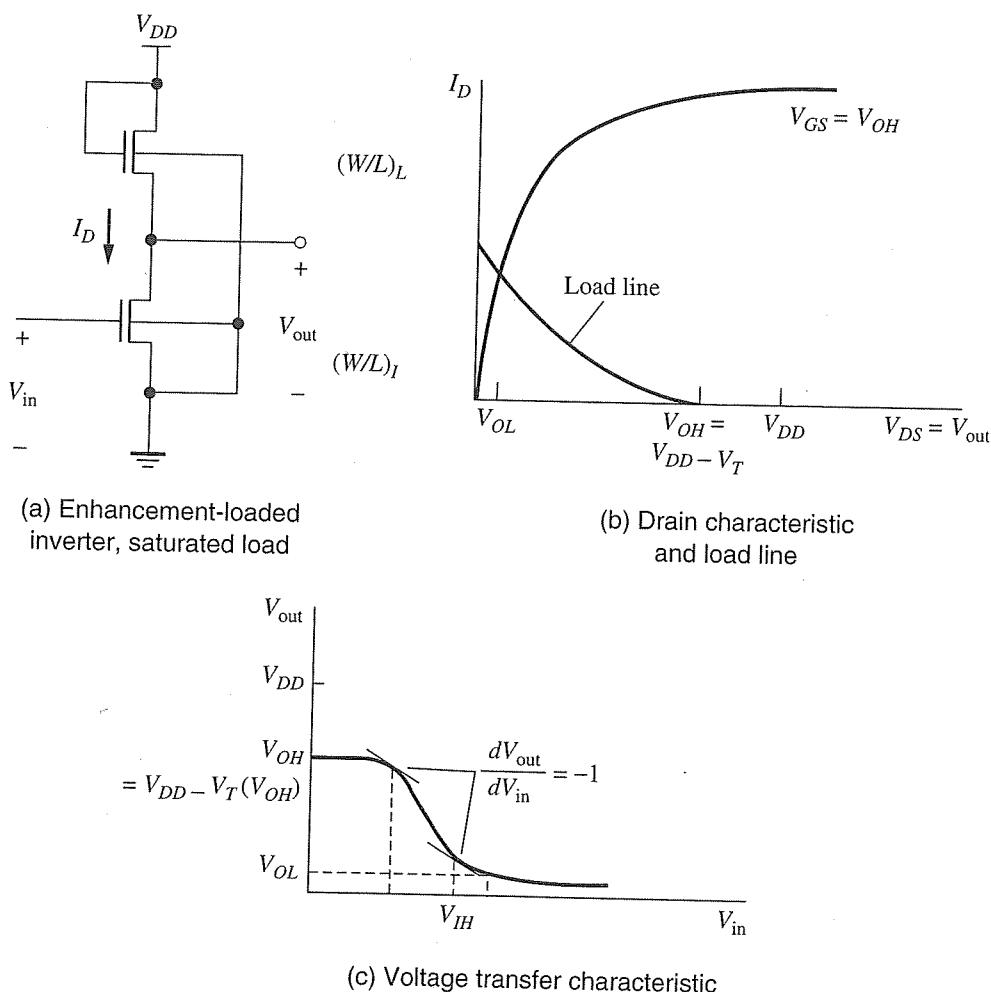


Figure 4.10

NMOS inverter, saturated enhancement load.

grounded because it is usually common to all transistors in a single well. Therefore, this transistor will experience body effect.

The other NMOS device pulls down on the output node to try to discharge it and, therefore, is referred to as the pull-down (or inverting) device. The relative sizes of the two transistors determine the output voltages, and so it is called a *ratioed* inverter. The importance of the device ratio is that it determines whether or not the inverter will function properly. If, for example, the sizes are chosen such that V_{OL} is much higher than the threshold voltage, then it will not operate as an inverter. In this sense, a ratioed inverter requires careful sizing of the pull-up and pull-down devices for proper operation. Conversely, a *ratioless* inverter does not require a specific ratio between the pull-up and pull-down devices to operate properly.

The ratio K_R for a MOS inverter⁶ is defined as

$$K_R = \frac{k_{\text{invert}}}{k_{\text{load}}} = \frac{k'(W/L)_I}{k'(W/L)_L} = \frac{(W/L)_I}{(W/L)_L} \quad (4.10)$$

Subscripts I and L are added to distinguish between inverting (I) and load (L) devices. The value of K_R in ratioed inverters determines the value of V_{OL} . In a well-designed inverter, V_{OL} will be about 5% of V_{DD} . If the value of K_R is large, V_{OL} is small. It should also be clear that increasing the value of K_R increases the circuit area.

The load line construction and the VTC for the inverter with the saturated load are shown in Figure 4.10b and c. A careful examination of the VTC curve reveals a serious deficiency of the saturated enhancement load that has been overlooked so far. The output high level V_{OH} is not equal to V_{DD} as it was for the resistor load. The pull-up transistor ceases to conduct after its gate-source voltage decreases to the threshold voltage. The result is that the output node does not rise above $V_{DD} - V_{TL}$. The threshold voltage of the load device, V_{TL} , is no longer V_{T0} as it is for the pull-down device, because the full output voltage appears as a body bias between source and body of the load device. The threshold voltage is given by the equation derived in Chapter 2:

$$V_{TL} = V_{T0} + \gamma(\sqrt{V_{SB} + 2|\phi_F|} - \sqrt{2|\phi_F|}) \quad (4.11)$$

where $V_{SB} = V_{OH}$. The resulting value of V_{OH} is one threshold voltage below the supply voltage. If $V_{DD} = 1.2$ V and $V_{TL} = 0.45$ V (including body-bias), then $V_{OH} = 1.2 - 0.45 = 0.75$ V. This may be too small to be acceptable as a high input to the next gate, making it difficult to design circuits that will operate with safe noise margins on a low supply voltage. In fact, this is one reason why early MOS circuits required higher voltage supplies.

Show that the pull-up device of Figure 4.10a is indeed in saturation or cutoff during normal operation.

Exercise 4.2

⁶ Some books use the term *beta ratio* (β_R) in place of the K_R . We avoid this terminology to avoid confusion with the current gain *beta* of a bipolar transistor.

The VTC for the saturated load inverter is computed as before. Here we will illustrate the approach to obtain only V_{OH} and V_{OL} . First, to find V_{OH} recall that the output node can only rise to a threshold drop below V_{DD} before the load device ceases to conduct. Therefore, we can write

$$\begin{aligned} V_{OH} &= V_{DD} - V_T(V_{OH}) \\ &= V_{DD} - [V_{T0} + \gamma(\sqrt{V_{OH} + 2|\phi_F|} - \sqrt{2|\phi_F|})] \end{aligned} \quad (4.12)$$

To obtain the solution, use a substitute variable, $x = \sqrt{V_{OH} + 2|\phi_F|}$, and solve the quadratic equation in x . Only the positive root is meaningful from a physical standpoint. The fact that the output voltage cannot rise above $V_{DD} - V_{TL}$ adversely affects the NM_H value. The reduction is so severe that it is not a suitable load for most logic gates in a CMOS design. However, we may occasionally encounter this type of load due to the operating voltages of a given circuit.

Example 4.4 Computation of V_{OH} for a Saturated Enhancement Load Inverter

Problem:

Calculate the V_{OH} of the NMOS inverter shown in Figure 4.10a. Assume the following parameters for a $0.13\text{ }\mu\text{m}$ technology. Comment on the magnitude of the result:

$$2|\phi_F| = 0.88\text{ V}, \quad \gamma = 0.2\text{ V}^{1/2}, \quad V_T = 0.4\text{ V}, \quad V_{DD} = 1.2\text{ V}$$

Solution:

$$\begin{aligned} V_{OH} &= V_{DD} - V_{TL} = V_{DD} - V_{TL}(V_{SB}) \\ &= V_{DD} - [V_{T0} + \gamma(\sqrt{V_{SB} + 2|\phi_F|} - \sqrt{2|\phi_F|})] \\ &= 1.2 - [0.4 + 0.2(\sqrt{V_{OH} + 2|-0.44|} - \sqrt{2|-0.44|})] \\ &= 0.8 - 0.2\sqrt{V_{OH} + 0.88} + 0.2\sqrt{0.88} \\ &= 0.99 - 0.2\sqrt{V_{OH} + 0.88} \end{aligned}$$

Solve by iteration, starting at the (optimistic) value $V_{OH} = 1.2\text{ V}$.

Old V_{OH}	New V_{OH}
1.20	0.70
0.70	0.74
0.74	0.73
0.73	0.73

$$\therefore V_{OH} = 0.73\text{ V}$$

Clearly, a V_{OH} of 0.73 V from a 1.2 V supply is not acceptable as input to the next gate. Also, note that faster convergence would have been obtained had we started with $V_{OH} = 1.2 - 0.4 = 0.8\text{ V}$.

To find V_{OL} , we equate the currents in the two transistors. V_{OL} is the output voltage from this inverter when its input voltage is V_{OH} , the output high level of another identical inverter. For correct results, we need to know which form of the drain current equation (e.g., linear or saturated) to use for each transistor.

When conducting, the load transistor in this circuit is always saturated since $V_{DS} = V_{GS}$. As for the inverting transistor, with $V_{GS} = V_{OH}$, the output voltage should be lower than V_{T0} if there is to be any noise margin. Therefore, our assumption is that the pull-down transistor is in the linear region. Now equate the drain currents, using the appropriate equations:

$$I_{DI}(\text{lin}) = I_{DL}(\text{sat})$$

We can substitute in the current equations to obtain

$$\frac{W_I}{L_I} \frac{\mu_n C_{ox}}{\left(1 + \frac{V_{out}}{E_{CN} L_I}\right)} \left[(V_{in} - V_{Tl}) V_{out} - \frac{V_{out}^2}{2} \right] = \frac{W_L v_{sat} C_{ox} (V_{DD} - V_{out} - V_{TL})^2}{(V_{DD} - V_{out} - V_{TL}) + E_{CN} L_L}$$

This expression is used whenever the pull-up is saturated and the pull-down is in the linear region. But note that the load transistor threshold voltage V_{TL} is modified by the body effect, due to the voltage V_{out} , which appears between its source and the (grounded) substrate. When Equation (4.11) is used to evaluate V_{TL} (V_{out}) in the above, the resulting expression is a high-order polynomial. Direct solution by hand to find V_{OL} is tedious and susceptible to error. Simulation using SPICE or an equivalent computer program is a desirable alternative. If hand analysis is necessary, choose several values of V_{out} , substitute in the above equation, and solve the resulting linear first-order equation in V_{in} . Fortunately, V_{out} is small and will not have a significant effect on V_{TL} .

One other shortcoming of this type of inverter is similar to the resistive-load inverter: there is dc power dissipation when the output is low. Since current flowing directly from V_{DD} to Gnd when the output is at V_{OL} , the power dissipation is significant, especially if many such gates exist in the logic design. This is another reason why this type of inverter is rarely used.

Design of Saturated Load Inverter

Example 4.5

Problem:

Design the saturated enhancement load inverter of Figure 4.10a such that it delivers a low output voltage of $V_{OL} = 0.1$ V when the input is V_{DD} . Assume a $0.13\ \mu\text{m}$ technology, with $L = 100\ \text{nm}$ and let the pull-up device be of unit size. Use the parameters

$$\mu_n = 270\ \text{cm}^2/\text{V-s}, \quad C_{ox} = 1.6\ \mu\text{F}/\text{cm}^2, \quad V_{T0} = 0.4\ \text{V}, \quad V_{DD} = 1.2\ \text{V}$$

$$E_{CN}L = 0.6\ \text{V}, \quad v_{sat} = 8 \times 10^6\ \text{cm/s}, \quad \gamma = 0.2\ \text{V}^{1/2}$$

Solution:

Use the current equation for the pull-down in the linear region and the pull-up in the saturation region.

$$\frac{W_I}{L_I} \frac{\mu_n C_{ox}}{1 + \frac{V_{out}}{E_{CN} L_I}} \left[(V_{in} - V_{TL}) V_{out} - \frac{V_{out}^2}{2} \right] = \frac{W_L v_{sat} C_{ox} (V_{DD} - V_{out} - V_{TL})^2}{(V_{DD} - V_{out} - V_{TL}) + E_{CN} L_L}$$

We can eliminate C_{ox} and then set $V_{in} = V_{DD}$ (we should really set it to V_{OH} but we keep things simple in this problem) and $V_{out} = V_{OL}$:

$$\therefore \frac{W_I}{L_I} \frac{\mu_n}{1 + \frac{V_{OL}}{E_{CN} L_I}} \left[(V_{DD} - V_{TL}) V_{OL} - \frac{V_{OL}^2}{2} \right] = \frac{W_L v_{sat} (V_{DD} - V_{OL} - V_{TL})^2}{(V_{DD} - V_{OL} - V_{TL}) + E_{CN} L_L}$$

Since we know that the output voltage is $V_{OL} = 0.1$ V, we can compute the correct V_{TL} :

$$\begin{aligned} V_T &= [V_{TO} + \gamma(\sqrt{V_{SB} + 2|\phi_F|} - \sqrt{2|\phi_F|})] \\ &= [0.4 + 0.2(\sqrt{0.1 + 2|-0.44|} - \sqrt{2|-0.44|})] \\ &= 0.4 + 0.2(\sqrt{0.1 + 0.88} - \sqrt{0.88}) \\ &= 0.41 \text{ V} \end{aligned}$$

The degree of body effect is small, as one would expect with a body-bias of 0.1 V. We can substitute the known quantities and solve for the transistor sizes:

$$\begin{aligned} \therefore \frac{W_I}{0.1(10^{-4})\text{cm}} \frac{(270 \text{ cm}^2/\text{V}\cdot\text{s})}{(1 + (0.1/0.6))} \left[(1.2 - 0.4)0.1 - \frac{0.1^2}{2} \right] \\ = \frac{W_L (8 \times 10^6 \text{ cm/s}) (1.2 - 0.1 - 0.41)^2}{(1.2 - 0.1 - 0.41) + 0.6} \\ \therefore K_R = \frac{W_I}{W_L} = \frac{2.95}{1.75} = 1.7 \end{aligned}$$

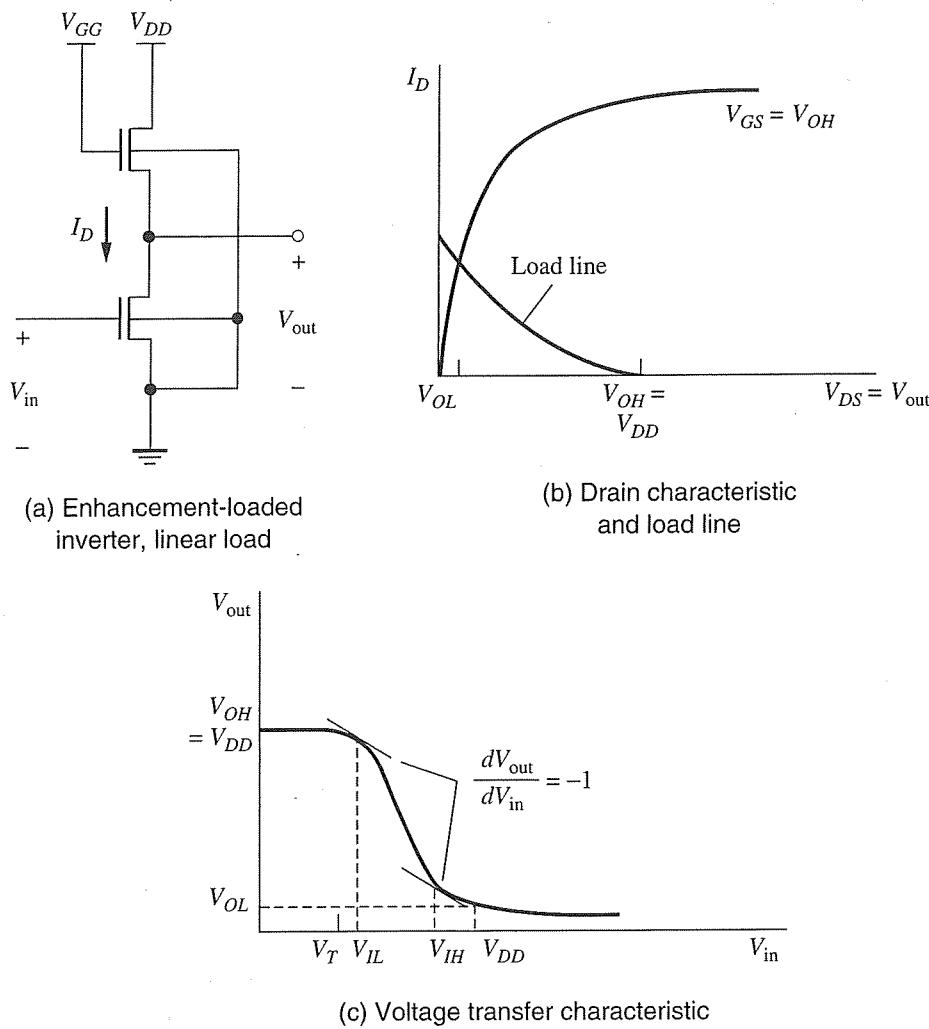
If $W_L = 100$ nm, then $W_I = 170$ nm.

4.5.2 Linear Enhancement Load

The output high level V_{OH} of the previous example can be increased simply by connecting the gate of the load transistor to a dc voltage of V_{GG} , which is sufficiently greater than V_{DD} . This is shown in Figure 4.11a. The desirable value for V_{GG} is given by

$$V_{GG} > V_{DD} + V_{TL}(V_{DD}) \quad (4.13)$$

where the last term is simply the value of V_{TL} with a body bias of V_{DD} . When this condition is met, the load device can pull the output all the way to V_{DD} . The pull-up device, when considering the quadratic model equations, operates in the linear

**Figure 4.11**

NMOS inverter, *linear* enhancement load.

region over the entire range of V_{out} , since throughout this range the load device is operating with $V_{DS} < V_{GS} - V_{TL}$. For this reason, the names *linear load*, *nonsaturated load*, or *triode load* were applied to this mode of operation. With the velocity saturation model, this name no longer applies. The device can be in the linear region or the saturation region, depending on the output voltage; in fact, when the output is low, it is saturated. However, we will use the term *linear* to reflect its original name, even though this name is a misnomer for short-channel devices.

Figure 4.11b shows the load line construction for this type of inverter. The linear enhancement load has several disadvantages when used in static inverters and gates. More chip area is required, since an extra voltage source V_{GG} with associated additional interconnections on the chip is needed. The required value of K_R is even larger than for a saturated enhancement load. Furthermore, there is dc power dissipation in the output low state. These types of inverters are rarely, if ever, used as stand-alone gates. However, the operating conditions of the circuit may occasionally create this type of situation.

4.6 Complementary MOS (CMOS) Inverters

The inverters discussed thus far have certain inherent limitations in terms of noise margin and power dissipation. Ideally, it is desirable to have a pull-up device that is complementary in behavior to the pull-down device. The reader, by now, should be aware that the CMOS inverter overcomes many of the problems cited above for other logic families. We now explore the features of CMOS inverters that form the basis of most gates used in IC logic design today.

4.6.1 DC Analysis of CMOS Inverter

Figure 4.12 shows a CMOS inverter along with its load line construction and VTC. In Figure 4.12a, the gate of the upper PMOS device is connected to the gate of the

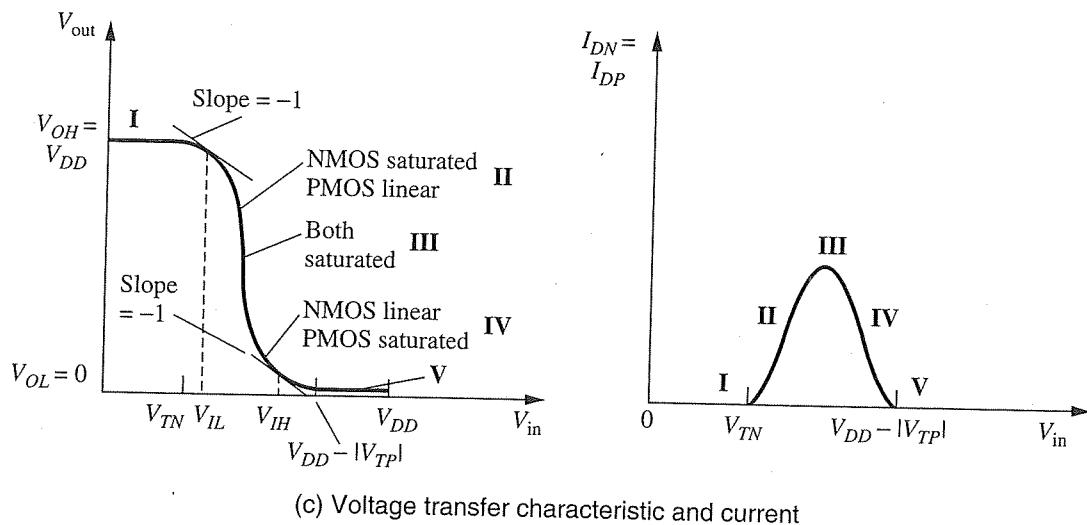
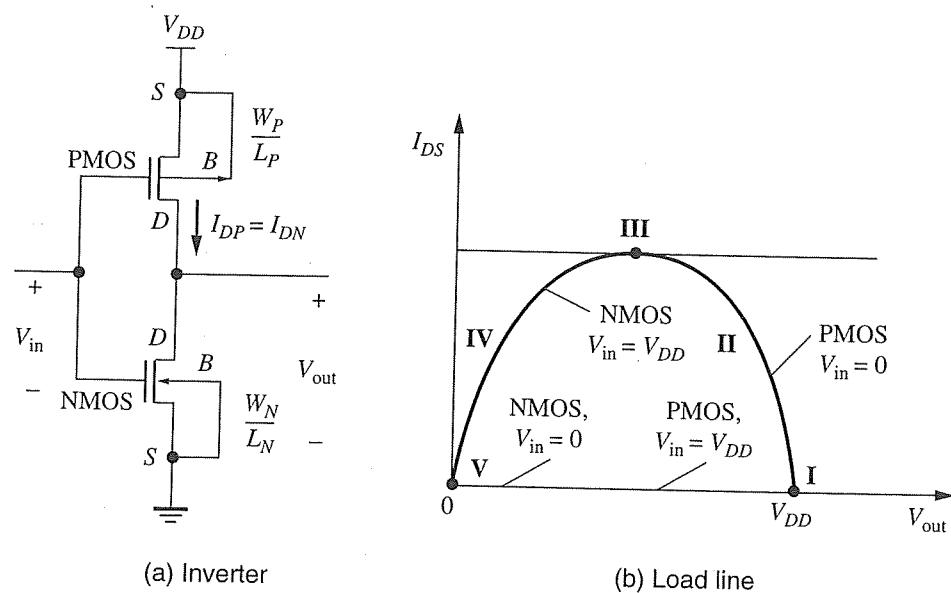


Figure 4.12

CMOS inverter characteristics.

lower NMOS device. The PMOS device has a substrate connection to V_{DD} while the NMOS body is connected to Gnd. This is required to keep the source/drain junctions reverse-biased.

The operation of the CMOS inverter is as follows. When the input is at V_{DD} , the NMOS device is conducting while the PMOS device (which has $V_{GS} = 0$ V) is cut off. The result is that V_{out} is 0 V. Hence the drain current of the NMOS is limited to the very small leakage current of the PMOS, even though a highly conductive channel is present in the NMOS device. When the input is at ground, the NMOS device is cut off while the PMOS device is conducting. Only the small leakage current of the NMOS device can flow, so V_{out} is equal to V_{DD} . Therefore, we find that $V_{OH} = V_{DD}$ and $V_{OL} = 0$ V, which implies large noise margins for CMOS inverters. A completely symmetrical VTC is obtained if $V_{TP} = -V_{TN}$ and $k_p = k_n$.

The gate is static in that there is always a conducting path from the output to V_{DD} or Gnd. The only currents that flow in steady state are the subthreshold currents of the nominally off device and leakage currents of reverse-biased pn junctions. Quiescent power dissipation is in the nanowatt range. Although no steady-state current flows, the *on* transistor supplies current to drive any load resistance or capacitance whenever the output voltage differs from 0 V or V_{DD} . The tiny steady-state power consumption of CMOS in either logic state is its most attractive feature. The resistive-load and NMOS-only circuits all suffer from power dissipation in the output low state and therefore are inappropriate for use in high-density deep sub-micron designs.

The load line construction for the CMOS inverter requires some explanation. Shown in Figure 4.12b are I_{DS} versus V_{out} curves for each transistor. When V_{GS} is equal in magnitude to V_{DD} for either device, the drain current curves are given by the curves on Figure 4.12b. When gate-source voltage is zero for either device, drain current is zero. Thus, the two devices behave almost as ideal loads for each other. The two logic states lie at $V_{out} = 0$ and $V_{out} = V_{DD}$ along the horizontal V_{out} axis, both at $I_{DS} = 0$.

The drain current of an *n*-channel device can be computed if the polarities in Figure 4.12a are observed. In normal operation of a *p*-channel device, V_{GS} , V_{DS} , and V_{TP} all have negative values. Drain current I_{DS} , defined as flowing *out* of the drain for PMOS, has a positive value. Another way to handle the change of signs for PMOS analysis is simply to use absolute values of all voltages in the drain current equations. If you use this approach, check to be sure that the PMOS devices are properly biased for normal operation. The gate must be negative with respect to source by more than a threshold voltage if current is to flow.

The complete VTC and noise margins can now be determined. The form of the VTC is shown in Figure 4.12c. It is useful to divide the curve into five regions as shown in the figure. Sequencing from region I to region V is carried out by starting with the input voltage V_{in} at 0 V and increasing its value until V_{DD} is reached.

Region I: *n*-off, *p*-lin. The NMOS device is off while the PMOS device is linear, although it has a $V_{DS} = 0$. The output remains at $V_{OH} = V_{DD}$ up to the point when $V_{in} = V_{TN}$.

Region II: *n*-sat, *p*-lin. The NMOS device is on and in saturation, while the PMOS device is still in the linear region since its V_{DS} is relatively small. The current equation in this region is given by

$$I_{DN}(\text{sat}) = I_{DP}(\text{lin})$$

Therefore,

$$\begin{aligned} & \frac{W_N v_{\text{sat}} C_{\text{ox}} (V_{\text{in}} - V_{TN})^2}{(V_{\text{in}} - V_{TN}) + E_{CN} L_N} \\ &= \frac{W_P}{L_P} \frac{\mu_p C_{\text{ox}}}{\left(1 + \frac{V_{DD} - V_{\text{out}}}{E_{CP} L_P}\right)} \left[(V_{DD} - V_{\text{in}} - |V_{TP}|)(V_{DD} - V_{\text{out}}) - \frac{(V_{DD} - V_{\text{out}})^2}{2} \right] \end{aligned}$$

The VTC in this region is computed by setting V_{in} and solving for V_{out} .

Region III: *n*-sat, *p*-sat. This involves the nearly vertical segment of the VTC where both transistors are saturated. By equating the saturated drain current equations for the two transistors, we obtain

$$I_{DN}(\text{sat}) = I_{DP}(\text{sat})$$

Therefore,

$$\frac{W_N v_{\text{sat}} C_{\text{ox}} (V_{\text{in}} - V_{TN})^2}{(V_{\text{in}} - V_{TN}) + E_{CN} L_N} = \frac{W_P v_{\text{sat}} C_{\text{ox}} (V_{DD} - V_{\text{in}} - |V_{TP}|)^2}{(V_{DD} - V_{\text{in}} - |V_{TP}|) + E_{CP} L_P}$$

Note that V_{out} is not present in this expression since the gain is (ideally) infinite due to modeling limitations of the current equations. The voltage V_S at which $V_{\text{in}} = V_{\text{out}}$ falls within this segment of the VTC. The value can be determined by setting V_{in} to V_S in the above expression:

$$\frac{W_N v_{\text{sat}} C_{\text{ox}} (V_S - V_{TN})^2}{(V_S - V_{TN}) + E_{CN} L_N} = \frac{W_P v_{\text{sat}} C_{\text{ox}} (V_{DD} - V_S - |V_{TP}|)^2}{(V_{DD} - V_S - |V_{TP}|) + E_{CP} L_P}$$

To solve for V_S is not straightforward, so a few simplifications are needed. Since V_S is expected to be approximately $V_{DD}/2$, the leading term in the denominator of both sides of the equation is smaller than the $E_C L$ term. To greatly simplify the equation, we remove these terms from both sides and cancel out $v_{\text{sat}} C_{\text{ox}}$ to obtain

$$\frac{W_N (V_S - V_{TN})^2}{E_{CN} L_N} = \frac{W_P (V_{DD} - V_S - |V_{TP}|)^2}{E_{CP} L_P}$$

Taking the square root of both sides, we solve for V_S :

$$V_S = \frac{V_{DD} - |V_{TP}| + \chi V_{TN}}{1 + \chi} \quad (4.14)$$

where

$$\chi = \sqrt{\frac{\frac{W_N}{E_{CN}L_N}}{\frac{W_P}{E_{CP}L_P}}} = \sqrt{\frac{\mu_n W_N}{\mu_p W_P}} \quad (4.15)$$

The switching threshold is very useful when sketching the VTC. As we increase χ , the switching threshold shifts left while decreasing the value shifts the switching threshold to the right. We can shift the VTC to the left by increasing the size of the NMOS device, or shift it to the right by increasing the size of the PMOS device. This shift in VTC produces a *skewed* inverter.

Computation of V_S for CMOS Inverter in 0.13 μm Technology

Example 4.6

Problem:

Calculate the switching voltage of a CMOS inverter with each the following PMOS sizes: $W_P = 400 \text{ nm}$ and $W_P = 100 \text{ nm}$. W_N is the same size of 100 nm for both cases. How does change in W_P affect the switching voltage?

Solution:

Compute V_S using Equations (4.14) and (4.15):

$$W_N = 0.1 \text{ } \mu\text{m} \quad W_P = 0.4 \text{ } \mu\text{m}$$

$$\chi = \sqrt{\frac{\frac{W_N}{E_{CN}L_N}}{\frac{W_P}{E_{CP}L_P}}} = \sqrt{\frac{W_N E_{CP}}{W_P E_{CN}}} = \sqrt{\frac{(0.1)(24)}{(0.4)(6)}} = 1.0$$

$$V_S = \frac{0.8 + (0.4)1.0}{1 + 1.0} = 0.6 \text{ V}$$

Re-computing for the second case:

$$W_N = 0.1 \text{ } \mu\text{m} \quad W_P = 0.1 \text{ } \mu\text{m}$$

$$\chi = \sqrt{\frac{\frac{W_N}{E_{CN}L_N}}{\frac{W_P}{E_{CP}L_P}}} = \sqrt{\frac{W_N E_{CP}}{W_P E_{CN}}} = \sqrt{\frac{(0.1)(24)}{(0.1)(6)}} = 2.0$$

$$V_S = \frac{0.8 + (0.4)2.0}{1 + 2.0} = 0.53 \text{ V}$$

We find that lowering W_P decreases V_S which shifts the VTC to the left.

We now continue on with the region-wise analysis.

Region IV: *n*-lin, *p*-sat. The NMOS device is on and in the linear region since V_{DS} is small, while the PMOS device is still in saturation. The current equation in this region is given by

$$I_{DN}(\text{lin}) = I_{DP}(\text{sat})$$

Therefore,

$$\begin{aligned} \frac{W_N}{L_N} \frac{\mu_n C_{ox}}{1 + \frac{V_{out}}{E_{CN} L_N}} & \left[(V_{in} - V_{TN}) V_{out} - \frac{V_{out}^2}{2} \right] \\ &= \frac{W_P v_{sat} C_{ox} (V_{DD} - V_{in} - |V_{TP}|)^2}{(V_{DD} - V_{in} - |V_{TP}|) + E_{CP} L_P} \end{aligned}$$

The VTC in this region is computed by setting V_{in} and solving for V_{out} .

Region V: *n*-lin, *p*-off. The applied input voltage is above $V_{DD} - |V_{TP}|$, so the output is now 0 V and will remain that way up to and beyond $V_{in} = V_{DD}$.

It is instructive to examine the transfer characteristic of the dc current versus voltage, as shown in Figure 4.12c. Here, current does not flow except in the range $V_{TN} < V_{in} < V_{DD} - |V_{TP}|$, again reinforcing the advantage of the CMOS inverter in terms of power dissipation. The peak current occurs in region III with both devices in saturation.

The noise margin calculations begin with the observations that $V_{OH} = V_{DD}$ and $V_{OL} = 0$ V. Next, V_{IL} and V_{IH} are determined from the $\partial V_{out}/\partial V_{in} = -1$ condition. At $V_{in} = V_{IL}$, the inverter is in region II and so the NMOS device is saturated while the PMOS device is in the linear region:

$$I_{DN}(\text{sat}) = I_{DP}(\text{lin})$$

Applying the two current equations to the above, we obtain

$$\begin{aligned} \frac{W_N v_{sat} C_{ox} (V_{in} - V_{TN})^2}{(V_{in} - V_{TN}) + E_{CN} L_N} &= \frac{W_P}{L_P} \frac{\mu_p C_{ox}}{1 + \frac{V_{DD} - V_{out}}{E_{CP} L_P}} \\ &\times \left[(V_{DD} - V_{in} - |V_{TP}|)(V_{DD} - V_{out}) - \frac{(V_{DD} - V_{out})^2}{2} \right] \end{aligned}$$

This equation looks somewhat cumbersome but can be greatly simplified. First, note that the term $(V_{in} - V_{TN})$ in the denominator of the *n*-channel device is small

relative to $E_{CN}L_N$ since the input voltage is slightly above V_{TN} . Second, the expression $(V_{DD} - V_{out})/E_{CP}L_P$ in the denominator of the *p*-channel device is also small, since V_{out} is around V_{DD} . These two terms are negligible when the inverter operating point is in region II around V_{IL} . This results in the following equation:

$$\frac{W_N v_{sat} C_{ox} (V_{in} - V_{TN})^2}{E_{CN} L_N} = \frac{W_P}{L_P} \mu_p C_{ox} \times \left[(V_{DD} - V_{in} - |V_{TP}|)(V_{DD} - V_{out}) - \frac{(V_{DD} - V_{out})^2}{2} \right]$$

Recalling that $v_{sat} = \mu_n E_{CN}/2$ and substituting into the above produces

$$\frac{W_N}{L_N} \frac{\mu_n C_{ox}}{2} (V_{in} - V_{TN})^2 = \frac{W_P}{L_P} \mu_p C_{ox} \times \left[(V_{DD} - V_{in} - |V_{TP}|)(V_{DD} - V_{out}) - \frac{(V_{DD} - V_{out})^2}{2} \right] \quad (4.16)$$

Essentially, we are back to the first-order MOS model equations for this calculation. To solve for V_{IL} , the equation above must be differentiated with respect to V_{in} :

$$\frac{W_N}{L_N} \mu_n (V_{in} - V_{TN}) = \frac{W_P}{L_P} \mu_p \left[(V_{DD} - V_{in} - |V_{TP}|) \left(-\frac{\partial V_{out}}{\partial V_{in}} \right) + (V_{DD} - V_{out})(-1) - (V_{DD} - V_{out}) \left(-\frac{\partial V_{out}}{\partial V_{in}} \right) \right]$$

After substituting in k_N and k_P , and setting $\partial V_{out}/\partial V_{in} = -1$ and $V_{in} = V_{IL}$, the result is

$$V_{IL} = \frac{2V_{out} - V_{DD} - |V_{TP}| + (k_N/k_P)(V_{TN})}{1 + (k_N/k_P)} \quad (4.17)$$

This equation is dependent on V_{out} , which is still unknown. The next step is to solve Equations (4.16) and (4.17) together to obtain V_{IL} and the corresponding $V_{out} = V_{OUL}$.

A similar process is used to obtain the equations for V_{IH} . At $V_{in} = V_{IH}$, the inverter is most likely in region IV and, therefore, the NMOS device is linear while the PMOS device is saturated.

$$I_{DN}(\text{lin}) = I_{DP}(\text{sat})$$

Applying the two current equations, we obtain:

$$\frac{W_N}{L_N} \frac{\mu_n C_{ox}}{1 + \frac{V_{out}}{E_{CN} L_N}} \left[(V_{in} - V_{TN}) V_{out} - \frac{V_{out}^2}{2} \right] = \frac{W_P v_{sat} C_{ox} (V_{DD} - V_{in} - |V_{TP}|)^2}{(V_{DD} - V_{in} - |V_{TP}|) + E_{CP} L_P}$$

This equation can also be greatly simplified. Again, note that the term $V_{out}/E_{CN} L_N$ in the denominator of the *n*-channel device is small and the expression $(V_{DD} - V_{in} - |V_{TP}|)$ in the denominator of the *p*-channel device is also small relative to $E_{CP} L_P$. These two terms are negligible when the inverter operating point is in region IV around V_{IH} . This results in the following equation:

$$\frac{W_N}{L_N} \mu_n C_{ox} \left[(V_{in} - V_{TN}) (V_{out}) - \frac{(V_{out})^2}{2} \right] = \frac{W_P v_{sat} C_{ox} (V_{DD} - V_{in} - |V_{TP}|)^2}{E_{CP} L_P}$$

Noting that $v_{sat} = \mu_p E_{CP}/2$ and substituting into the above produces

$$\frac{W_N}{L_N} \mu_n C_{ox} \left[(V_{in} - V_{TN}) (V_{out}) - \frac{(V_{out})^2}{2} \right] = \frac{W_P}{L_P} \frac{\mu_p C_{ox}}{2} (V_{DD} - V_{in} - |V_{TP}|)^2 \quad (4.18)$$

Again, we are back to the first-order MOS model equations. To solve for V_{IH} , the equation above must be differentiated with respect to V_{in} :

$$\begin{aligned} \frac{W_N}{L_N} \mu_n C_{ox} & \left[(V_{in} - V_{TN}) \frac{\partial V_{out}}{\partial V_{in}} + (V_{out}) - V_{out} \frac{\partial V_{out}}{\partial V_{in}} \right] \\ &= \frac{W_P}{L_P} \mu_p C_{ox} (V_{DD} - V_{in} - |V_{TP}|) (-1) \end{aligned}$$

After substituting in k_N and k_P , and setting $\partial V_{out}/\partial V_{in} = -1$ and $V_{in} = V_{IH}$, this yields

$$V_{IH} = \frac{2V_{out} + V_{TN} + (k_p/k_N)(V_{DD} - |V_{TP}|)}{1 + (k_p/k_N)} \quad (4.19)$$

This expression is substituted in the above drain current equation to obtain explicit solutions for V_{IH} and the corresponding value of $V_{out} = V_{OEH}$.

Example 4.7

SPICE Analysis and Hand Calculation of CMOS Inverter in 0.18 μm Technology

Problem:

Compare SPICE analysis and hand analysis for a CMOS inverter of Figure 4.12a, with $W_N = 400$ nm and $W_P = 800$ nm, for the VTC parameters. Use 0.18 μm technology parameters as follows:

$$\begin{aligned}\mu_n &= 270 \text{ cm}^2/\text{V-s}, \quad \mu_p = 70 \text{ cm}^2/\text{V-s}, \\ C_{ox} &= 1.0 \text{ } \mu\text{F/cm}^2, \quad V_{TN} = 0.5 \text{ V}, \quad V_{TP} = -0.5 \text{ V} \\ E_{CN}L &= 1.2 \text{ V}, \quad E_{CP}L = 4.8 \text{ V}, \quad v_{sat} = 8 \times 10^6 \text{ cm/s}, \\ V_{DD} &= 1.8 \text{ V}, \quad L = 200 \text{ nm}\end{aligned}$$

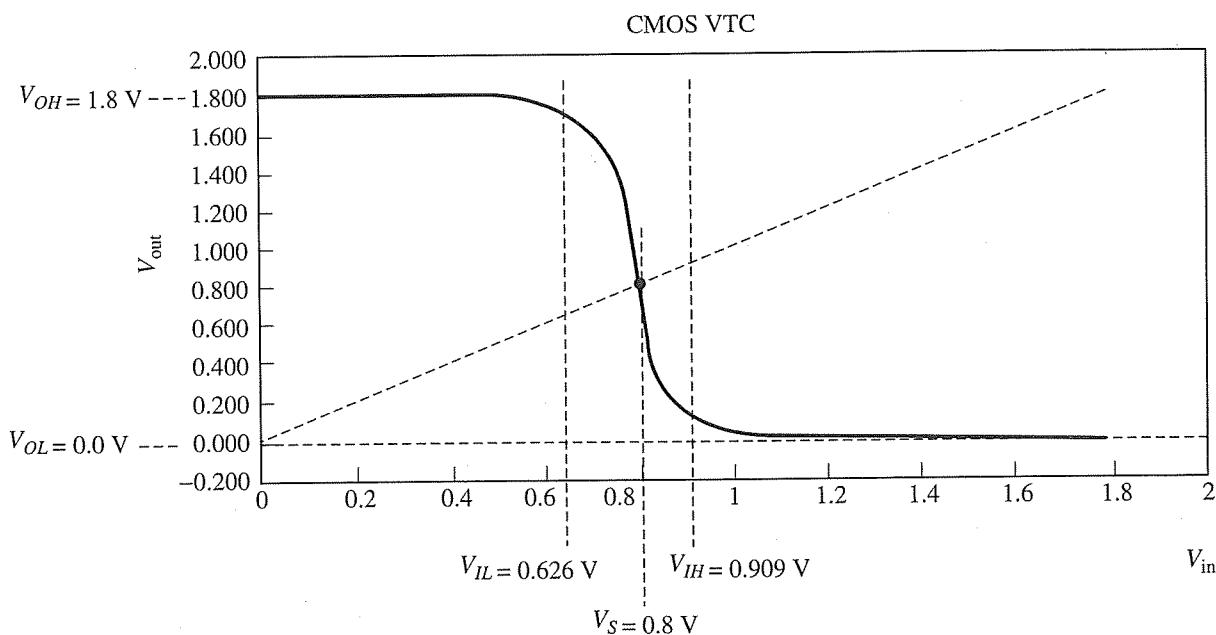
Solution:

```

*CMOS Inverter Example
*Set supply and library
.param Supply=1.8          *for 0.18 technology
.include 'bsim3v3.cmosp18.lib'
.opt scale=0.1u              *Set lambda
.global Vdd Gnd
Vdd      Vdd  Gnd 'Supply'   *Supply is set by .lib call
*Top level simulation netlist
mp        out    in     Vdd  Vdd PMOS1 l=2 w=8 ad=0 pd=0 as=0 ps=0
mn        out    in     Gnd  Gnd NMOS1 l=2 w=4 ad=0 pd=0 as=0 ps=0
*Top level simulation netlist
Vin      in     Gnd 'Supply'
*Simulation
.dc      Vin      0      'Supply'  'Supply/50'
.plot    dc      V(out)
.end

```

From the SPICE simulation using the above input, the plot shown below is produced.



The VTC parameters from SPICE are $V_{OH} = 1.8 \text{ V}$, $V_{OL} = 0 \text{ V}$, $V_{IL} = 0.626 \text{ V}$, $V_{IH} = 0.909 \text{ V}$, $V_S = 0.8 \text{ V}$. Therefore, the noise margins are

$$NM_H = 1.8 - 0.909 = 0.891 \text{ V}$$

$$NM_L = 0.626 - 0 = 0.626 \text{ V}$$

For hand analysis, we already know that $V_{OH} = 1.8 \text{ V}$, and $V_{OL} = 0 \text{ V}$.

The computation of V_S can be carried out as before, with $0.18 \mu\text{m}$ parameters:

$$W_N = 0.4 \mu\text{m} \quad W_P = 0.8 \mu\text{m}$$

$$X = \sqrt{\frac{\frac{W_N}{E_{CN}L_N}}{\frac{W_P}{E_{CP}L_P}}} = \sqrt{\frac{W_NE_{CP}}{W_P E_{CN}}} = \sqrt{\frac{(0.4)(24)}{(0.8)(6)}} = 1.41$$

$$V_S = \frac{1.8 - 0.5 + (0.5)1.41}{1 + 1.41} = 0.83 \text{ V}$$

This value is quite accurate relative to the SPICE results. The values for V_{IL} and V_{IH} require some effort but eventually produce $V_{IL} \approx 0.7 \text{ V}$ and $V_{IH} \approx 1.0 \text{ V}$. Therefore,

$$NM_H = 1.8 - 1.0 = 0.8 \text{ V}$$

$$NM_L = 0.7 - 0 = 0.7 \text{ V}$$

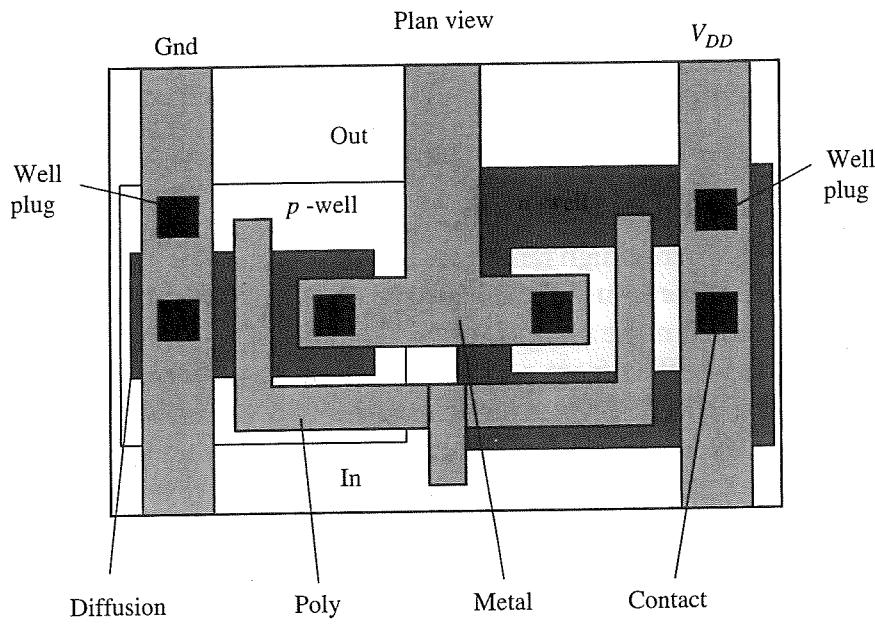
While these are not accurate results, they are within 15% of SPICE and are good enough for hand analysis.

4.6.2 Layout Design of CMOS Inverter

This section illustrates the layout of a simple CMOS inverter. The layers used to implement the design are as follows:

- Metal 1
- Poly
- Contact
- n^+ diffusion
- p^+ diffusion
- n -well
- p -well

These are the main mask layers for the inverter cell. Each of these layers must follow the set of design rules as described in Chapter 3. The n -channel transistors are created with n^+ diffusion in p -wells. The p -channel transistors are created with p^+ diffusion in n -wells. Contacts are used to connect Metal 1 to poly or diffusion.

**Figure 4.13**

Layout of CMOS inverter.

A plan view of a CMOS inverter layout is shown in Figure 4.13. In the plan view, there are two wells, separated by a safe distance, that contain the *n* and *p* devices. The *n*-devices are in the *p*-well and the *p*-devices are in the *n*-well. The inputs for the two devices are formed on the poly layer. When poly crosses diffusion, a transistor is formed. In this case, poly crossings over the n^+ and p^+ regions form the two transistors. The input signal is on poly since it will be driving the gates of the two transistors. All input signals must eventually be connected to the poly layer since all transistor gates are poly gates.

The *n*-channel transistor has its source connected to Gnd while its drain is connected to the output. The *p*-channel device has its source connected to V_{DD} and its drain connected to the output. The output is taken on the Metal 1 layer since metal can be connected directly to the diffusion regions. This output will drive another gate so it must switch to poly at some point using a metal-to-poly contact, not shown in the figure.

There are a total of four contacts associated with the inverter itself: two for the *n*-device and two for the *p*-device. There are also two other contacts that are shown here. These contacts connect the wells to Gnd and V_{DD} , respectively, to keep the pn junctions reverse-biased. They are often referred to as "well plugs."

The layout is important in that the dimensions and areas needed for simulation are extracted from them. Specifically, SPICE simulations require L , W , AS , AD , PS , and PD , as illustrated earlier. The area and perimeter data are primarily for the transient analysis so they have not been used in the SPICE input examples in this chapter. We will be specifying this information in later chapters. Therefore, it is useful to

study this layout to determine how to obtain the desired information for accurate transient simulation.

4.7 Pseudo-NMOS Inverters

There are situations where a high fanin gate is needed in a logic circuit, and for these situations, we would like to use the advantage of NMOS loads, without incurring the penalty of degraded noise margins or two supplies. The saturated enhancement NMOS load inverter suffers from a lower V_{OH} than the other configurations. The linear enhancement NMOS load inverter requires two voltage supplies to produce $V_{OH} = V_{DD}$. While we rarely see these types of inverters, they do have one significant advantage: when designing multi-input (multifanin) gates, we only require one load regardless of the number of inputs. Because of the *push-pull* arrangement in standard CMOS, shown again in Figure 4.14a, we will need roughly twice as many transistors to implement multi-input gates.

One solution is to use a PMOS load configuration that is reminiscent of the NMOS load, as shown in Figure 4.14b. We call this a *pseudo-NMOS* configuration as a reminder of its relation to the NMOS-style loads. The gate of the PMOS device is connected to ground so that the transistor is always on. This device is able to pull the output to V_{DD} when the NMOS device is off without the need for additional supplies. However, it will fight the NMOS transistor when it is on. Therefore, it must be sized along with the NMOS device to deliver the desired V_{OL} . This makes it a ratioed inverter, whereas the standard CMOS inverter is ratioless. When the output is low, power is dissipated in the pseudo-NMOS inverter just as in the other NMOS inverter configurations. This limitation directs us to use this approach sparingly to minimize overall chip power. However, the real value of the pseudo-NMOS configuration will be clear when we examine multi-input gates: it will have a much smaller area than the standard CMOS gate.

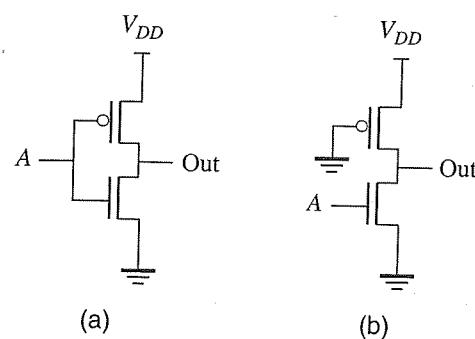


Figure 4.14

(a) CMOS. (b) Pseudo-NMOS inverters.

The voltage transfer characteristics of the pseudo-NMOS inverter can be computed in the usual way and will not be described here. However, V_{OH} and V_{OL} calculations are illustrated. From our discussion above, we already know that

$$V_{OH} = V_{DD} \quad (4.20)$$

The derivation of V_{OL} is as follows:

$$I_{DP}(\text{sat}) = I_{DN}(\text{lin})$$

$$\frac{W_P v_{\text{sat}} C_{\text{ox}} (V_{DD} - |V_{TP}|)^2}{(V_{DD} - |V_{TP}|) + E_{CP} L_P} = \frac{W_N}{L_N} \frac{\mu_n C_{\text{ox}}}{\left(1 + \frac{V_{OL}}{E_{CN} L_N}\right)} \left[(V_{DD} - V_{TN})(V_{OL}) - \frac{(V_{OL})^2}{2} \right]$$

$$(4.21)$$

Since V_{OL} is small, this equation can be simplified to

$$V_{OL} = \frac{I_{DP}(\text{sat})}{k_N(V_{DD} - V_{TN})}$$

The sizes of the two devices can be computed based on the desired V_{OL} .

Design of a Pseudo-NMOS Inverter in 0.18 μm with SPICE Comparison

Example 4.8

Problem:

Design a pseudo-NMOS inverter of Figure 4.14b, to deliver $V_{OH} = V_{DD} = 1.8$ V, and $V_{OL} = 0.065$ V. Plot the VTC using SPICE to confirm your results. Use the following 0.18 μm technology parameters:

$$\mu_n = 270 \text{ cm}^2/\text{V-s}, \quad \mu_p = 70 \text{ cm}^2/\text{V-s}, \quad C_{\text{ox}} = 1.0 \text{ } \mu\text{F/cm}^2,$$

$$V_{TN} = 0.5 \text{ V}, \quad V_{TP} = -0.5 \text{ V}, \quad E_{CN} L = 1.2 \text{ V}, \quad E_{CP} L = 4.8 \text{ V},$$

$$v_{\text{sat}} = 8 \times 10^6 \text{ cm/s}, \quad V_{DD} = 1.8 \text{ V}, \quad L = 200 \text{ nm}$$

Solution:

To design the inverter of Figure 4.14b with the proper V_{OH} , we set $V_{DD} = 1.8$ V. To compute the transistor widths to deliver $V_{OL} = 0.065$ V, use Equation (4.21). The required value of W_N/W_P can be obtained as follows:

$$\frac{W_P v_{\text{sat}} C_{\text{ox}} (V_{DD} - |V_{TP}|)^2}{(V_{DD} - |V_{TP}|) + E_{CP} L_P} = \frac{W_N}{L_N} \frac{\mu_n C_{\text{ox}}}{\left(1 + \frac{V_{OL}}{E_{CN} L_N}\right)} \left[(V_{DD} - V_{TN})(V_{OL}) - \frac{(V_{OL})^2}{2} \right]$$

$$\therefore \frac{W_p(8 \times 10^6)(1.8 - 0.5)^2}{(1.8 - 0.5) + 4.8} = \frac{W_N}{0.2(10^{-4})} \frac{270}{(1 + (0.065/1.2))} \left[(1.8 - 0.5)(0.065) - \frac{(0.065)^2}{2} \right]$$

$$\therefore \frac{W_N}{W_p} \approx 2.0$$

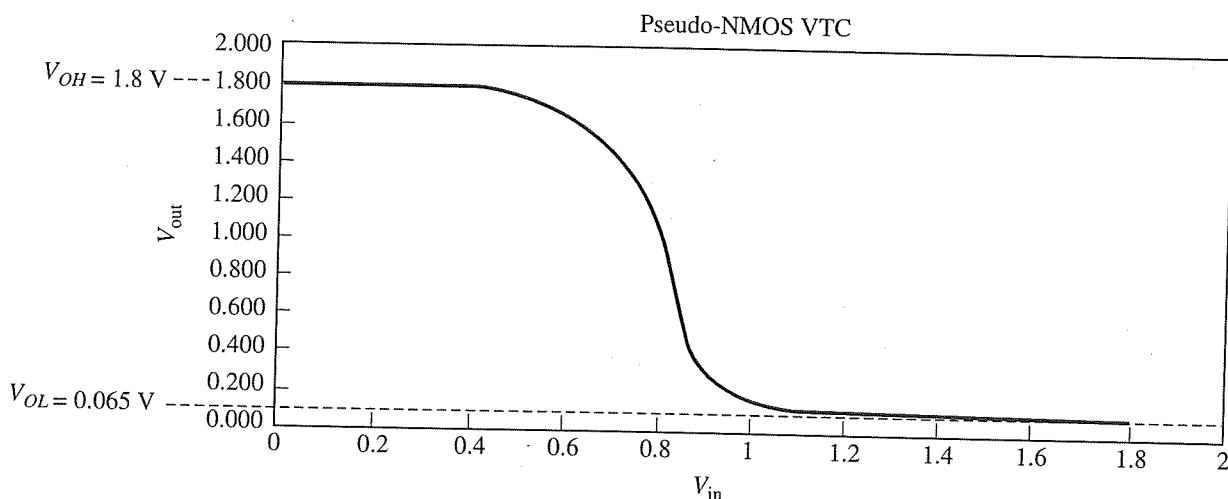
If we set $W_p = 400$ nm, then $W_N = 800$ nm. This should produce the desired $V_{OL} = 0.065$ V.

Now check the results with SPICE.

*Pseudo-NMOS Inverter Example

*Set supply and library

```
.param Supply = 1.8          *for 0.18 technology
.include 'bsim3v3.cmosp18.lib'
.opt scale = 0.1u            *Set lambda
.global Vdd Gnd
Vdd      Vdd Gnd 'Supply'   *Supply is set by .lib call
*Toph level simulation netlist
mp        out    Gnd    Vdd      Vdd PMOS1 l=2 w=4 ad=0 pd=0 as=0 ps=0
mn        out    in     Gnd      Gnd NMOS1 l=2 w=8 ad=0 pd=0 as=0 ps=0
*Toph level simulation netlist
Vin      in     Gnd    'Supply'
*Simulation
.dc      Vin    0      'Supply' 'Supply/50'
.plot    dc      V(out)
.end
```



The SPICE results show that $V_{OH} = 1.8$ V and $V_{OL} = 0.065$ V.

Sizing Inverters

Inverter sizing is the problem of selecting the W/L values for the pull-up and pull-down transistors to satisfy a set of design specifications. The proper selection of device sizes for the basic inverters discussed thus far involves a tradeoff between timing, power, area, and noise margins. Usually one or two of these factors controls the ultimate device sizes.

For the CMOS inverter, we usually focus on the timing to determine the device sizes, since the inverter is ratioless and standby power is small. For the pseudo-NMOS inverter, the sizes depend primarily on the desired V_{OL} (noise margin) and timing.

It is useful to compare the sizing problem for the CMOS inverter and the pseudo-NMOS inverter. In order to do so, we must first develop a simple timing model. More advanced timing models will be the subject of Chapter 6. Here we seek to introduce a simple approach that is extremely valuable for hand calculations.

Figure 4.15a illustrates two important timing specifications: t_{PHL} and t_{PLH} . These are the high-to-low and low-to-high propagation delays, as shown using the waveforms for the input and output voltages of an inverter. A step input is applied to the inverter and the output propagation time is measured at the 50% point. To make the delay calculation simple, we assume that the inverter can be modeled with an effective on-resistance, R_{eff} , driving a load capacitance, C_L .

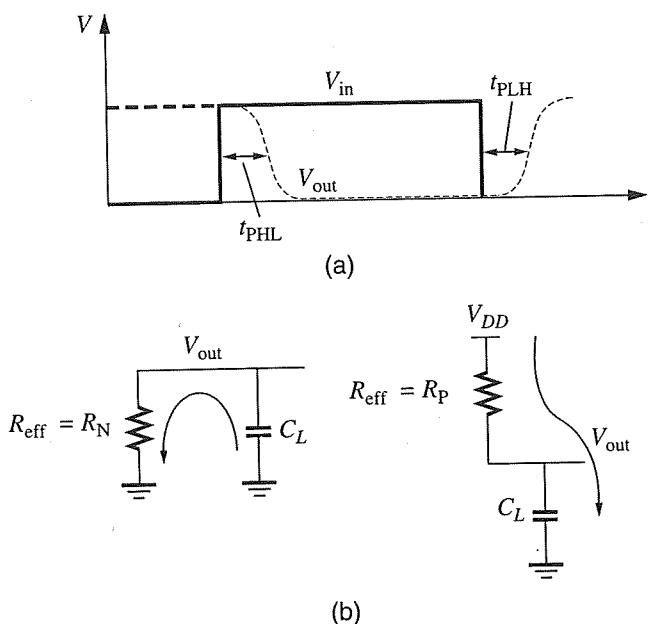


Figure 4.15

Simple timing model for inverter sizing.

The output response to a step input is an exponential waveform. For the falling case, a pull-down resistance, $R_{\text{eff}} = R_N$, is connected to the capacitive load and the output response is given by

$$V_{\text{out}}(t) = V_{DD} e^{-t/R_N C_L} \quad (4.22a)$$

For the rising case, the on-resistance is $R_{\text{eff}} = R_P$ and the output is given by

$$V_{\text{out}}(t) = V_{DD}(1 - e^{-t/R_P C_L}) \quad (4.22b)$$

In both cases, the 50% point occurs at

$$\tau = 0.69 R_{\text{eff}} C_L \quad (4.22c)$$

The two values of R_{eff} are defined as the average on-resistances in the pull-up and pull-down cases, respectively. These values are both inversely proportional to W/L since the resistance decreases as we increase W/L . The NMOS and PMOS devices have different on-resistances that we can compute using SPICE simulations on unit-sized devices when driving a known capacitance. If we measure t_{PHL} and t_{PLH} , we can extract the equivalent on-resistances. From such SPICE measurements, we find that:⁷

$$R_{eqn} = 12.5 \text{ k}\Omega/\square \quad (4.23a)$$

$$R_{eqp} = 30 \text{ k}\Omega/\square \quad (4.23b)$$

where R_{eqn} is associated with the NMOS device and R_{eqp} is associated with the PMOS device. The units of the resistance quantities are ohms per square, similar to the units used for interconnect resistance. However, here the *square* refers to the aspect ratio L/W of the transistor. The actual pull-up and pull-down resistances are computed using the equations:

$$R_N = R_{eqn} \times \frac{L_N}{W_N} \quad (4.24a)$$

$$R_P = R_{eqp} \times \frac{L_P}{W_P} \quad (4.24b)$$

For example, a unit-sized NMOS device has an on-resistance of 12.5 kΩ whereas a unit-sized PMOS device has an on-resistance of 30 kΩ. If we double the width of each device, the NMOS resistance would be 6.25 kΩ and the PMOS resistance would be 15 kΩ. Based on these equations, we can size the NMOS and PMOS devices to deliver the required delay.

⁷ These values are approximately correct for a number of different technologies including 0.35 μm, 0.18 μm, and 0.13 μm.

Comparison of CMOS and Pseudo-NMOS Inverters

Example 4.9

Problem:

Given the CMOS and pseudo-NMOS inverters of Figure 4.14, size the two inverters to achieve the following specifications, assuming that the output loading is 50 fF. Use 0.13 μm technology parameters:

$$\mu_n = 270 \text{ cm}^2/\text{V}\cdot\text{s}, \quad \mu_p = 70 \text{ cm}^2/\text{V}\cdot\text{s}, \quad C_{ox} = 1.6 \text{ } \mu\text{F/cm}^2,$$

$$V_{TN} = 0.4 \text{ V}, \quad V_{TP} = -0.4 \text{ V}$$

$$E_{CN}L = 0.6 \text{ V}, \quad E_{CP}L = 2.4 \text{ V}, \quad v_{sat} = 8 \times 10^6 \text{ cm/s},$$

$$V_{DD} = 1.2 \text{ V}, \quad L = 100 \text{ nm}$$

CMOS Inverter: $t_{PHL} = t_{PLH} < 50 \text{ ps}$; $V_{OH} = 1.2 \text{ V}$ and $V_{OL} = 0 \text{ V}$; low dc power; minimum area

Pseudo-NMOS Inverter: $t_{PHL} < 50 \text{ ps}$; $V_{OH} = 1.2 \text{ V}$ and $V_{OL} = 0.1 \text{ V}$; minimum dc power; minimum area

Solution:

For the CMOS inverter, timing is the highest priority. The noise margin parameters of V_{OH} and V_{OL} are satisfied by choosing $V_{DD} = 1.2 \text{ V}$. The power requirement is satisfied since it is a standard CMOS gate. The area requirement can be satisfied by meeting but not exceeding the timing specification. That is, we do not size the devices any larger than needed. For the timing requirement, we make use of (4.22c):

$$t = 0.7R_{\text{eff}}C_L = 0.7R_{\text{eff}} \times 50 \text{ fF} = 50 \text{ ps} \Rightarrow R_{\text{eff}} = 1.4 \text{ k}\Omega$$

Using (4.24a), we can determine the NMOS (W/L) value:

$$R_N = R_{eqn} \times \frac{L}{W} = 12.5 \text{ k}\Omega \times \frac{L}{W} = 1.4 \text{ k}\Omega \Rightarrow \frac{W_N}{L_N} = 8.75$$

Now, (4.24b) can be used to obtain the value of the PMOS device sizes:

$$R_P = R_{eqp} \times \frac{L}{W} = 30 \text{ k}\Omega \times \frac{L}{W} = 1.4 \text{ k}\Omega \Rightarrow \frac{W_P}{L_P} = 21$$

Since $L = 0.1 \mu\text{m}$, we now have the sizes that satisfy the design specifications:

$$\frac{W_N}{L_N} = \frac{0.875 \mu\text{m}}{0.1 \mu\text{m}} \quad \frac{W_P}{L_P} = \frac{2.1 \mu\text{m}}{0.1 \mu\text{m}}$$

For the pseudo-NMOS gate, we prioritize V_{OL} in the sizing process since it is a ratioed circuit. We can satisfy the V_{OH} requirement by choosing $V_{DD} = 1.2 \text{ V}$. The sizing of the NMOS device should be the same as in the CMOS case since its timing requirement is the same. We can satisfy the power and area requirements by

meeting but not exceeding the V_{OL} and t_{PLH} requirements. To compute the PMOS device size, use (4.21):

$$\frac{W_P v_{sat} C_{ox} (V_{DD} - |V_{TP}|)^2}{(V_{DD} - |V_{TP}|) + E_{CP} L_P} = \frac{W_N}{L_N} \frac{\mu_n C_{ox}}{\left(1 + \frac{V_{OL}}{E_{CN} L_N}\right)} \left[(V_{DD} - V_{TN})(V_{OL}) - \frac{(V_{OL})^2}{2} \right]$$

Using this equation, we apply the known values and solve for W_P :

$$\begin{aligned} \frac{W_P 8(10^6)(1.6)(10^{-6})(1.2 - 0.4)^2}{(1.2 - 0.4) + 2.4} \\ = 8.75 \times \frac{430 \text{ } \mu\text{A/V}^2}{\left(1 + (0.065/0.6)\right)} \left[(1.2 - 0.4)(0.065) - \frac{(0.065)^2}{2} \right] \\ \therefore W_P = 0.665 \text{ } \mu\text{m} \end{aligned}$$

Based on this result, we can specify the device sizes:

$$\frac{W_N}{L_N} = \frac{0.875 \text{ } \mu\text{m}}{0.1 \text{ } \mu\text{m}} \quad \frac{W_P}{L_P} = \frac{0.665 \text{ } \mu\text{m}}{0.1 \text{ } \mu\text{m}}$$

For the pseudo-NMOS case, the two device sizes are about the same. This implies that t_{PLH} will be well above 100 ps (i.e., more than twice as large as t_{PHL}). This is the price that must be paid for using this type of load. There will also be more power dissipated with the output low as in the NMOS load cases. However, the area is far less than CMOS for high-fanin gate circuits, as will be seen later.

4.9 Tristate Inverters

To close out this chapter, we introduce a special type of inverter or buffer used in situations where a bus is to be driven by multiple drivers. In addition to driving the bus high or low, the inverter must be able to assume a high-impedance state at the output so that it is effectively disconnected from the bus. This can be achieved with the inverter configuration shown in Figure 4.16. In this circuit, there are two additional transistors that can be enabled or disabled with the EN signal and its complement. When EN is high, both the p -channel and n -channel devices are *on*. The inverter operates as a regular inverter except that there are two transistors in the signal path.

When the EN signal is low, both the p -channel and n -channel devices are turned off and the output enters a high-impedance condition (sometimes called a *high-Z* state). The inverter no longer controls the output voltage. Rather, some other bus driver can set the value of the bus. Typically, only one set of drivers should be driving a bus at a given point in time. Otherwise, conflicts would arise if multiple drivers try to gain control of the bus at the same time. If no other driver is controlling the bus, the previous value is stored at the output. Therefore, the EN signal is critical in determining when a given driver exercises control or relinquishes control of the bus. Other forms of tristate drivers exist and will be described later in the book.

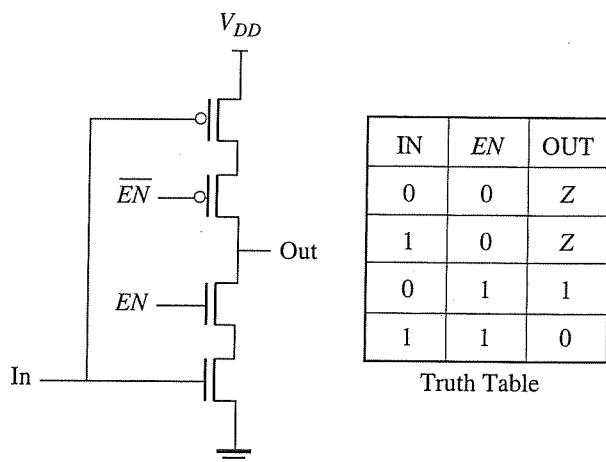


figure 4.16
Enable-state buffer.

.10 Summary

useful technology parameters for $0.18 \mu\text{m}$ CMOS:

$$\begin{aligned} \mu_n &= 270 \text{ cm}^2/\text{V-s}, \quad \mu_p = 70 \text{ cm}^2/\text{V-s}, \quad t_{ox} = 35 \text{ \AA}, \\ C_{ox} &= 1.0 \text{ } \mu\text{F/cm}^2, \quad V_{TN} = 0.5 \text{ V}, \quad V_{TP} = -0.5 \text{ V} \\ E_{CN}L &= 1.2 \text{ V}, \quad E_{CP}L = 4.8 \text{ V}, \quad v_{sat} = 8 \times 10^6 \text{ cm/s}, \\ V_{DD} &= 1.8 \text{ V}, \quad 2|\phi_F| = 0.85 \text{ V}, \quad \gamma = 0.3V^{1/2}, L = 2\lambda = 200 \text{ nm} \end{aligned}$$

Useful technology parameters for $0.13 \mu\text{m}$ CMOS:

$$\begin{aligned} \mu_n &= 270 \text{ cm}^2/\text{V-s}, \quad \mu_p = 70 \text{ cm}^2/\text{V-s}, \quad t_{ox} = 22 \text{ \AA}, \\ C_{ox} &= 1.6 \text{ } \mu\text{F/cm}^2, \quad V_{TN} = 0.4 \text{ V}, \quad V_{TP} = -0.4 \text{ V} \\ E_{CN}L &= 0.6 \text{ V}, \quad E_{CP}L = 2.4 \text{ V}, \quad v_{sat} = 8 \times 10^6 \text{ cm/s}, \\ V_{DD} &= 1.2 \text{ V}, \quad 2|\phi_F| = 0.88 \text{ V}, \quad \gamma = 0.2V^{1/2}, L = 2\lambda = 100 \text{ nm} \end{aligned}$$

CMOS noise margins:

- Single-Source Noise Margins

$$\begin{aligned} SSNM_H &= V_{OH} - V_S \quad V_S \cong \frac{V_{DD} - |V_{TP}| + \chi V_{TN}}{1 + \chi} \quad \chi = \sqrt{\frac{W_N/E_{CN}L_N}{W_P/E_{CP}L_P}} \\ SSNM_L &= V_S - V_{OL} \end{aligned}$$

- Multiple-Source Noise Margins

$$NM_H = V_{OH} - V_{IH}$$

$$NM_L = V_{IL} - V_{OL}$$

On resistance calculations:

- NMOS

$$R_{eqn} = 12.5 \text{ k}\Omega/\square \quad R_N = R_{eqn} \times \frac{L_N}{W_N}$$

- PMOS

$$R_{eqp} = 30 \text{ k}\Omega/\square \quad R_P = R_{eqp} \times \frac{L_P}{W_P}$$

REFERENCES

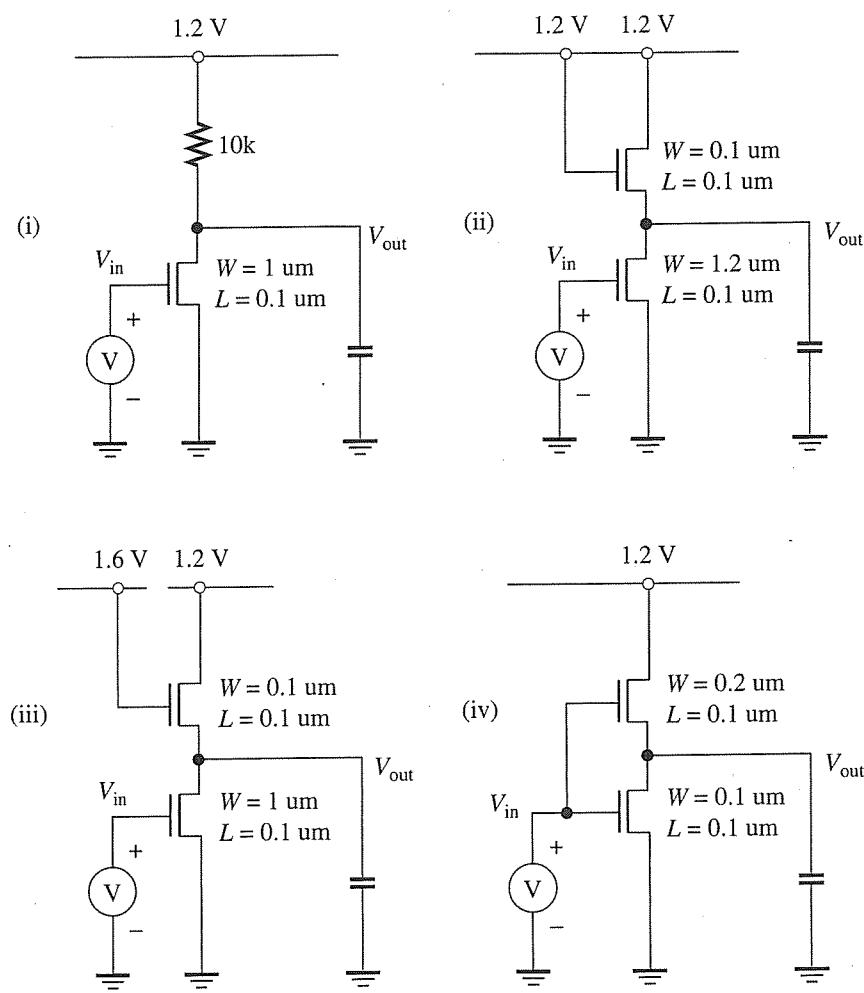
A long list of references exists for this material. Here are a few recent ones.

1. S. M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits, Analysis and Design*, 3rd ed., McGraw-Hill, New York, NY, 2003.
2. J. Rabaey, *Digital Integrated Circuits: A Designer Perspective*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ, 2003.
3. H. Veendrick, *Deep-Submicron CMOS ICs*, 2nd ed., Kluwer Academic Publishers, Boston, MA, 2000.
4. J. P. Uyemura, *CMOS Logic Circuit Design*, Kluwer Academic Publishers, Boston, MA, 1999.

PROBLEMS

All problems assume 0.13 μm technology parameters, although 0.18 μm technology may be used with the appropriate modifications.

- P4.1.** The circuits of Figure P4.1 show different implementations of an inverter whose output is connected to a capacitor.
- (a) Which one of the circuits consumes static power when the input is high?
 - (b) Which one of the above circuits consumes static power when the input is low?
 - (c) V_{OH} of which circuit(s) is 1.2 V?
 - (d) V_{OL} of which circuit(s) is 0 V?
 - (e) The proper functionality of which circuit(s) depends on the size of the devices?

**Figure P4.1**

P4.2. Calculate V_{OH} and V_{OL} of the circuits of Figure P4.2.

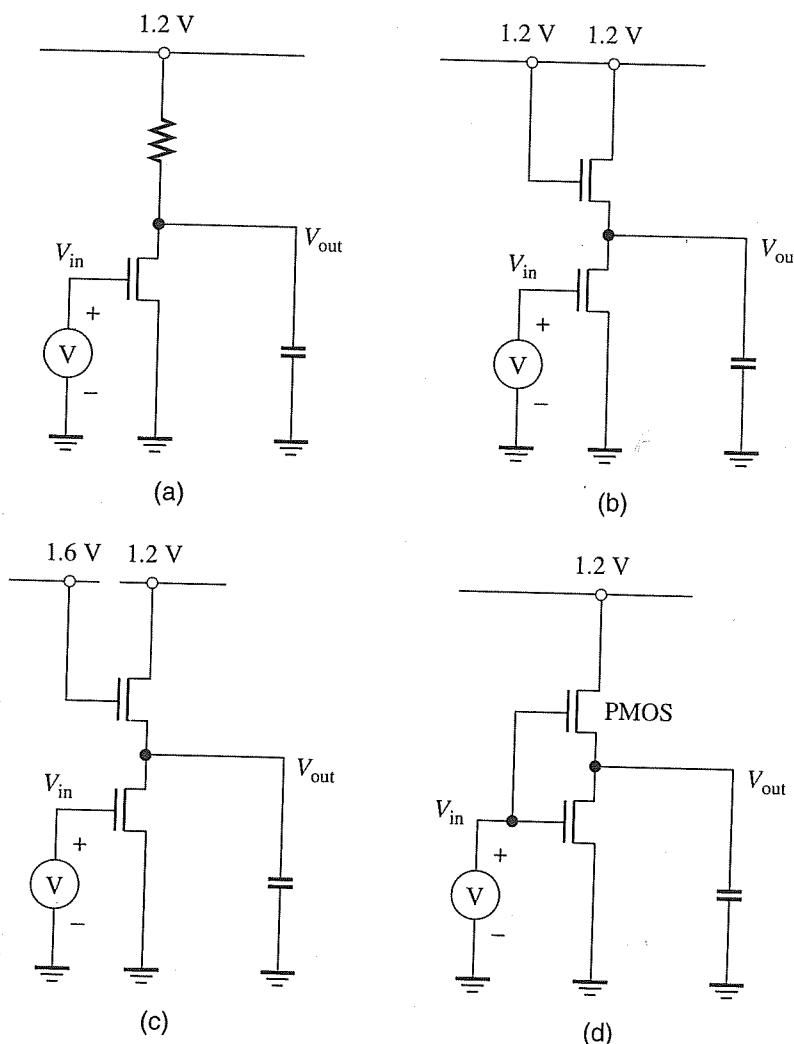


Figure P4.2

- P4.3.** Compute V_{IL} and V_{IH} of a CMOS inverter with device sizes $W_P = 16\lambda$ and $W_N = 4\lambda$, and one with $W_P = 8\lambda$ and $W_N = 4\lambda$. Compare the resulting multistage noise margins, NM_H and NM_L .
- P4.4.** A saturated enhancement load inverter has a pull-up and pull-down size of $W = 4\lambda$. Calculate V_{OH} , V_{OL} , V_{IH} and V_{IL} . Compare your results with SPICE. Which values are not as accurate in relation to the SPICE results and why?
- P4.5.** Design a CMOS inverter with a switching voltage V_S of $\frac{2}{3}V_{DD}$. What is the resulting ratio of W_P/W_N ?

P4.6. The circuit of Figure P4.6a has eight inverters in series:

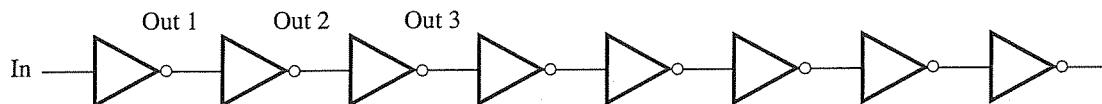


Figure P4.6a

Each inverter is designed as in Figure P4.6b:

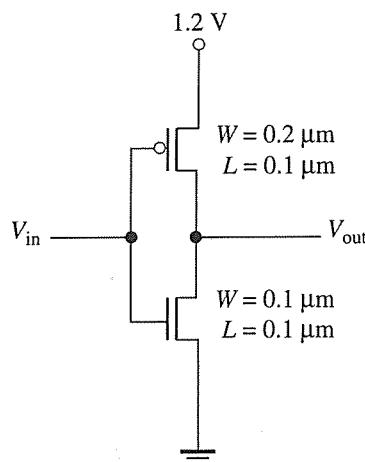


Figure P4.6b

If the input voltage is zero, the output of the first inverter will be 1.2 V. However, if the input is 0.45 V because of noise, the output deviates from 1.2 V. Using SPICE simulation, determine after how many stages of inverters, the deviation of the output signal is less than 2%. At what value of input does the inverter chain flip its output values?

P4.7. A new logic family is proposed whereby the NMOS device acts as a load device and the PMOS device acts as the inverting device, as shown in Figure P4.7. What are the advantages and disadvantages of this *pseudo-PMOS* architecture over the pseudo-NMOS architecture?

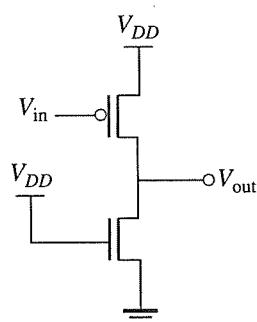


Figure P4.7

A pseudo-PMOS device.

- P4.8. (a) What is the intended function of the circuit shown in Figure P4.8? What is the output swing?

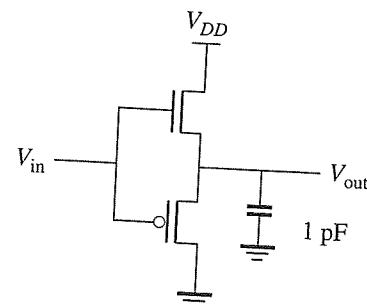


Figure P4.8

- (b) Draw the dc voltage transfer characteristic of the above gate. Label V_{OL} and V_{OH} , and any other interesting values in the VTC. Since the gate has hysteresis, be sure to handle both the rising and falling cases.
- (c) What is the gain of the circuit? Is this a valid gate (i.e., does it have the needed noise rejection properties)?
- (d) Use SPICE to validate your solution by plotting the VTC.

- P4.9. In the circuits of Figure P4.9, design the widths of the pull-down transistors so that $V_{OL} = 0.1 \text{ V}$. (All transistors are minimum size, $L = 0.1 \mu\text{m}$.) Explain the results.

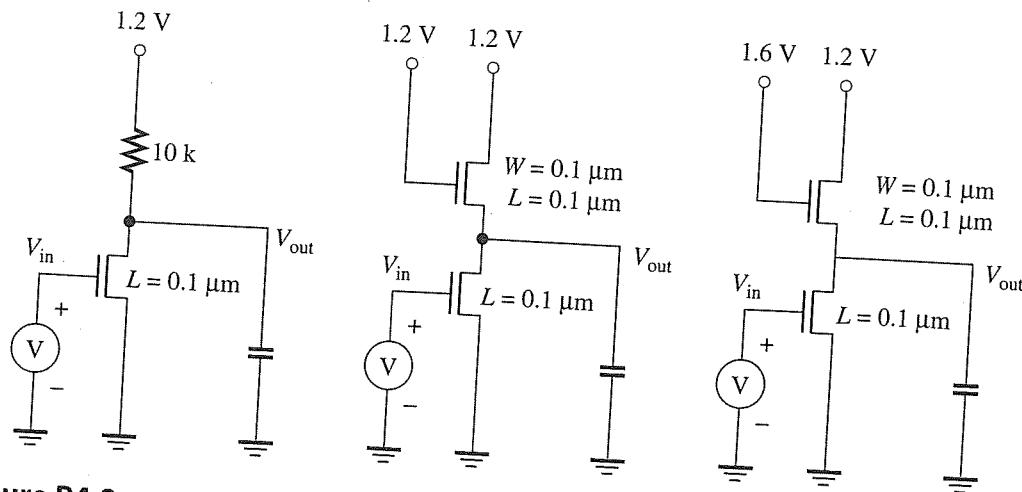
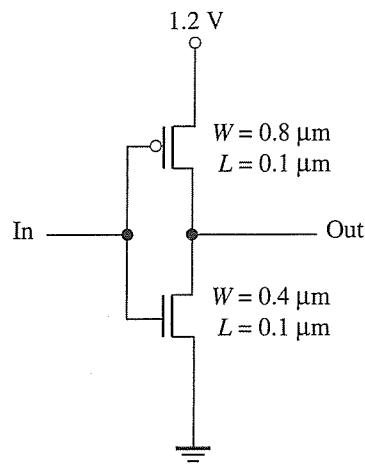


Figure P4.9

- P4.10. Calculate V_{OH} , V_{OL} , V_{IL} , V_{IH} , and V_s , for Figure P4.10. How do you expect each value to change if the width of the PMOS transistor doubles? Calculate V_s again with a double-sized PMOS device to verify your answer. Sketch the VTC for the two cases.

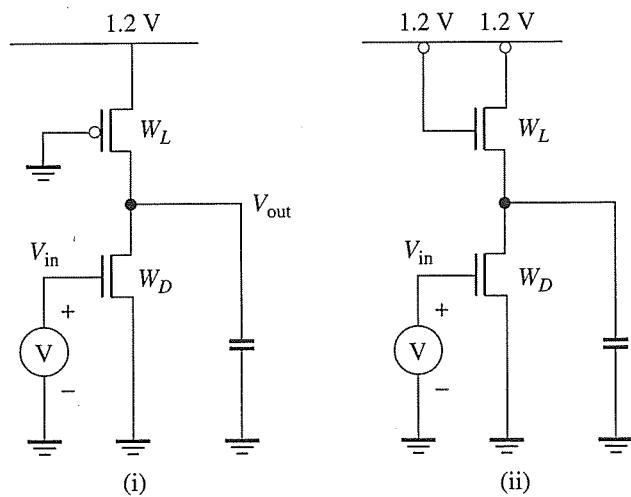
**Figure P4.10**

P4.11. Compute the peak dc current of the inverter of Figure P4.10. At what voltage does this occur?

P4.12. For the circuits of Figure P4.12, compute the ratio of W_D/W_L for the following cases, given that $V_{in} = 1.2 \text{ V}$:

- (a) $V_{out} = V_{OL} = 0.1 \text{ V}$
- (b) $V_{out} = V_{TN} = 0.4 \text{ V}$
- (c) $V_{out} = V_S \cong 0.6 \text{ V}$

What happens to the ratio as the desired output voltage increases?

**Figure P4.12**

P4.13. Using SPICE and Figure P4.13, determine the circuit noise margins (NM_H , NM_L). How does the VTC compare with expected results for this type of inverter?

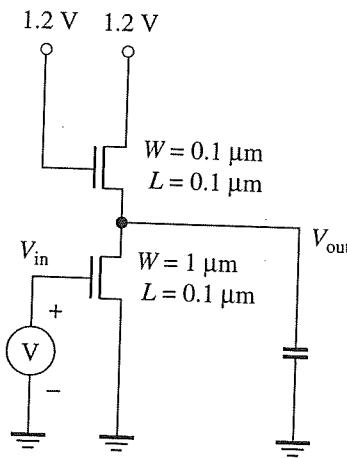


Figure P4.13

P4.14. Design a static CMOS inverter with

- (a) $V_S = 0.6 V_{DD}$
- (b) $V_S = 0.4 V_{DD}$
- (c) What happens to the propagation delays in these two cases? Compute t_{PHL} and t_{PLH} . Explain the tradeoffs.

P4.15. A new type of logic gate is designed with the voltage transfer characteristics shown in Figure P4.15.

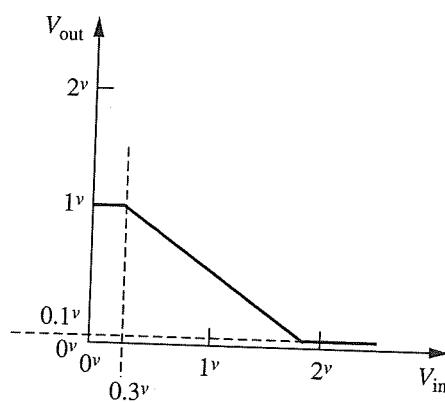


Figure P4.15

- (a) Does this inverter have the regenerative property. Why or why not?
- (b) Consider a chain of such gates with 100 stages. For an input of 0^v , what is the output of the last inverter?
- (c) How much noise can the inverter tolerate? Is it possible to define the noise margin for this gate?
- P4.16.** The two CMOS circuits of Figure P4.16 are intended to be tristate buffers. Do both of these gates work as tristate buffers? Is one better than the other? Explain.

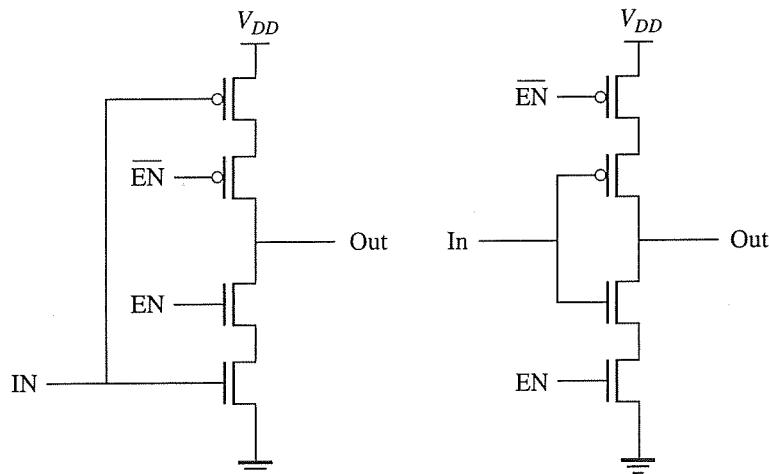


Figure P4.16