1-1-2007

Dynamic power dissipation of basic logic gates

Katayoun Pourbahri
Ryerson University

DYNAMIC POWER DISSIPATION OF BASIC LOGIC GATES

by:

Katayoun Pourbahri
B.Eng, Ryerson University, 2005

A thesis
presented to Ryerson University
in partial fulfillment of the
requirement for the degree of
Master of Applied Science
in the Program of
Electrical and Computer Engineering.

Toronto, Ontario, Canada, 2007

Instructions on Borrowers

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Katayoun Pourbahri

Dynamic Power Dissipation of Basic Logic Gates
MASc, Electrical and Computer Engineering
Ryerson University, 2007.

In this thesis, a model is proposed to estimate the dynamic power dissipation of CMOS logic gate that is loaded with identical logic gates. The proposed model is based on parasitic capacitance, input capacitance and input-to-output coupling capacitance of the gate. Also, model takes into account transistor width and has a second order relation with fanout. Using $0.13\mu m$ CMOS technology and netlist (Cadence), basic logic gates are designed and loaded by identical basic logic gates. Basic logic gates dynamic power are simulated and compared with the calculated values of proposed model over a range of transistor sizes and capacitive load condition. The proposed model shows a good agreement with the simulated value of dynamic power dissipation of basic logic gates that are loaded by identical basic logic gates.

Index Terms-- Dynamic Power Dissipation, Basic Logic Gates.

Acknowledgments

I would like to express my deep appreciation to my supervisor, Dr. Raahemifar, for h
guidance and for providing excellent research opportunities during my studies. Thank you

Contents

1 Introduction 1
- 1.1 Motivation of the Thesis . . . 2
- 1.2 Objective of the Thesis . . . 3
- 1.3 Main Contribution of the Thesis . . . 4
- 1.4 Organization of the Thesis Chapters . . . 4

2 Survey and Technical Background 5
- 2.1 MOS Transistor Theory . . . 5
- 2.2 I-V Characteristics of MOS Transistors . . . 5
  - 2.2.1 Non-ideal I-V Effects . . . 10
  - 2.2.2 Velocity Saturation and Mobility Degradation . . . 11
  - 2.2.3 Sub-threshold Conduction . . . 13
- 2.3 C-V Characteristics of MOS Transistors . . . 14
- 2.4 Propagation Delay . . . 17
- 2.5 Turn-on and Turn-off Delay . . . 19

3 Proposed Model 20
- 3.1 Switching Power Dissipation . . . 20
  - 3.1.1 Operating Frequency and Voltage . . . 22
- 3.2 Overshoot Power Dissipation . . . 23
- 3.3 Short Circuit Power Dissipation . . . 28
- 3.4 Dynamic Power Dissipation Proposed Model . . . 32

4 Model Validation 35
- 4.1 Dynamic Power Dissipation Due to Actual Gate . . . 36
- 4.2 Dynamic Power Dissipation Due to Equivalent Lumped Capacitance . . . 45
- 4.3 Dynamic Power Dissipation and Transistor Width . . . 53

5 Conclusions and Future Works 59

Bibliography 61

List of Figures

1.1 Increase of power dissipation due to device scaling. 2
1.2 Dynamic power dissipation currents. 3

2.1 Average gate to channel voltage 6
2.2 NMOS transistor cutoff region of operation. 8
2.3 NMOS transistor linear region of operation. 8
2.4 NMOS transistor saturation region of operation. 9
2.5 I-V characteristics of ideal NMOS and PMOS transistors. 9
2.6 Carrier velocity vs. electric field 11
2.7 I-V characteristics for NMOS transistor with velocity saturation 13
2.8 MOS transistor capacitances model 15
2.9 Overlap capacitance 16
2.10 Diffusion region geometry 17
2.11 RC ladder for Elmore delay 18

3.1 Generic representation of a CMOS logic gate. 20
3.2 Input/output voltage and PMOS/NMOS current of inverter. 23
3.3 Three operating regions and inverter models in each region. 24
3.4 CMOS inverter input/output and transistors current for a falling input . . . 27
3.5 Current behavior of an inverter without load. 28
3.6 Inverter input/output and short circuit current under different capacitive load condition. 30
3.7 Short circuit current 31

4.1 Transistor level circuits of basic logic gates 35
4.2 Inverter input/output voltage and output transistor power dissipation due to actual gate ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.26\mu m$). 37
4.3 Nand2 input/output voltage and output transistor power dissipation due to actual gate ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.52\mu m$). 38
4.4 Nor2 input/output voltage and output transistor power dissipation due to actual gate ($W_{PMOS} = 1.04\mu m$ and $W_{NMOS} = 0.26\mu m$). 39
4.5 Nand3 input/output voltage and output transistor power dissipation due to actual gate ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.78\mu m$). 40

4.6 Nor3 input/output voltage and output/transistor power dissipation due to actual gate ($W_{PMOS} = 1.56\mu m$ and $W_{NMOS} = 0.26\mu m$). 41
4.7 Average and dynamic power dissipation of inverter with no-load. 42
4.8 Comparison of proposed model (P_{model}) and basic logic gates dynamic power dissipation due to actual gate (P_{act}) simulation results under different loading conditions. 44
4.9 Inverter input/output voltage and transistor/output power dissipation due to lumped capacitance. 46
4.10 Nand2 input/output voltage and output/transistor power dissipation due to lumped capacitance ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.52\mu m$). 47
4.11 Nor2 input/output voltage and output/transistor power dissipation due to lumped capacitance($W_{PMOS} = 1.04\mu m$ and $W_{NMOS} = 0.26\mu m$). 48
4.12 Nand3 input/output voltage and output/transistor power dissipation due to lumped capacitance ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.78\mu m$). 49
4.13 Nor3 input/output voltage and output/transistor power dissipation due to lumped capacitance ($W_{PMOS} = 1.56\mu m$ and $W_{NMOS} = 0.26\mu m$). 50
4.14 Comparison of basic logic gates power dissipation due to actual gate (P_{act}) and equivalent lump capacitance (P_{cap}) under different load conditions 52
4.15 Comparison of proposed model (P_{model}) and inverter P_{act} with respect to transistor width under different load conditions. 54
4.16 Comparison of proposed model (P_{model}) and 2-input nand P_{act} with respect to transistor width under different load conditions. 55
4.17 Comparison of proposed model (P_{model}) and 2-input nor P_{act} with respect to transistor width under different load conditions. 56
4.18 Comparison of proposed model (P_{model}) and 3-input nand P_{act} with respect to transistor width under different load conditions. 57
4.19 Comparison of proposed model (P_{model}) and 3-input nor P_{act} with respect to transistor width under different load conditions. 58

List of Tables

2.1 The approximate values of gate capacitance components 16

3.1 The activity factor (α) of basic logic gates. 21

4.1 Transistor width of minimum size basic logic gates (L=0.13μm) 35
4.2 Basic logic gates C_M, $C_{gate,in}$ and $C_{gate,par}$. 42
4.3 Comparison of proposed model (P_{model}) and basic logic gates dynamic power dissipation simulation results due to actual gate (P_{act}) under different loading conditions (power values are in μW and $E\% = \frac{P_{act}-P_{model}}{P_{act}} \times 100$). 43
4.4 The value of equivalent lumped capacitance of basic logic gates. 45
4.5 Comparison of basic logic gates power dissipation due to actual gate (P_{act}) and equivalent lump capacitance (P_{cap}) under different load conditions (power values are in μW and $E\% = \frac{P_{act}-P_{cap}}{P_{act}} \times 100$). 51
4.6 Comparison of proposed model (P_{model}) and inverter P_{act} with respect to transistor width under different load conditions (power values are in μW). 54
4.7 Comparison of proposed model (P_{model}) and 2-input nand P_{act} with respect to transistor width under different load conditions (power values are in μW). 55
4.8 Comparison of proposed model (P_{model}) and 2-input nor P_{act} with respect to transistor width under different load conditions (power values are in μW). . 56
4.9 Comparison of proposed model (P_{model}) and 3-input nand P_{act} with respect to transistor width under different load conditions (power values are in μW). 57
4.10 Comparison of proposed model (P_{model}) and 3-input nor P_{act} with respect to transistor width under different load conditions (power values are in μW). . 58

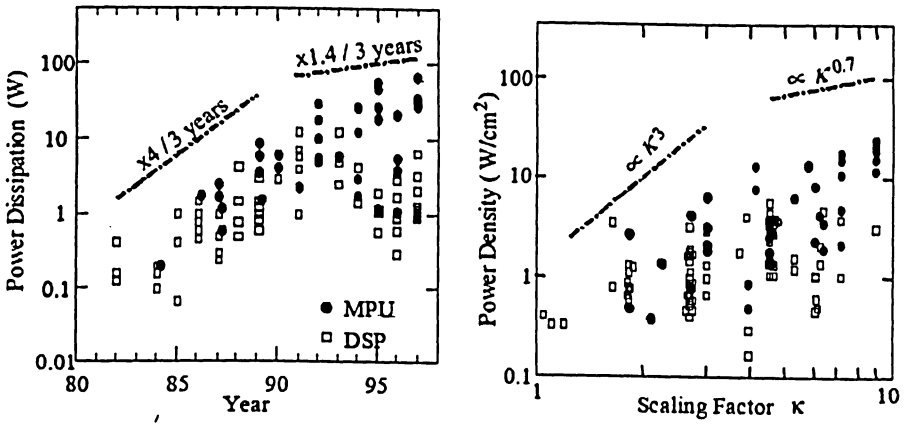
Chapter 1

Introduction

In 1965, Gordon Moore, a co-founder of Intel, prepared an article for the 35th anniversary of Electronics Magazine on the future of the semiconductor industry. By tracking the history of the growing integrated chip over the period 1959-65, he made a prediction about its future. Moore said that the number of components on a single silicon chip would continue to double every year. In the 1975 IEEE International Electron Devices meeting, Moore revised his prediction to the number of transistors on a chip doubling every two years. This prediction became known as Moore's Law.

In fact in that time, technology development was associated with the increasing of integration and speed of circuits with taking cost and reliability into account. While power dissipation, for the most part, has been ignored since power dissipation in those generations has stayed within acceptable limits.

As shown in the Figure 1.1 [1], CMOS power dissipation has been increasing due to the increase in power density. The power dissipation increased 4 times every 3 years until the early 1990's, due to a constant voltage scaling. Then, a constant field scaling has been applied to reduce power dissipation, where the power density is increased in proportion to the 0.7th power of scaling factor, resulting in power increase by 1.4 times every 3 years. It is considered that the power dissipation of CMOS chips will steadily increase as a natural result of device scaling [1]. However, as we move toward sub-micron technologies and multi-GHz frequencies, the power dissipation exceeds acceptable limits jeopardizing performance



(a) Power dissipation vs. year (b) Power density vs. scaling factor

Figure 1.1: Increase of power dissipation due to device scaling.

and reliability. For example, Toshiba has developed a chip that can read and write 200 megabytes per second. Also, CMOS technology is improved from $0.35\mu m$ to $65nm$ node.[1]

1.1 Motivation of the Thesis

Increasing the power dissipation results in raising the chip temperature, which in turn can make the chip runs slower. Worst-case scenario: when the amount of power dissipation exceeds the specification, the chip can be permanently damaged. In order to keep chip temperature down a cooling device has to be added to system which increases the cost as well. Power costs money and portable power costs even more and power dissipation shortens the battery life of portable devices which have to be charged or replaced and all of these cost money. In order to reduce the power dissipation design engineers have to know where and why power is dissipated and need to have an accurate estimation model that can work as a hand tool in addition to simulation tools such as Cadence for power management. Not only optimization of chip design but also choosing the correct IC package and optimization the box or board containing the chip will be possible through accurate power dissipation model.

[1]At the time of writing this thesis

1.2 Objective of the Thesis

The power dissipation is composed of dynamic and static components. The primary distinguishing factor between these two components is that dynamic power dissipation is frequency dependent, but static power dissipation is not. Dynamic power dissipation is the main focus of this thesis and consists of switching power dissipation, short circuit power dissipation and overshoot power dissipation. Dynamic power dissipation (P_{dyn}) can be expressed as the sum of these three components as [2]:

$$P_{dyn} = P_{sw} + P_{sc} + P_{ov}, \tag{1.1}$$

where switching power dissipation (P_{sw}) is due to charging and discharging of the capacitive load that is connected to the output node and short circuit power dissipation (P_{sc}) is caused by the current that flows from the power supply to the ground through the NMOS and PMOS transistors that conduct simultaneously during input signal transitions. The overshoot power dissipation (P_{ov}) is due to the current that flows from the input to the output through the input-to-output coupling capacitance [2]. Figure 1.2 illustrates the path of the currents that are responsible for the dynamic power dissipation of an Inverter.

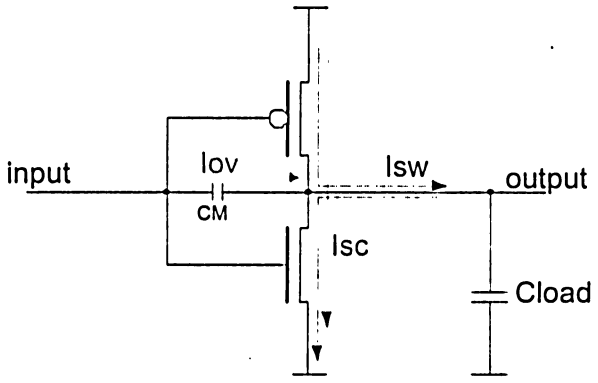


Figure 1.2: Dynamic power dissipation currents.

The simulation results of dynamic power dissipation of basic logic gates due to actual gate shows a second-order relation with the fanout. However, the existing dynamic power dissipation model has a first-order relation with the fanout and shows a big discrepancy with

the simulation results at high fanout. The objective of this thesis is to modify the existing model of dynamic power dissipation and get better accuracy at high fanout. This model gives design engineers a better hand tool to estimate dynamic power dissipation without using Cadence and decrease dynamic power dissipation. In order to use the proposed model to optimize circuit delay-power through transistor sizing method in the future works, it is required that the model represent the dynamic power dissipation relation to transistor size.

1.3 Main Contribution of the Thesis

By taking input-to-output coupling capacitance into account through modifying the overshoot power dissipation, the second-order proposed model was developed to estimate the dynamic power dissipation of basic logic gates. Also, the short circuit power dissipation model as a part of dynamic power dissipation model was modified to improve the proposed model accuracy. The proposed model achieved a good accuracy with the simulation results at low and high fanouts. Since the proposed model of dynamic power dissipation takes into account transistor width, it can be used to size transistor in order to optimize the dynamic power dissipation.

1.4 Organization of the Thesis Chapters

The remainder of this thesis is organized as follows: Chapter 2 is about technical background; Chapter 3 reviews the previous works on dynamic power dissipation and its components, also presents the theory behind the proposed model; Chapter 4 includes design of basic logic gates using 0.13μm CMOS technology and netlist and simulation to validate the proposed model of basic CMOS logic gate; Chapter 5 summarizes the conclusions and proposes areas for future works.

Chapter 2

Survey and Technical Background

2.1 MOS Transistor Theory

The MOS transistor is a majority carrier device, in which the current in a conducting channel between the source and drain is controlled by the voltage applied to the gate. In an NMOS transistor, the majority carriers are electrons and in a PMOS transistor the majority carriers are holes. In a simple MOS structure, the top layer of the structure is a good conductor called the gate. Early transistors used metal gates, but modern transistors generally use polysilicon. The middle layer is a very thin insulating film of sio_2 called the gate oxide. The bottom layer is the doped silicon body. The gate oxide is a good insulator and when body is grounded and a voltage is applied to the gate almost zero current flows from the gate to the body. The behavior of MOS transistors could be explained as I-V (current-voltage) and C-V (capacitance-voltage) characteristics of transistor [3].

2.2 I-V Characteristics of MOS Transistors

According to ideal Shockley current model and considering an NMOS transistor, when the gate-to-source voltage (V_{gs}) is less than threshold voltage (V_t), there is no channel and almost no current flows from drain to source. When the V_{gs} exceeds the V_t, the gate attracts carriers (electrons) to form a channel (the vertical electric field). The electrons drift from source to drain at a rate proportional to the electric field between these regions (the lateral electric

field). Thus the currents can be computed by knowing the amount of charge in the channel and the rate at which it moves. Since the charge on each plate of a capacitor is $Q = CV$, the charge in the channel can be expressed as:

$$Q = C_g(V_{gc} - V_t), \tag{2.1}$$

where C_g is the capacitance of the gate to the channel and $(V_{gc} - V_t)$ is the amount of voltage attracting charge to the channel beyond the minimum required to invert from p to n. The gate voltage is referenced to the channel, which is not grounded. If the source is at V_s and the drain is at V_d, the channel voltage V_c is equal $V_s + V_{ds}/2$. Therefore, the mean difference between the gate and channel potentials can be express as $V_{gc} = V_{gs} - V_{ds}/2$, as shown in Figure 2.1 [3].

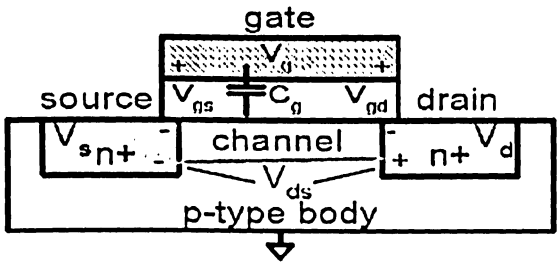


Figure 2.1: Average gate to channel voltage

The gate can be modeled as a parallel plate capacitor with capacitance proportional to area over thickness and expressed as $C_g = C_{ox}WL$, where C_{ox} is the capacitance per unit area of the gate oxide, L and W are the length and width of the gate, respectively.

Each carrier in the channel is accelerated to an average velocity proportional to the lateral electric field (the field between source and drain). The constant of proportionality μ is called the mobility and average velocity of the carrier in the channel can be expressed as:

$$v = \mu E, \tag{2.2}$$

where the electric field E (the lateral electric field) is the voltage difference between drain and source (V_{ds}) divided by the channel length as:

$$E = \frac{V_{ds}}{L}, \tag{2.3}$$

where the time required for carriers to cross the channel is the channel length divided by the carrier velocity (L/v). Therefore, the current between source and drain is the total amount of charge in the channel divided by the time required to cross the channel:

$$I_{ds} = \frac{Q}{L/v}, \tag{2.4}$$

replacing the equivalent values of V_{gc} and C_g in Equation (2.1) total energy is given by:

$$Q = C_{ox}WL(V_{gs} - V_t - \frac{V_{ds}}{2}),$$

and by replacing the equivalent values of Q and v in Equation 2.4 the I_{ds} can be expressed as:

$$I_{ds} = \mu C_{ox}\frac{W}{L}(V_{gs} - V_t - \frac{V_{ds}}{2})V_{ds}, \tag{2.5}$$

where Equation (2.5) describes the linear region of operation for $V_{gs} > V_t$ and relatively small V_{ds}. It is called linear because $\frac{V_{ds}}{2} << V_{gs} - V_t$ and I_{ds} increases almost linearly with V_{ds} just like an ideal resistor. However, when $V_{ds} > V_{gs} - V_t$, the channel is no longer inverted in the nearby of drain and it is pinched off. Beyond this point, it is called drain saturation voltage and increasing the drain voltage has no further effect on current. By substituting $V_{ds} = (V_{gs} - V_t)$ into 2.5, an expression can be developed for the saturation current that is independent of V_{ds}. This expression is valid for $V_{gs} > V_t$ and $V_{ds} > V_{gs} - V_t$ as:

$$I_{ds} = \mu C_{ox}\frac{W}{2L}(V_{gs} - V_t)^2, \tag{2.6}$$

also, I_{dsat} can be defined as the current of a transistor that is fully ON ($V_{gs} = V_{dd}$):

$$I_{dsat} = \mu C_{ox}\frac{W}{2L}(V_{dd} - V_t)^2 \tag{2.7}$$

In summery, according to ideal Shockley current model, an NMOS transistor has three regions of operation [3]:

- Cutoff region

  When $V_{gs} < V_t$, the source and drain have free electrons and the body has free holes

but no free electrons. The junctions between the body and the source or drain are reverse-biased, so almost zero currents flows. This mode of operation is called cutoff as shown in Figure 2.2 [3].

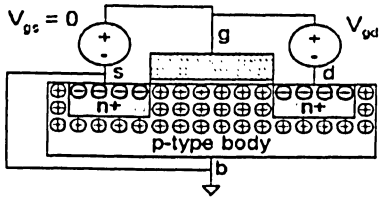


Figure 2.2: NMOS transistor cutoff region of operation.

- Linear region

  When $V_{ds} < (V_{gs} - V_t)$, an inversion region of electrons (majority carriers) called the channel connects the source and drain, creating a conductive path. The number of carriers and the conductivity increases with the gate voltage. The potential difference between drain and source is $V_{ds} = V_{gs} - V_{gd}$ and if $V_{ds} = 0$, there is no electric field tending to push current from drain to source. As shown in Figure 2.3 [3] when a small positive potential V_{ds} is applied to the drain, current I_{ds} flows through the channel from drain to source. The current increases with both the drain voltage and gate voltage and this mode of operation is called linear.

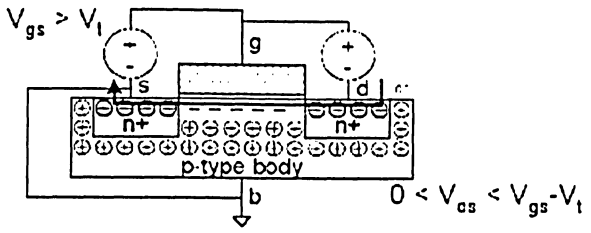


Figure 2.3: NMOS transistor linear region of operation.

- Saturation region

  When $V_{ds} > (V_{gs} - V_t)$, the channel is no longer inverted near the drain and becomes pinched off as shown in Figure 2.4 [3]. However, conduction is still brought about by

the drift of electrons under the influence of the positive drain voltage. As electrons reach the end of the channel, they are injected into the depletion region near the drain and accelerated toward the drain. Above this drain voltage the current I_{ds} is controlled only by the gate voltage and ceases to be influenced by the drain voltage. This mode is called saturation.

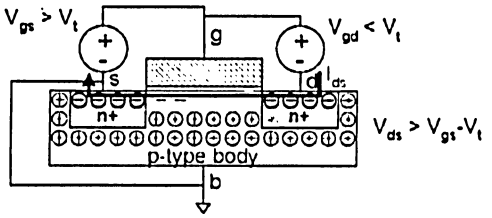
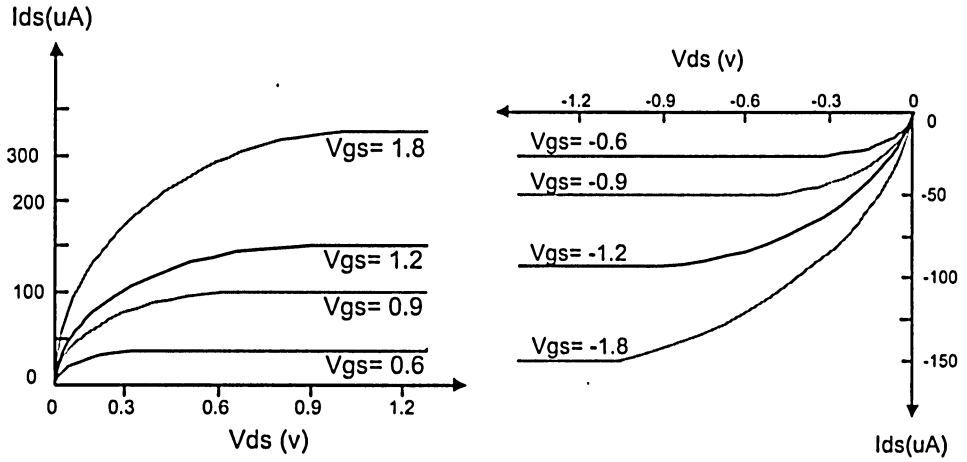


Figure 2.4: NMOS transistor saturation region of operation.

PMOS transistors behave the same way, but the signs are reversed and I-V characteristics is in the third quadrant as shown in Figure 2.5 [3].



(a) NMOS transistors (b) PMOS transistors

Figure 2.5: I-V characteristics of ideal NMOS and PMOS transistors.

The mobility of holes in silicon is typically lower than that of electrons. This means that PMOS transistors provide less current than NMOS transistors of comparable size and hence are slower. Although MOS transistors are symmetrical, by convention we say that majority

carriers flow from source to drain. Because electrons are negatively charged, the source of an NMOS transistor is the more negative of the two terminals. Holes are positively charged so the source of a PMOS transistor is the more positive of two terminals. In complementary CMOS gates, the source is the terminal closer to the supply rail and drain is the terminal closer to the output.

2.2.1 Non-ideal I-V Effects

The ideal I-V model of transistors drain current neglects many effects that are important to modern devices. For example, the saturation current decreases with increasing V_{gs}. This is caused by two effects: velocity saturation and mobility degradation. At high electric field between source and drain (lateral electric field) strengths (V_{ds}/L), carrier velocity ceases to increase linearly with field strength. This is called velocity saturation and results in lower I_{ds} than expected at high V_{ds}. At high vertical electric field strengths (V_{gs}/t_{ox}), the carriers scatter more often. This mobility degradation effect also leads to less current than expected at high V_{gs}. The saturation current of the non-ideal transistor increases slightly with V_{ds}. This is caused by channel length modulation, in which higher V_{ds} increases the size of the depletion region around the drain and thus effectively shortens the channel.

There are several sources of leakage resulting in current flow in nominally OFF transistors. When $V_{gs} < V_T$ the current drops off exponentially rather than abruptly becoming zero. This is called sub-threshold conduction. The threshold voltage itself is influenced by the voltage difference between the source and body: this is called the body effect. The source and drain diffusions are reverse-biased diodes and also experience junction leakage into the substrate or well. The current into the gate (I_g) is ideally zero. However, as the thickness of gate oxides reduces to only a small number of atomic layers, electrons tunnel through the gate, causing some gate current.

Both mobility and threshold voltage decrease with rising temperature. The mobility effect is most important for ON transistors, resulting in lower I_{ds} at high temperature. The threshold effect is most important for OFF transistors, resulting in higher leakage current at high

temperature. Clearly, characteristics degrade with temperature.

2.2.2 Velocity Saturation and Mobility Degradation

According to Equation (2.2) carrier drift velocity and hence current increase linearly with the lateral electric field $E_{lat} = V_{ds}/L$ between source and drain. This is only true for weak fields; at high field strength, drift velocity rolls off due to carrier scattering and eventually saturates at v_{sat} as shown in Figure 2.6 [3]. The carrier velocity may be fitted with Equation (2.8)

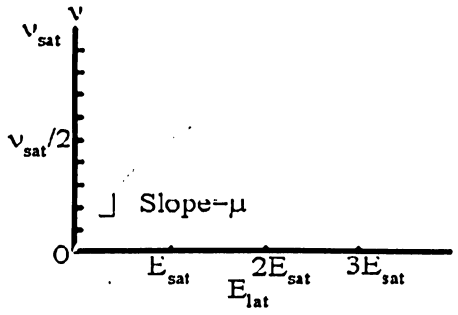


Figure 2.6: Carrier velocity vs. electric field

where E_{sat} is determined empirically. $v_{sat} = \mu E_{sat}$ is in the range of $6 - 10 \times 10^6$ cm/s for electrons and $4 - 8 \times 10^6$ cm/s for holes. This corresponds to a saturation field on the order of 2×10^4 V/cm for NMOS transistors.

$$v = \frac{v_{sat}}{1 + \frac{E_{lat}}{E_{sat}}}, \tag{2.8}$$

recall that without velocity saturation, the saturation current is:

$$I_{ds} = \mu C_{ox} \frac{W}{2L} (V_{gs} - V_t)^2, \tag{2.9}$$

if the transistor were completely velocity saturated, $v = v_{sat}$ and the saturation current becomes:

$$I_{ds} = C_{ox} W (V_{gs} - V_t) v_{sat}, \tag{2.10}$$

where the drain current is quadratically dependent on voltage without velocity saturation and linearly dependent when fully velocity saturated. For moderate supply voltages, transistors operate in a region where the velocity no longer increases linearly with field, but also

is not completely saturated. The α power law model given in Equation (2.11) to Equation (2.13) provides a simple approximation to capture this behavior. α is called the velocity saturation index and is determined by curve fitting measured I-V data. Transistors with long channels or low V_{dd} display quadratic I-V characteristics in saturation and are modeled with $\alpha = 2$. As transistors become more velocity saturated, increasing V_{gs} has less effect on current and α decreases, reaching 1 for transistors that are completely velocity saturated. For simplicity, the model uses a straight line in the linear region. Overall, the model is based on parameters such as $\alpha - \beta - P_c$ and P_v that can be determined empirically from a curve fit of I-V.

Cutoff ($V_{gs} < V_t$):

$$I_d = 0. \tag{2.11}$$

Linear ($V_{ds} < V_{dsat}$):

$$I_d = I_{dsat} \frac{V_{ds}}{V_{dsat}}. \tag{2.12}$$

Saturation ($V_{ds} > V_{dsat}$):

$$I_d = I_{dsat}. \tag{2.13}$$

where

$$I_{dsat} = P_c \mu C_{ox} \frac{W}{2L} (V_{gs} - V_t)^{\alpha}. \tag{2.14}$$

$$V_{dsat} = P_v (V_{gs} - V_t)^{\alpha/2}.$$

As channel lengths become shorter, the lateral field increases and transistors become more velocity saturated (α closer to 1) if the supply voltage is held constant. For example, a transistor with a 2μm channel length begins to show the effects of velocity saturation at V_{dd} above 4V, while a 0.18μm long transistor begins to experience velocity saturation above 0.36V. Figure 3.1 [3] compares I_{ds} for a velocity-saturated NMOS transistor with that of

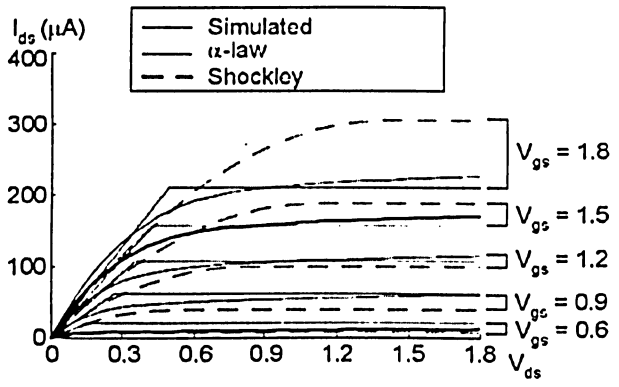


Figure 2.7: I-V characteristics for NMOS transistor with velocity saturation

an ideal transistor and with that predicted by the $\alpha-$power law. The Shockley model grossly over-predicts current at high voltage but the $\alpha-$power fit is reasonably good. As the transistor becomes severely velocity saturated, there is no performance benefit to raising V_{dd}. The low-field mobility of holes is much lower than that of electrons, so PMOS transistors experience less velocity saturation than NMOS for a given V_{dd}. This shows up as a larger value of α for PMOS than for NMOS transistors. Strong vertical electric fields resulting from large V_{gs} cause the carriers to scatter against the surface and also reduce the carrier mobility μ with a smaller μ_{eff}. The $\alpha-$power law captures this effect in the parameter α.

2.2.3 Sub-threshold Conduction

Where the ideal transistor I-V model assumes current only flows from source to drain when $V_{gs} > V_t$. In real transistors, current does not abruptly cut off below threshold, but rather drops off exponentially as given in Equation 2.15

$$I_{ds} = I_{ds0} e^{\frac{V_{gs} - V_t}{n V_T}} \left(1 - e^{\frac{-V_{ds}}{V_T}}\right). \tag{2.15}$$

where this conduction is also known as leakage and often results in undesired current when a transistor is nominally OFF and,

$$I_{ds0} = \mu C_{ox} \frac{W}{L} V_T^2 e^{1.8}. \tag{2.16}$$

where I_{ds0} is the current at threshold and is dependent on process and device geometry, v_T is the thermal voltage and the $e^{1.8}$ term was found empirically. n is a process-dependent term affected by the depletion region characteristics and is typically in the range of 1.4 to 1.5 for CMOS processes. The term $(1 - e^{\frac{-V_{ds}}{v_T}})$ indicates that leakage current is 0 if $V_{ds} = 0$, but increases to its full value when V_{ds} is a few multiples of the thermal voltage v_T (e.g., when $V_{ds} > 50$mV). Conduction through an OFF transistor discharges the capacitor unless it is periodically refreshed or a trickle of current is available to counter the leakage. Leakage also contributes to power dissipation in idle circuits. Leakage increases exponentially as V_t decreases or as temperature rises, so it is becoming a major problem for chips using low supply and threshold voltages. Sub-threshold conduction is exacerbated by drain-induced barrier lowering (DIBL) in which a positive V_{ds} effectively reduces V_t. This effect is especially pronounced in short-channel transistors. It can be modeled as [3]:

$$\acute{V_t} = V_t - \eta V_{ds}, \tag{2.17}$$

where η is the DIBL coefficient, typically in the range of 0.02 to 0.1.

2.3 C-V Characteristics of MOS Transistors

An MOS transistor can be viewed as a four-terminal device with capacitances between each terminal pair as shown in Figure 2.8 [3] where these capacitances are nonlinear and voltage dependent (C-V). However, they can be approximated as simple capacitors when their behavior is averaged across the switching voltages of a logic gate. An MOS transistor consists of gate and diffusion capacitances where the gate capacitance includes components to the body, source and drain, or source alone, depending on operating region [3]. The body is usually tied to the ground (for NMOS) or to the V_{dd} (for PMOS). However, the body of PMOS is also drawn to the ground in small-signal models.

The gate capacitance (C_g) is necessary to attract charge to invert the channel, so high C_g is required to obtain high drain to source current. The gate capacitor can be viewed as a parallel plate capacitor with the gate on top and channel on bottom with the thin oxide dielectric

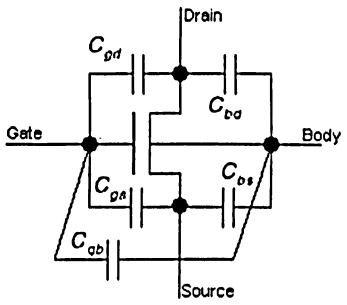


Figure 2.8: MOS transistor capacitances model

between. The source diffusion capacitance (C_{bs}) and drain diffusion capacitance (C_{bd}) arise from the reverse-biased P-N junctions between the source or drain diffusion and the body. These diffusion capacitances are not fundamental to operation of the transistors, but impact circuit performance and hence are called parasitic [3]. The MOS transistor capacitance can be expressed as:

- Gate Capacitance Model:

  When the transistor is OFF (cutoff mode and $V_{gs} = 0$), the channel is not inverted and charge on the gate is matched with opposite charge from the body. This is called gate-to-body capacitance (C_{gb}). As V_{gs} increases but remains below a threshold, a depletion region forms at the surface. This effectively moves the bottom plate downward from the oxide, reducing the capacitance.

  When $V_{gs} > V_t$ (linear mode) the channel inverts and again serves as a good conductive bottom plate. However, the channel charge is connected to the source and drain, rather than the body. At low values of V_{ds}, the channel charge is roughly shared between source and drain, so $C_{gs} = \frac{1}{2}C_{ox}WL$. As V_{ds} increases, the region near the drain becomes less inverted, so a greater fraction of the capacitance is attributed to the source and a smaller fraction to the drain.

  At $V_{ds} > (V_{gs} - V_t)$, the transistor saturates and the channel pinches off. At this point, all the intrinsic capacitance is to the source. Because of pinch off, the capacitance in saturation reduces to $C_{gs} = \frac{2}{3}C_{ox}WL$ for an ideal transistor. The behavior in these

three regions can be approximated as shown in Table 2.1.

Operation Mode	Approximate Value		
Linear	$C_{gs} = \frac{1}{2}C_{ox}WL$	$C_{gd} = \frac{1}{2}C_{ox}WL$	$C_{gb} \approx 0$
Saturation	$C_{gs} = \frac{2}{3}C_{ox}WL$	$C_{gd} \approx 0$	$C_{gb} \approx 0$
Cutoff	$C_{gs} \approx 0$	$C_{gd} \approx 0$	$C_{gb} = C_{ox}WL$

Table 2.1: The approximate values of gate capacitance components

The capacitance per unit area of the gate oxide (C_{ox}) can be expressed as:

$$C_{ox} = \frac{K\epsilon_o}{t_{ox}}, \tag{2.18}$$

where K is the gate oxide dielectric constant, ϵ_o is the permittivity of free space and t_{ox} is the gate oxide thickness.

The gate overlaps the source and drain by a small amount in a real device and also has fringing fields terminating on the source and drain as shown in Figure 2.9 [3] and this leads to additional overlap capacitances and these capacitances are proportional to the width of the transistor [3].

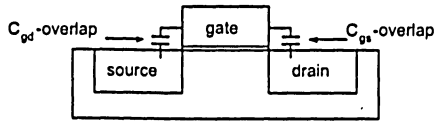


Figure 2.9: Overlap capacitance

Gate-source overlap capacitance ($C_{gs-overlap}$) and gate-drain overlap capacitance ($C_{gd-overlap}$) can be expressed as:

$$C_{gs-overlap} = C_{gso}W \qquad and \qquad C_{gd-overlap} = C_{gdo}W, \tag{2.19}$$

therefore, the C_g can be expressed as:

$$C_g = C_{gs-overlap} + C_0 + C_{gd-overlap}. \tag{2.20}$$

- MOS Diffusion Capacitance Model:

  The drain diffusion capacitor (C_{bd}) depends on both the area and sidewall perimeter

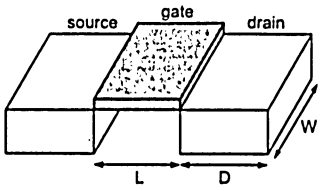


Figure 2.10: Diffusion region geometry

of the drain diffusion region and Figure 2.10 [3] illustrates the drain diffusion region geometry [3].

The drain diffusion area (A_{drain}) and perimeter (P_{drain}) can be expressed as:

$$A_{drain} = WD \qquad and \qquad P_{drain} = 2W + 2D,$$

and

$$C_{bd} = A_{drain} \times C_j + P_{drain} \times C_{jsw}, \tag{2.21}$$

where the area junction capacitance (C_j) has units of capacitance per area and the sidewall capacitance (C_{jsw}) has units of capacitance per length. Since the depletion region thickness depends on the reverse bias, these parasitics capacitance are nonlinear [3]. The source diffusion has a similar parasitic capacitance dependent on A_{source} and P_{source}.

2.4 Propagation Delay

High performance circuit can be defined as high speed, low power dissipation and area also the circuit has to be reliable and cost effective. In every designs there are some paths that are critical and require attention to timing details. When an input changes, the output will retain its old value for at most the propagation delay period. Propagation delay is the time interval at which the input and output voltages cross the half-supply voltage.

Although transistors have complex nonlinear current voltage characteristics, they can be approximated as a switch in series with a resistor, where the effective resistance is chosen to match the average current delivered by transistor. Moreover, transistor gate and diffusion

nodes have capacitance. The delay of basic logic gates can be modeled as RC product of the effective driver resistance and the load. capacitance. Usually, the gate that charges or discharges a node is called the driver and the gates being driven is called load and gates use minimum length devices for least delay, area and power dissipation [3]. Given this, the delay of a logic gate depends on the widths of the transistors in the gate and the capacitance of the load that must be driven.

An NMOS transistor with width of one unit has effective resistance R. The unit width PMOS has a higher resistance that depends on its mobility relative to the NMOS transistor and PMOS effective resistance is usually assumed 2R. Wider transistors have lower resistance for example a PMOS transistor of double unit width has effective resistance R. Parallel and series transistors combine like conventional resistors. When multiple transistors are in parallel, the resistance is lower if they are all ON. In many gates, the worst-case delay occurs when only one of several parallel transistors is ON. In that case, the effective resistance is just that of the single transistor.

According to Elmore delay model and as shown in Figure 2.11 [3] ON transistors are considered as resistors and a chain of transistors can be represented as a RC ladder.

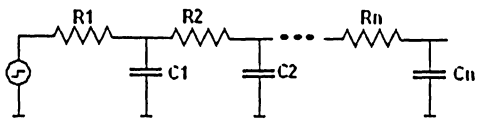


Figure 2.11: RC ladder for Elmore delay

The Elmore delay model estimates the delay of an RC ladder as the sum over each node in the ladder of the resistance R_{n-i} between that node and a supply multiplied by the capacitance on the node [3] as:

$$\tau_{Elmore} = \sum_{i=1}^{n} R_{i-to-source} C_i, \tag{2.22}$$

for example, the RC delay of the circuit given in the Figure 2.11 is

$$\tau = (R_1)C_1 + (R_1 + R_2)C_2 + ... + (R_1 + R_2 + ... + R_n)C_n.$$

According to linear delay model, the propagation delay of a basic logic gates can be written as [11]:

$$\tau_{linear} = gh + p, \tag{2.23}$$

where g is the complexity of the gate, p is the parasitic delay inherent to the gate when no load is attached to the driver and h is fanout. If the load is not identical copies of the driver gate, h can be computed as the capacitance of the external load divided by input capacitance of driver.

Even though the exact analysis of the delay is quite complex, if the nonlinear characteristic of a CMOS gate is taken into account, the delay model can be defined as [4]:

$$\tau = \frac{C_{load} \times V_{dd}}{I} = \frac{C_{load} \times V_{dd}}{\mu C_{ox}(W/2L)(V_{dd} - V_t)^2}, \tag{2.24}$$

where the model shows that delay drastically increases as V_{dd} approaches the V_t. Moreover, decreasing the size of the transistor due to decreasing its current drive makes the gate slower. Therefore, sizing the transistors can be used to optimize the delay of the basic logic gates.

2.5 Turn-on and Turn-off Delay

An accurate delay model of gate switching must be input slope dependent and must distinguish between falling and rising signals [10]. The output transition time of a CMOS structure can be obtained from the modeling of the charging or discharging current that flows during the switching process of the structure and from the amount of charge to be exchanged with the output node as:

$$\tau_{HL} = \frac{C_{load} \cdot \Delta V}{I_{n,max}}, \tag{2.25}$$

and

$$\tau_{LH} = \frac{C_{load} \cdot \Delta V}{I_{p,max}}, \tag{2.26}$$

where ΔV is the output voltage variation, $I_{n,max}$ and $I_{p,max}$ are NMOS and PMOS transistors maximum saturation currents, respectively [2].

Chapter 3

Proposed Model

3.1 Switching Power Dissipation

Any CMOS circuit can be represented by its PMOS network, NMOS network, and the capacitive load connected to its output as shown in Figure 3.1.

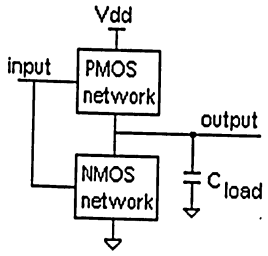


Figure 3.1: Generic representation of a CMOS logic gate.

The switching power dissipation as a component of dynamic power dissipation arises when the capacitive load (C_{load}) of a CMOS circuit is charged through PMOS network to make a voltage transition from 0 to the supply voltage (V_{dd}). During this charging process, current flows from the power supply to the C_{load} and the total power delivered by the power supply to C_{load} per transition can be express as:

$$P_{sw} = \int_0^{\infty} V_{dd} i_{sw}(t) dt = \int_0^{\infty} V_{dd}(C_{load}\frac{dV_{out}}{dt})dt = V_{dd}C_{load}\int_0^{V_{dd}} dV_{out}$$

$$\Rightarrow P_{sw} = C_{load}V_{dd}^2, \tag{3.1}$$

however, the total power stored in the C_{load} per transition can be represented as:

$$P_{C_{load}} = \int_0^\infty V_{out} i_{sw}(t) dt = \int_0^\infty V_{out}(C_{load}\frac{dV_{out}}{dt})dt = C_{load}\int_0^{V_{dd}} V_{out} dV_{out}$$

$$\Rightarrow P_{Cload} = \frac{1}{2} C_{load} V_{dd}^2, \tag{3.2}$$

as shown in Equation (3.2) half of the power delivered by the power supply is stored in the C_{load} and the other half is dissipated as heat in the PMOS network during charging process. During discharging process, no charge is drawn from the power supply, and the energy stored in the C_{load} is dissipated as heat in the conducting NMOS network [3].

The dynamic power dissipation of basic logic gates can be reduced by lowering the power supply voltage. However, since both capacitance and threshold voltage are constant, the speed of the basic logic gates will also decrease with this voltage scaling.

If the voltage of C_{load} switches between V_{dd} and zero at a clock rate (f_{clk}), the power drawn from the power supply would be $f_{clk} C_{load} V_{dd}^2$. However, the switching will not occur at the clock rate, but rather at some reduced rate which is best described as an activity factor (α) times the clock rate [4, 5] and the power drawn from the power supply can be expressed as:

$$P_{sw} = \alpha f_{clk} C_{load} V_{dd}^2. \tag{3.3}$$

A clock has an activity factor of one because it rises and falls every cycle. 2-Input NAND gate has an activity factor of 0.25 due to 25% probability of occurring logic-1 as output. Table 3.1 shows the value of α for several basic logic gates.

	inverter	NAND2	NOR2	NAND3	NOR3
α	0.5	0.25	0.25	0.125	0.125

Table 3.1: The activity factor (α) of basic logic gates.

One common technique for reducing power is to reduce the supply voltage. For CMOS circuits the cost of lower supply voltage is lower performance. Scaling the threshold voltage can limit this performance loss somewhat but results in increased static power dissipation

3.1.1 Operating Frequency and Voltage

The relation between operating frequency and supply voltage of CMOS circuits can be represented as [6]:

$$f \propto \frac{(V_{dd} - V_t)^{\alpha}}{V_{dd}}, \tag{3.4}$$

where the exponent α is experimentally derived (approximately equals to 1.3 for current technology) [6]. This equation can be used to develop a linear equation relating frequency and supply voltage as:

$$V_{norm} = \beta_1 + \beta_2 . f_{norm}, \tag{3.5}$$

where $\beta_1 = \frac{V_t}{V_{max}}$, $\beta_2 = 1 - \beta_1$, V_{norm} is the operation voltage which is normalized to the maximum operating voltage (V_{max}), and f_{norm} is operation frequency which is normalized to the maximum operating frequency (f_{max}). As Equation (3.4) indicates, the frequency drops to zero when supply voltage is reduced to V_t and Equation (3.5) shows that f=0 corresponds to $V_{norm} = \frac{V_t}{V_{max}}$ which for today's technology is approximately 0.3. Reducing the operating frequency by a particular percentage from f_{max} will reduce the operating voltage by a smaller percentage as Equation(3.5) shows. For instance, assume $\beta_1 = 0.3$, reducing the frequency by 50% ($f_{norm} = 0.5$) will reduce the operating voltage by 35% ($V_{norm} = 0.65$). Reducing the voltage by half ($V_{norm} = 0.5$) will reduce the operating frequency by more than half ($f_{norm} = 0.3$). When the dominant source of power dissipation was switching power ($P_{total} = \alpha f_{clk} C_{load} V_{dd}^2$) by halving the supply voltage, the power dissipation will reduce by a factor of four, while the processor's maximum operating frequency will decrease by more than half [6].

In [7], two models have been presented; one to estimate the timing behavior and the other to estimate the power consumption of digital circuits . These models rely on several cell-based extracted parameters. It is indicated that delay at output is dependent on transition at the inputs of a gate and the load capacitance. The delay (τ) was formulated as:

$$\tau = \tau_{in} h + \tau_{par}, \tag{3.6}$$

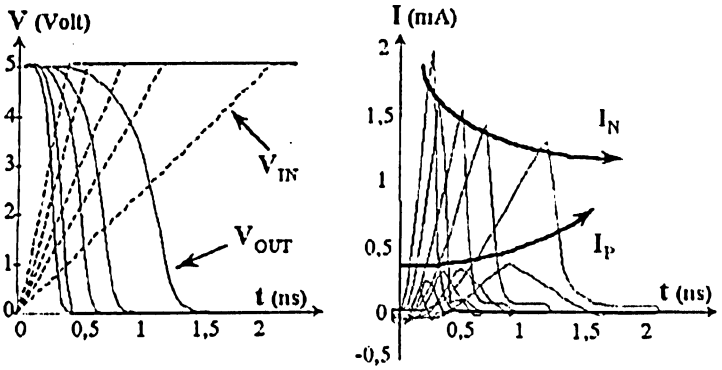


Figure 3.2: Input/output voltage and PMOS/NMOS current of inverter.

where τ_{in} is the time increment per fanout, h is the fanout, and τ_{par} is the delay of no-load [7]. Consequently, the switching power consumption (P_{sw}) of a gate was developed as [7, 8, 9]:

$$P_{sw} = P_{in}h + P_{par}, \tag{3.7}$$

Where P_{in} is the power increment per fanout, h is the fanout, and P_{par} is the power dissipation of no-load condition and they can be express as:

$$P_{in} = \alpha f_{ctk} V_{dd}^2 C_{in} \qquad and \qquad P_{par} = \alpha f_{clk} V_{dd}^2 C_{par},$$

where C_{in} is the input capacitance of fanout, and C_{par} is the internal capacitance that has to be charged for the transition. Therefore, Equation (3.7) can be rewritten as:

$$P_{sw} = \alpha f_{clk} V_{dd}^2 (C_{in}h + C_{par}). \tag{3.8}$$

3.2 Overshoot Power Dissipation

From Figure 3.2 [2] it can be observed that in the case of fast input transition from logic-1 to logic-0, the PMOS current is first negative while the output voltage is greater than V_{dd}. This phenomenon is known as overshoot due to the current that flows from the input to the output through the input-to-output coupling capacitance (C_M) of the basic logic gate [2].

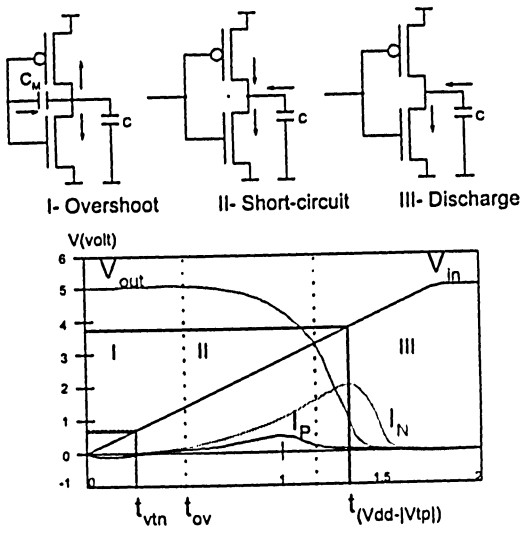


Figure 3.3: Three operating regions and inverter models in each region.

Figure 3.3 [2] shows the different operating regions and inverter models in each region and its associated power dissipation component as:

- I- Overshoot region:

  At the beginning of the input transition, the current that passes through the input-to-output capacitance drives the output voltage beyond the V_{dd}. Due to a partial discharge of the output capacitance from the output node toward the power supply, a negative PMOS transistor current (I_P) will be created. $t_{v_{tn}}$ is the time that input reaches to V_{tn} and during the time t_{vtn} to t_{ov} both the NMOS and PMOS transistors are simultaneously conducting. However, no direct short circuit current is flowing from power supply to ground due to the reversed linear operating mode of the PMOS transistor.

- II- Short circuit region:

  At t_{ov}, which the time necessary to discharge the extra output voltage due to the input-to-output coupling, the output voltage goes below V_{dd}. The PMOS recovers the direct linear operating mode and a direct path occurs between the power supply and ground, resulting in a direct short circuit current.

- III- Discharge region:

  At $t_{v_{dd}-v_{tp}}$ which is the time that the input reaches $V_{dd} - V_{tp}$, the PMOS turns off and the only output load discharge occurs through the NMOS transistor.

During the overshoot part of the transition (Region I), which can be approximated by the time interval defined between t_{vtn} and t_{ov}, the dissipated power is mainly due to the saturated NMOS transistor. The V_{ds} voltages are indeed equal to $V_{dd} + \Delta V_{ov}$ for the NMOS transistor and ΔV_{ov} for the PMOS one, with $\Delta V_{ov} << V_{dd}$. ΔV_{ov} appears here as the extra voltage value induced at the gate output node by the input-to-output coupling capacitance and overshoot power dissipation (P_{ov}) can be expressed as [2]:

$$P_{ov} = \alpha f_{clk} C_{ov} V_{dd}^2, \tag{3.9}$$

where this model is based on the equivalent overshoot capacitance (C_{ov}) that allows author to write overshoot power dissipation in a similar way as switching power dissipation. The C_{ov} can be defined as:

$$C_{ov} = \frac{1}{V_{dd}} \int_{tv_{tn}}^{t_{ov}} I_{n,sat}(t) dt, \tag{3.10}$$

where $I_{n,sat}$ is NMOS saturation current.

The influence of input-to-output coupling capacitance on CMOS inverter delay based on the α-power and n-power law MOS model were presented in [12] and [13], respectively. The load current in these models is assumed as the combination of linear and exponential during the overshooting period. The input-to-output coupling capacitance was modeled where the dynamic behavior of CMOS Inverter was described by a differential equation as [12, 13]:

$$C_{load} \frac{dV_{out}}{dt} = I_p - I_n + C_M \frac{d(V_{in} - V_{out})}{dt}, \tag{3.11}$$

where V_{in} and V_{out} are the gate input and output voltages, respectively. Moreover, I_p and I_n are the PMOS and NMOS transistor currents, respectively. Hence, the differential equation will be solved only for the falling input ramp is expressed as [12]:

$$V_{in} = \begin{cases} V_{dd} & t \leq 0 \\ (1 - \frac{t}{t_{in}}) V_{dd} & 0 \leq t \leq t_f \\ 0 & t > t_f \end{cases}$$

where t_f is the input falling time. The value of C_M, which is strongly voltage dependent, when the input is high (C_M^H) is computed considering the overlap capacitances of PMOS and NMOS drains and the gate to drain overlap capacitance of the NMOS transistor that operates in the linear region as [12]:

$$C_M^H = C_{ox}\left(\frac{W_{n,eff}L_{n,eff}}{2} + X_{D,p}W_{p,eff} + X_{D,n}W_{n,eff}\right), \tag{3.12}$$

with $W_{p,eff}$ and $W_{n,eff}$ are the effective channel width of PMOS and NMOS, respectively. $L_{n,eff}$ is the effective channel length of NMOS, while $X_{D,n}$ and $X_{D,p}$ are the gate drain underdiffusion for the NMOS and PMOS transistors, respectively. For a static input low, the capacitance C_M^L is obtained similarly. Using the α-power-law MOSFET model for short channel devices, the transistor current is given as shown in Equation (3.13) to Equation (3.15) as:

$$\text{when} \quad V_{gs} \leq V_t \quad \Rightarrow \quad I_d = 0, \tag{3.13}$$

$$\text{when} \quad V_{ds} < V_{d0}' \quad \Rightarrow \quad I_d = \frac{V_{ds}}{V_{d0}'} I_{d0}', \tag{3.14}$$

$$\text{when} \quad V_{ds} \geq V_{d0}' \quad \Rightarrow \quad I_d = I_{d0}'. \tag{3.15}$$

where

$$I_{d0}' = I_{d0}\left(\frac{V_{gs} - V_t}{V_{dd} - V_t}\right)^{\alpha}. \tag{3.16}$$

and

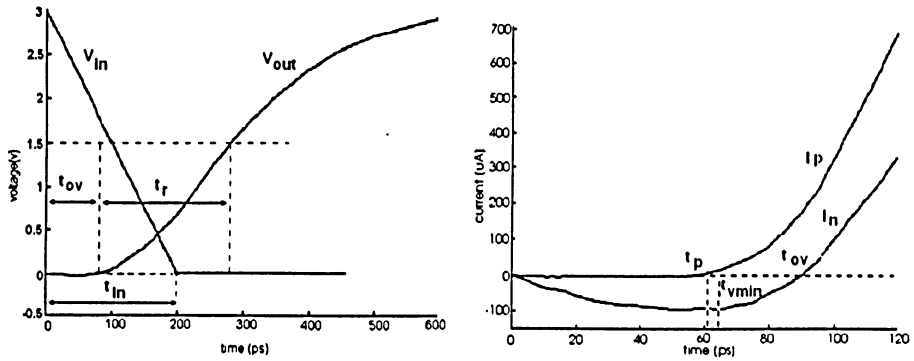
$$V_{d0}' = V_{d0}\left(\frac{V_{gs} - V_t}{V_{dd} - V_t}\right)^{\frac{\alpha}{2}}. \tag{3.17}$$

where the parameter α is velocity saturation index. The V_{d0} is the drain saturation voltage at $V_{gs} = V_{dd}$ and I_{d0} is the drain saturation current at $V_{gs} = V_{ds} = V_{dd}$ [12].

As shown in Figure 3.4 [2] during overshoot period two time points (t_p and t_{vmin}) can be defined [12]. Where $t_p = \frac{|V_{tp}|}{V_{dd}} t_{in}$ is the time when the PMOS transistor is off and t_{vmin} is the time when the output voltage is at its minimum.

The overshooting period can be divided into two parts:

1. When $t < t_p$, PMOS transistor is off and the voltage of the capacitor C_{load} begins to decrease from zero due to a partial discharge through input-to-output coupling capacitance.



(a) Input/output voltages of CMOS inverter

(b) The I_p and I_n during t_{ov}

Figure 3.4: CMOS inverter input/output and transistors current for a falling input

2. When $t > t_p$, PMOS transistor begins to conduct and PMOS current charges the load capacitance C_{load}. Then the voltage of capacitive load increases. Once the output voltage goes up to zero, the NMOS current becomes positive.

At the time $t = t_{ov}$, the charge of C_{load} is zero and the output voltage waveform rises from zero. Therefore, the Equation (3.12) can be expressed as:

$$I_{load} = I_p - I_n - I_{CM}. \tag{3.18}$$

when $t < t_p$ the PMOS is cutoff and the NMOS is in linear mode. Therefore, the Equation (3.12) can be expressed as:

$$(C_{load} + C_M)\frac{dV_{out}}{dt} + \mu C_{ox}\frac{W}{2L}(V_{in} - V_{tn})V_{out} = C_M\frac{dV_{in}}{dt}. \tag{3.19}$$

since V_{gs} is also a function of t this equation can not be easily solved. When $t_p < t < t_{vmin}$ PMOS is on and the expression for output voltage is even more complicated. When $t = t_{vmin}$, the current that flows into capacitive load is zero. Therefore, the Equation (3.12) can be expressed as [12]:

$$I_p(t_{vmin}) = I_n(t_{vmin}) - C_M\frac{dV_{in}}{dt}. \tag{3.20}$$

3.3 Short Circuit Power Dissipation

As the transistors size and threshold voltage become smaller, short circuit power dissipation is no longer a negligible factor since the number of transitions increases due to clock speed up. When the input signal is between V_{tn} and $V_{dd} - |V_{tp}|$, where the V_{tn} is the threshold voltage of NMOS and V_{tp} is the threshold voltage of PMOS, both NMOS and PMOS are turned on and form a DC path. As a result of the DC path, there will be a short circuit current flowing from the power supply to the ground. Short circuit current is dependent on the input transition time, capacitive load, and transistor sizes of the logic gate and it can be expressed as [14]:

$$P_{sc} = I_{mean} V_{dd}. \tag{3.21}$$

Also there is a difference in the short circuit dissipation of an inverter without load and that of an inverter with load. As shown in Figure 3.5 [14] for an inverter without load during the period of $t_1 - t_2$ that short circuit current increases from 0 to I_{max} and the output voltage will be larger than $V_{in} - V_{tn}$. As a consequence, the NMOS transistor will be in saturation

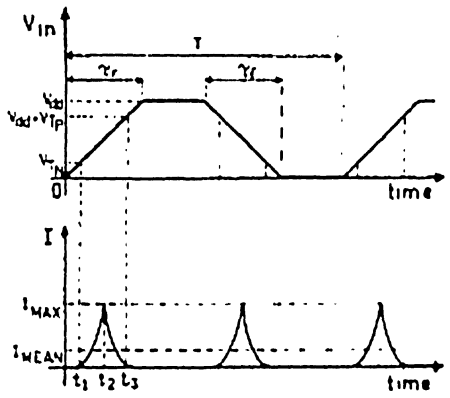


Figure 3.5: Current behavior of an inverter without load.

during this period of time. Therefore, I_{sc} can be expressed as:

$$I_{sc} = I_{sat,nmos} = \frac{\beta}{2}(V_{in} - V_t)^2 \qquad for\ 0 \leq I_{sc} \leq I_{max}. \tag{3.22}$$

assuming that inverter is symmetrical ($\beta_n = \beta_p$), this current will reach its maximum value when $V_{in} = \frac{V_{dd}}{2}$. The mean value of short circuit current (I_{mean}) during one period (T) can be written as:

$$I_{mean} = 2 \times \frac{2}{T} \int_{t_1}^{t_2} I_{sc}(t)dt = \frac{4}{T} \int_{t_1}^{t_2} \frac{\beta}{2}(V_{in}(t) - V_t)^2 dt, \tag{3.23}$$

assuming equal rise and fall times of the input signal and a linear relation between V_{in} and time during its transients as:

$$V_{in}(t) = \frac{V_{dd}}{\tau}.t, \tag{3.24}$$

from Figure 3.5 [14] t_1 and t_2 can be expressed as:

$$t_1 = \frac{V_T}{V_{dd}}.\tau \quad and \quad t_2 = \frac{\tau}{2}, \tag{3.25}$$

therefore,

$$I_{mean} = \frac{2\beta}{T} \int_{\tau/2}^{V_t.\tau/V_{dd}} (\frac{V_{dd}}{\tau}t - V_t)d(\frac{V_{dd}}{\tau}.t - V_t). \tag{3.26}$$

which has the solution as [14]:

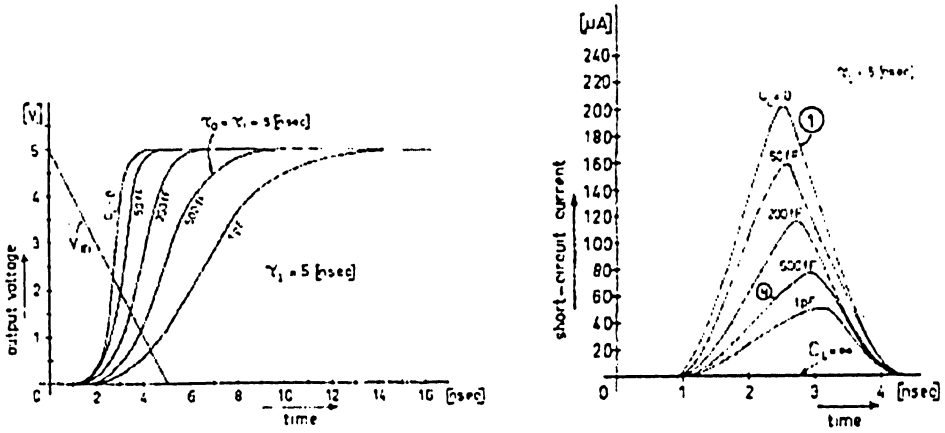
$$I_{mean} = \frac{\beta}{12V_{dd}} (V_{dd} - 2V_t)^3 \frac{\tau}{T}, \tag{3.27}$$

where $\beta = \beta_n = \beta_p$, $V_t = V_{tn} = |V_{tp}|$, $\tau = \tau_r = \tau_f$, $\frac{1}{T}$ = input frequency, τ_f = rising time and τ_r = falling time. From Equation (3.21) and Equation (3.27) the following expression can be derived for the short circuit dissipation of a CMOS inverter without load [14]:

$$P_{sc} = \frac{\beta}{12} (V_{dd} - 2V_t)^3 \frac{\tau}{T}, \tag{3.28}$$

as $\frac{1}{T} = f$, Equation (3.28) shows that short circuit power dissipation component is proportional to input signal frequency. Since V_{dd} and V_t are process-determined, the only design parameters that affect short circuit power dissipation are β, input τ_r and τ_f. For an inverter with capacitive load, the β's of the transistors are determined by requirements on output rise and fall times. In this case the short circuit dissipation depends only on the duration

of the input signal edges. The input signal edges should not be too long, especially in the case of driver circuits that have large β value. Figure 3.6 [14] shows the short circuit current behavior, during a time interval t_1 to t_3 as a function of load capacitance (C_L), for input fall and rise times of 5 ns. curve(1) shows the behavior of the inverter without load. At any time this current is the maximum short circuit current that can occur. Curve (4) shows the short circuit current behavior of the inverter when it is loaded with a characteristic capacitance C_L of $500fF$. In this case the rise and fall time on the output node are equal to the rise and fall time on the input. It was also reported in [14] that if the inverter is loaded in a way



(a) Input and Output (b) Short circuit current

Figure 3.6: Inverter input/output and short circuit current under different capacitive load condition.

causing output rise- and fall-time that are shorter than input rise- and fall-time, then the short circuit dissipation will increase. Therefore, to minimize dissipation, an inverter used as part of a buffer should be designed in such a way that the input rise and fall times are less than or about equal to the output rise- and fall- time in order to guarantee a small short circuit dissipation.

Considering the fact that short circuit power dissipation is delivered by the voltage supply, short circuit power dissipation was been modeled in [15] as:

$$P_{sc} = \alpha f_{sw} V_{dd} \int_T i_{sc}(t) dt, \tag{3.29}$$

where T is the switching period and $i_{sc}(t)$ is short circuit current.

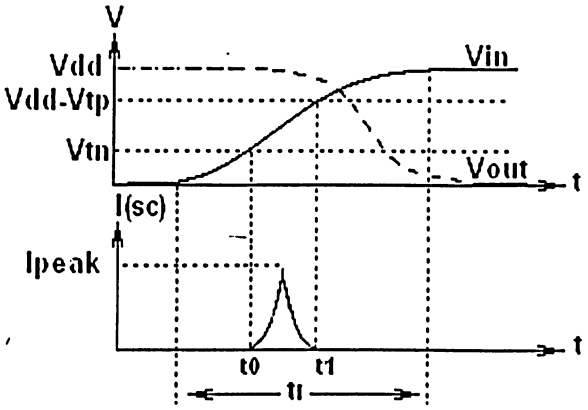


Figure 3.7: Short circuit current

Figure 3.7 [15] shows the short circuit current on PMOS device of an inverter that is driven by a rising ramp input. Assuming the input signal begins to rise at origin, the short circuit current starts to flow at t_0 when the NMOS transistor turns on, and ends at t_1 when the PMOS transistor turns off. During this time interval, the PMOS moves from linear to saturation region. Based on the ramp input signal with a rise time of τ_r, t_0 and t_1 can be expressed as:

$$t_0 = \tau_r \frac{V_{tn}}{V_{dd}} \qquad and \qquad t_1 = \tau_r \frac{V_{dd} - V_{tp}}{V_{dd}}, \tag{3.30}$$

the average short circuit power P_{sc} of the inverter can be specified as:

$$P_{sc} = V_{dd} \int_{t_0}^{t_1} \frac{i_{sc}(t)dt}{t_1 - t_0}, \tag{3.31}$$

when the input rise and fall times are much larger than the output rise and fall times then the short circuit path will be active for a longer period of time and as a result the short circuit current (i_{sc}) is significant. To minimize the short circuit current, it is desirable to have equal input and output rise and fall times. Also, if the supply is lowered to be below the sum of NMOS and PMOS transistors threshold ($V_{dd} < V_{tn} + |V_{tp}|$) the short circuit current can be eliminated because both devices will not be on at the same time for any value of

input voltage [4]. The short circuit power dissipation also decreases as capacitive load at output increases because with large loads the output switches a small amount during the input transition [14]. The short circuit power dissipation has been modeled same as the switching power dissipation using an equivalent capacitance concept [2, 16, 17]:

$$P_{sc} = \alpha f_{clk} C_{sc} V_{dd}^2, \tag{3.32}$$

and

$$C_{sc} = \frac{2}{V_{dd}} \int_0^{\tau_r/2} I_{d,sat}(t)dt, \tag{3.33}$$

where τ_r is the input rise time. The factor 2 and the integration over 0 to $\tau_r/2$ is because the short circuit current is considered symmetric with respect to its maximum value. The $I_{d,sat}$ equation of the Klaassen's model is [16]:

$$I_{d,sat} = \beta[V_g - V_t - \alpha V_{d,sat}] \frac{LE_p}{1 + \frac{E_p}{E_c}}, \tag{3.34}$$

where L is the effective length, E_p is the value of field where the gate loses control over the channel and E_c is the field characterizing velocity saturation.

3.4 Dynamic Power Dissipation Proposed Model

Since dynamic power dissipation that is simulated at output node has a second degree relation with fanout and it is equal sum of switching power dissipation and short circuit power dissipation. Based on C_M second order effect on basic logic gates delay [12] and dynamic power dissipation simulation results it can be concluded that P_{ov} has a second degree relation with fanout and modeled as:

$$P_{ov} = \alpha f_{clk} V_{dd}^2 C_M h^2. \tag{3.35}$$

Considering the fact that P_{sc} value is maximum at load zero and decreases by increasing the load, it can be modeled as:

$$P_{sc} = \alpha f_{clk} V_{dd}^2 (C_{par,sc} - C_{in,sc} h), \tag{3.36}$$

where $C_{par,sc}$ is the short circuit equivalent capacitance due to zero load condition and $C_{in,sc}$ is the short circuit equivalent capacitance due to each fanout. Assuming $C_{in,sc} << C_{par,sc}$ the value of short circuit power is maximum at load zero and is zero at $C_{par,sc} = \infty$.
The dynamic power dissipation due to lumped capacitance was modeled in [2, 5] as:

$$P_{dyn} = \alpha f_{clk} V_{dd}^2 [C_{load} + C_{ov} + C_{sc}], \tag{3.37}$$

however substituting (3.8), (3.9) and (3.32) into (1.1), the existing model of dynamic power dissipation can be defined as:

$$P_{dyn} = \alpha f_{clk} V_{dd}^2 [C_{in} h + C_{par} + C_{ov} + C_{sc}], \tag{3.38}$$

the existing dynamic power dissipation model does not consider C_M effects and has a first order relation with fanout. However by substituting (3.8), (3.35) and (3.36) into (1.1), the dynamic power dissipation of basic logic gates that are loaded with actual gate can be defined as:

$$P_{dyn} = \alpha f_{clk} V_{dd}^2 [C_M h^2 + (C_{in} - C_{in,sc}) h + (C_{par} + C_{par,sc})], \tag{3.39}$$

considering the capacitances that are responsible for power increment per fanout as $C_{in,gate}$ and parasitic power as $C_{par,gate}$, the dynamic power dissipation can be written as:

$$P_{dyn} = \alpha f_{clk} V_{dd}^2 [(C_M h^2 + C_{in,gate} h + C_{par,gate})], \tag{3.40}$$

in [3] it is indicated that capacitance of a circuit is proportional to transistor size, knowing that all the devices have the same length, then dynamic power dissipation is proportional to transistor width (W) and Z can be defined as:

$$Z = \frac{W_{gate}}{W_{min}}, \tag{3.41}$$

where, W_{gate} is the gates transistor width and W_{min} is the minimum transistor width. The final expression of the dynamic power dissipation model can be written as:

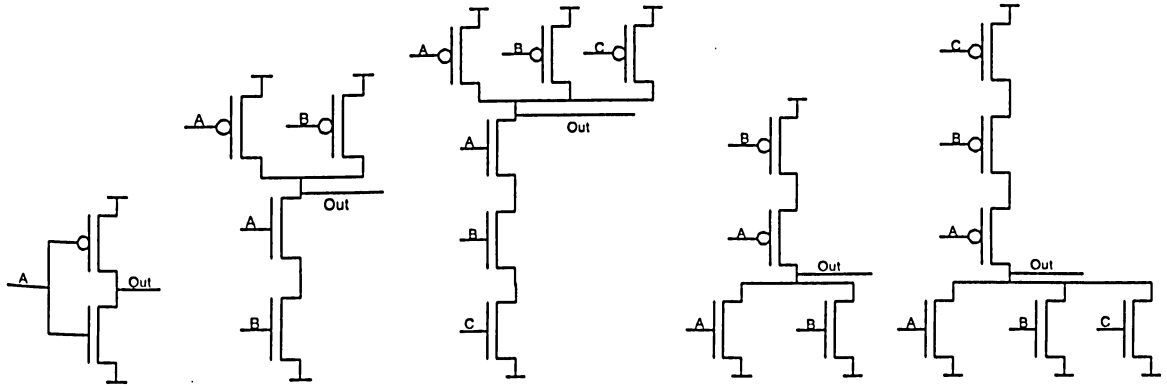
$$P_{dyn} = \alpha_{gate} f_{clk} V_{dd}^2 Z (C_M h^2 + C_{in,gate} h + C_{par,gate}). \tag{3.42}$$

Considering the fact that proposed model has a second-order relation with fanout shows that there is a great difference between using actual gate and equivalent lump capacitance as load. Since dynamic power dissipation caused by actual gates has a second-order relation with fanout and dynamic power dissipation caused by equivalent lumped capacitance has a first-order relation with fanout. It is not practical to use lumped capacitance instead basic logic gate as a load.

Chapter 4

Model Validation

Using 0.13μm CMOS technology and netlist (Cadence) to validate the proposed model, different basic logic gates have been considered, such as inverter, 2-, 3-input NAND and 2-, 3-input NOR gates as illustrated in the Figure 4.1.



(a) inverter (b) 2-input NAND (c) 3-input NAND (d) 2-input NOR (e) 3-input NOR

Figure 4.1: Transistor level circuits of basic logic gates

Transistors of the basic logic gates were sized as shown in Table 4.1.

	inverter	NAND2	NOR2	NAND3	NOR3
W_{PMOS}	0.52μm	0.52μm	1.04μm	0.52μm	1.56μm
W_{NMOS}	0.26μm	0.52μm	0.26μm	0.78μm	0.26μm

Table 4.1: Transistor width of minimum size basic logic gates (L=0.13μm)

Then a 5-stage inverter ring oscillator with minimum size transistors (W_{PMOS} = 0.52 μm and W_{NMOS} = 0.26 μm) was designed to reach the highest frequency that 0.13-μm CMOS

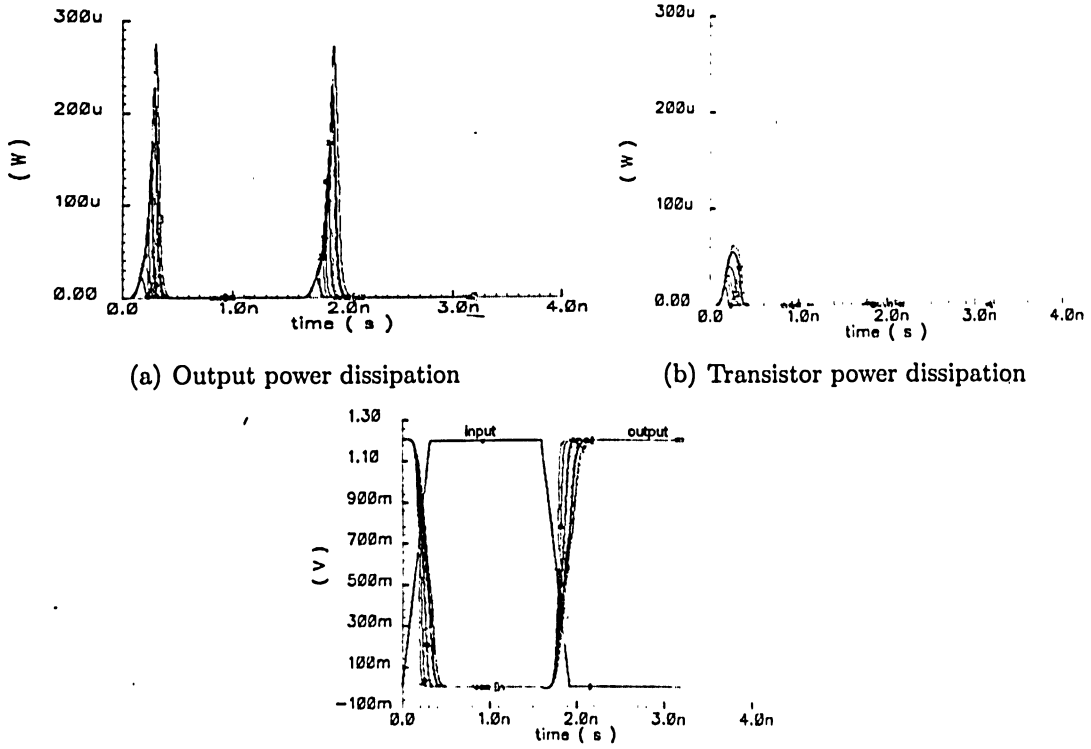
technology can support. The oscillation frequency was measured to be 2 GHz.

Since the 3-input NOR gate with fan-out-of-10, which has the highest delay, failed to perform at this frequency a pulse signal with frequency of about 0.3 GHz, duty cycle of 50%, rise and fall time of 320 ps (10% of the signal period) was generated. Then the pulse was applied to input (A) of the gates which are the input of the closest transistor to the output of the gates as shown in the Figure 4.1. The other inputs of 2- and 3-input NAND gates were connected to the V_{dd}. With respect to 2- and 3-input NOR gates, the other inputs were connected to the ground. Therefore, these transistors were ON and operating in the linear region during the output transitions.

4.1 Dynamic Power Dissipation Due to Actual Gate

In the first part of the simulation, actual gates were used as load (load gate) by connecting identical basic logic gates to the output node of the gate under test (driver gate). For each of the considered gates, the load condition was increased from fanout-of-0 to fanout-of-10 by one. For one charging and discharging period, the output and transistor power dissipation were simulated. Since the P_{sw} and P_{sc} currents pass through the basic logic gates transistor, the simulated power dissipation at transistor is equal to $(P_{sw} + P_{sc})$ and the simulated power at output node is equal to P_{dyn}. The basic logic gates input/output voltage, output/transistor power dissipation waveform are shown in Figure 4.2 to Figure 4.6. Output voltage delay and output/transistor power dissipation of basic logic gates increase with the fanout and complexity of the basic logic gates. The output power dissipation of basic logic gates are much greater than transistor power dissipation as a result of over-shoot power dissipation.

Transistor power dissipation of nand gates were simulated at PMOS and power dissipation of nor gates were simulated at NMOS. The transistor power of nand gates are dissipated during rising process and the transistor power of nor gates are dissipated during falling process. Other than these periods, the transistor power dissipations are almost zero.



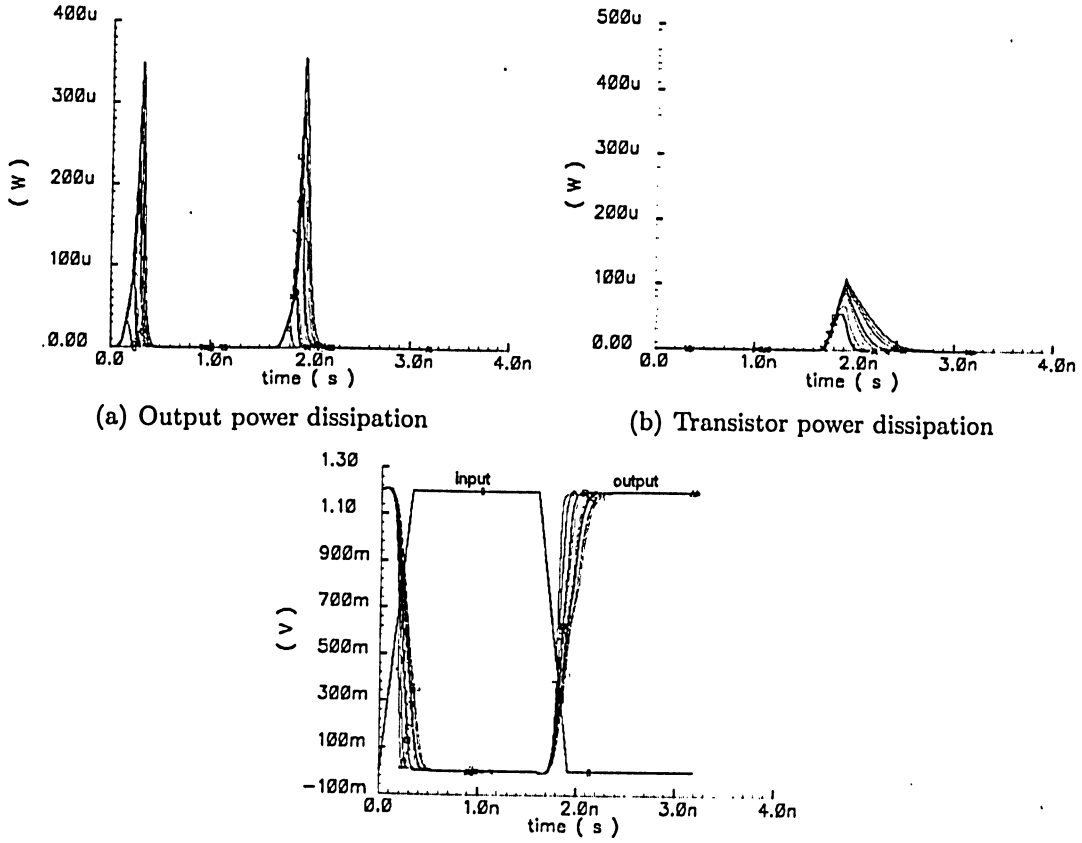
(a) Output power dissipation

(b) Transistor power dissipation

(c) Input and output voltage

Figure 4.2: Inverter input/output voltage and output/transistor power dissipation due to actual gate ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.26\mu m$).

Figure 4.2 illustrates the inverter output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. Also their waveforms increase from zero to maximum value during charging and discharging period of output capacitive load. Power is also dissipated for longer period of time during charging than discharging period. The output power dissipation of inverter with minimum size transistor is in the rang of μW and has almost equal maximum values at charging and discharging. The maximum value of output power dissipation of inverter is much greater than its transistor power dissipation as a result of over-shoot power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



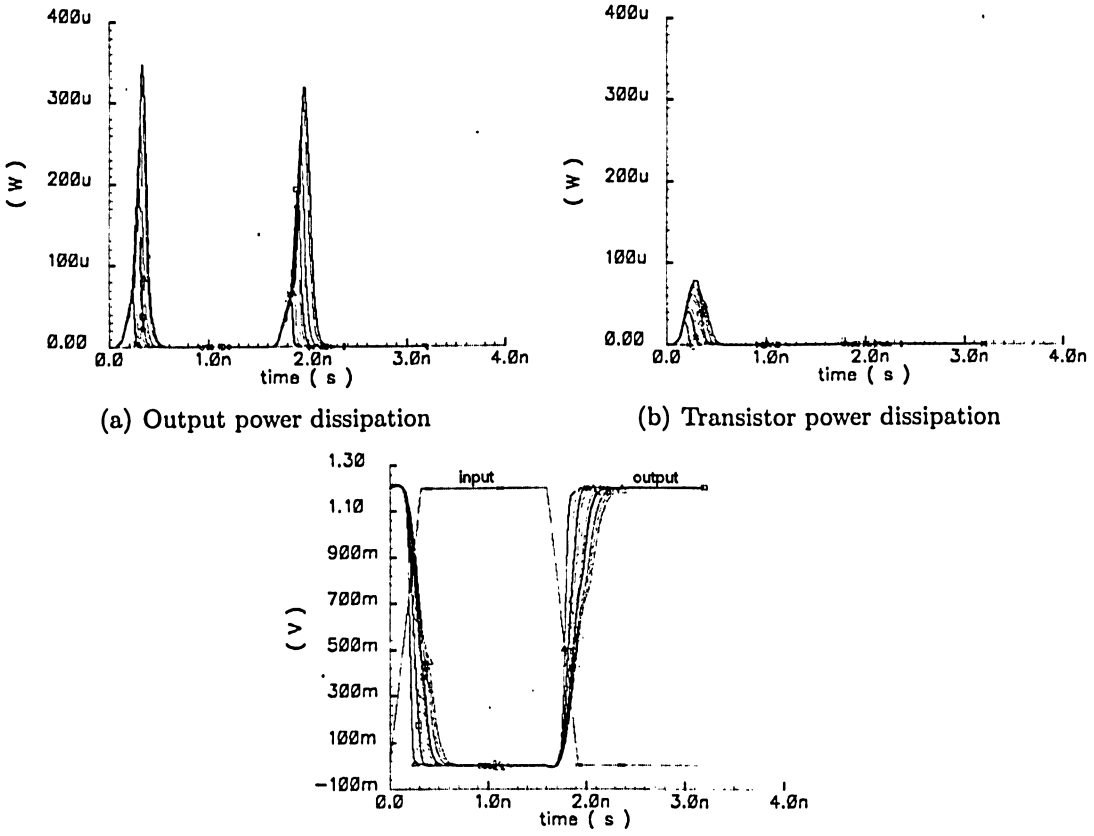
(a) Output power dissipation

(b) Transistor power dissipation

(c) Input and output voltage

Figure 4.3: Nand2 input/output voltage and output/transistor power dissipation due to actual gate ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.52\mu m$).

Figure 4.3 illustrates the 2-input nand gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. Power is also dissipated for longer period of time during charging than discharging period of capacitive load that is connected to output. The output power dissipation of 2-input nand gate with minimum size transistor is in the rang of μW. The maximum value of output power dissipation of 2-input nand gate is much greater than its transistor power dissipation as a result of over-shoot power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



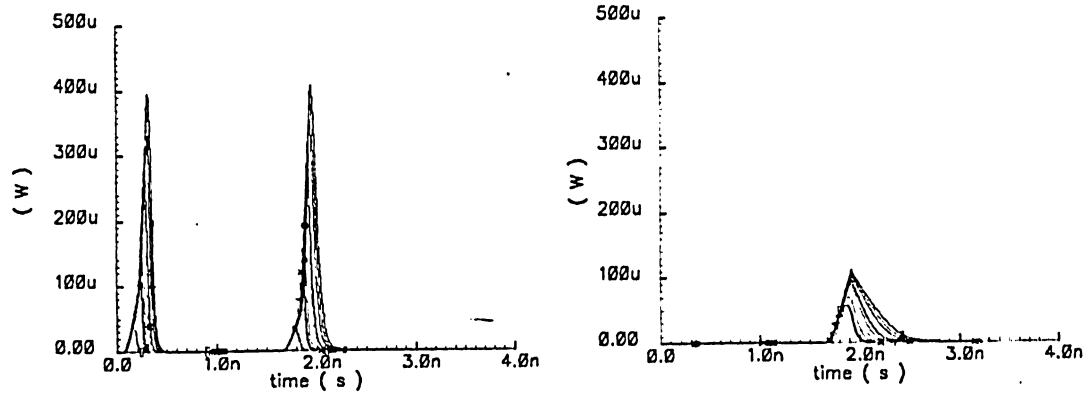
(a) Output power dissipation

(b) Transistor power dissipation

(c) Input and output voltage

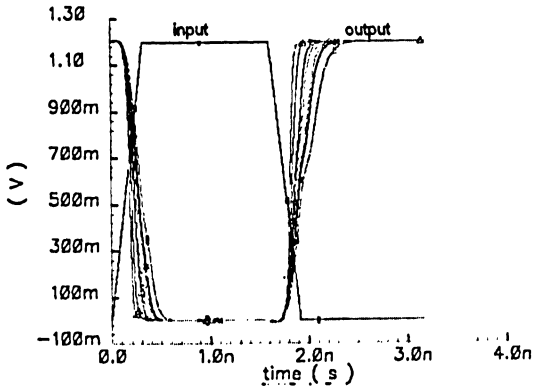
Figure 4.4: Nor2 input/output voltage and output/transistor power dissipation due to actual gate ($W_{PMOS} = 1.04\mu m$ and $W_{NMOS} = 0.26\mu m$).

Figure 4.4 illustrates the 2-input nor gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. Also their waveforms increase from zero to maximum value during charging and discharging period of output capacitive load. The output power dissipation of 2-input nor gate with minimum size transistor is in the rang of μW. The maximum value of output power dissipation of 2-input nor gate is much greater than its transistor power dissipation as a result of over-shoot power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



(a) Output power dissipation

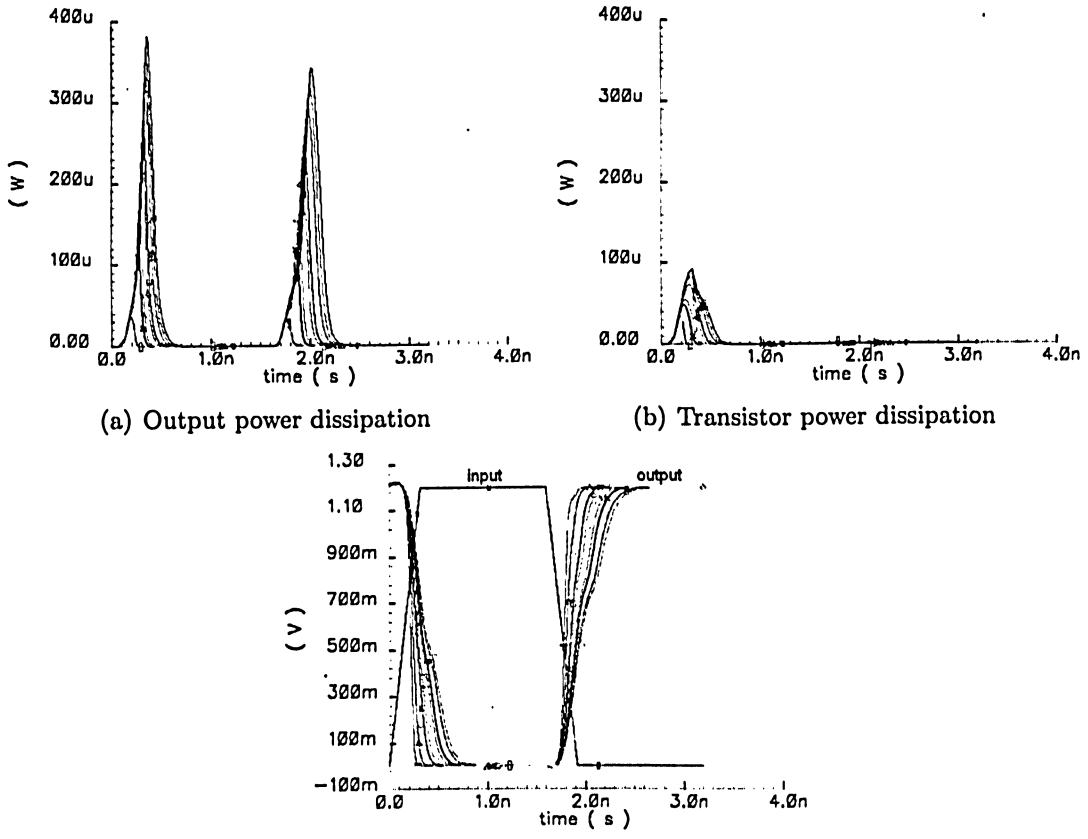
(b) Transistor power dissipation



(c) Input and output voltage

Figure 4.5: Nand3 input/output voltage and output/transistor power dissipation due to actual gate ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.78\mu m$).

Figure 4.5 illustrates the 3-input nand gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. Power is also dissipated for longer period of time during charging than discharging period of capacitive load that is connected to output. The maximum value of output power dissipation of 3-input nand gate is much greater than its transistor power dissipation as a result of over-shoot power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



(a) Output power dissipation

(b) Transistor power dissipation

(c) Input and output voltage

Figure 4.6: Nor3 input/output voltage and output/transistor power dissipation due to actual gate ($W_{PMOS} = 1.56\mu m$ and $W_{NMOS} = 0.26\mu m$).

Figure 4.6 illustrates the 3-input nor gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. Power is also dissipated for longer period of time during charging than discharging period of capacitive load that is connected to output. The maximum value of output power dissipation of 3-input nor gate is much greater than its transistor power dissipation as a result of over-shoot power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.

The average power dissipation of basic logic gates were obtained from the simulation plots. For example the simulation plots of average and dynamic power dissipation of inverter with no-load are shown in the Figure 4.7.

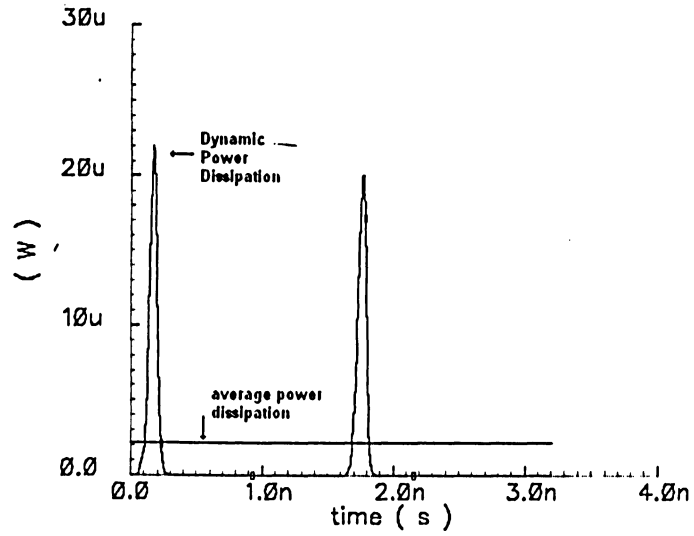


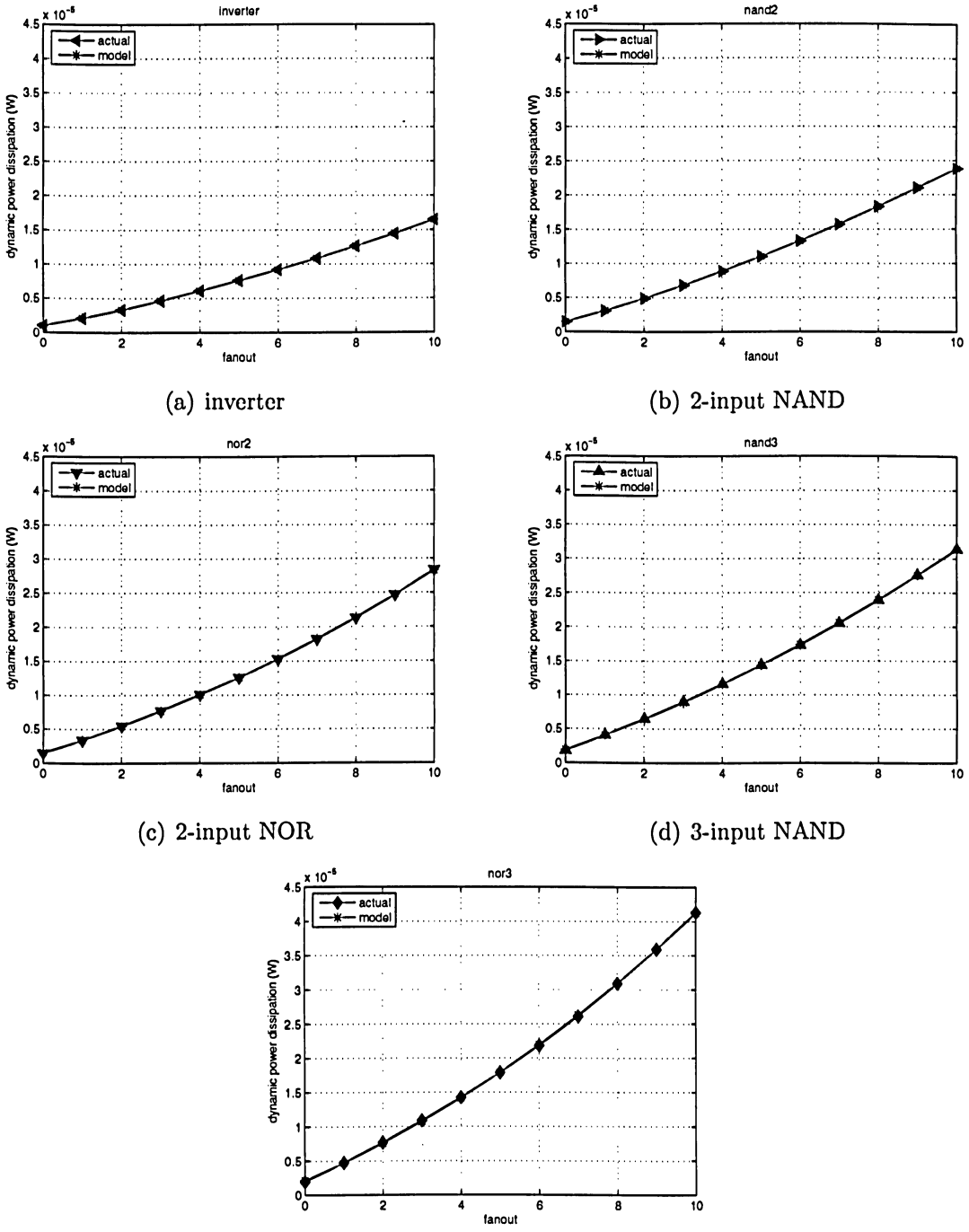
Figure 4.7: Average and dynamic power dissipation of inverter with no-load.

The value of C_M, $C_{gate,in}$ and $C_{gate,par}$ of minimum size basic logic gates were extracted from simulation results as shown in Table 4.2.

	inverter	NAND2	NOR2	NAND3	NOR3
C_M (fF)	0.10	0.14	0.21	0.20	0.33
$C_{gate,in}$ (fF)	2.45	3.52	3.87	4.54	5.34
$C_{gate,par}$ (fF)	2.11	3.22	3.49	4.32	4.78

Table 4.2: Basic logic gates C_M, $C_{gate,in}$ and $C_{gate,par}$.

The power dissipation of minimum size basic logic gates loaded with actual gate (P_{act}) were simulated and compared with the proposed model (P_{model}) under different loading conditions as shown in Table 4.3 and Figure 4.8. As illustrated in this table and figure, fanout is increased from zero to ten and the basic logic gates power dissipation is simulated. The basic logic gates power dissipation are in the range of μW and increase with fanout and the complexity of gate. The highest error between basic logic gates P_{act} and P_{model} happen at fanout-of-zero. The reason of high discrepancy at fanout-of-zero is that the output of basic



(a) inverter

(b) 2-input NAND

(c) 2-input NOR

(d) 3-input NAND

(e) 3-input NOR

Figure 4.8: Comparison of proposed model (P_{model}) and basic logic gates dynamic power dissipation due to actual gate (P_{act}) simulation results under different loading conditions.

logic gates is floating and it makes estimating of $C_{gate,par}$ difficult. The error at fanout-of-zero increase with the complexity of basic logic gates. The basic logic gates P_{act} have a second-order relation with fanout and P_{model} show a high agreement with the basic logic gates P_{act} with maximum error of 3.20% and average error of -0.28%. As it is shown in the Table 4.3 The model sometimes under estimate the value of dynamic power dissipation and when the graph of dynamic power dissipation cross the model the value of error is minimum. The minimum value of errors usually happens at fanout-of-4 where the delay for a process (in picoseconds) is 1/3 to 1/2 of the channel length (in nanometers) [3].

	Inverter			NAND2			NOR2			NAND3			NOR3		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	0.98	0.95	3.20	1.46	1.45	0.60	1.46	1.57	-7.42	1.80	1.95	-8.08	1.94	2.15	-10.8
1	2.07	2.09	-1.20	3.08	3.10	-0.70	3.40	3.40	-0.07	4.10	4.08	0.58	4.75	4.70	0.96
2	3.30	3.33	-0.94	4.88	4.88	-0.04	5.50	5.42	1.43	6.50	6.39	1.75	7.70	7.56	1.86
3	4.64	4.66	-0.38	6.80	6.79	0.11	7.72	7.63	1.22	9.00	8.87	1.39	10.9	10.7	1.76
4	6.09	6.07	0.26	8.86	8.83	0.31	10.1	10.0	0.93	11.6	11.5	0.50	14.3	14.2	0.71
5	7.62	7.58	0.51	11.0	11.0	-0.03	12.6	12.6	0.07	14.4	14.4	0.08	17.9	17.9	-0.05
6	9.20	9.18	0.24	13.3	13.3	-0.02	15.3	15.4	-0.34	17.3	17.4	-0.65	21.8	22.0	-0.72
7	10.9	10.9	-0.14	15.7	15.7	-0.21	18.2	18.3	-0.53	20.5	20.6	-0.57	26.1	26.3	-0.79
8	12.6	12.6	-0.02	18.3	18.3	0.03	21.3	21.4	-0.60	23.9	24.0	-0.41	30.9	31.0	-0.17
9	14.5	14.5	-0.35	21.0	21.0	0.08	24.7	24.7	-0.18	27.6	27.6	0.15	35.9	35.9	0.01
10	16.5	16.5	0.19	23.8	23.8	-0.02	28.4	28.2	0.54	31.4	31.3	0.32	41.3	41.1	0.37
	Maximum $E\%$		3.20												
	Average $E\%$		-0.28												

Table 4.3: Comparison of proposed model (P_{model}) and basic logic gates dynamic power dissipation simulation results due to actual gate (P_{act}) under different loading conditions (power values are in μW and $E\% = \frac{P_{act}-P_{model}}{P_{act}} \times 100$).

4.2 Dynamic Power Dissipation Due to Equivalent Lumped Capacitance

Equivalent lumped capacitance of each basic logic gates were estimated using simulation results as shown in Table 4.4.

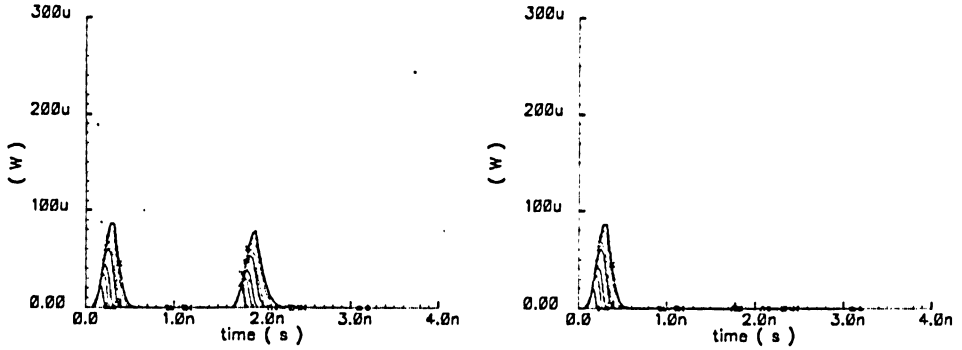
	inverter	NAND2	NOR2	NAND3	NOR3
Equivalent Lumped Capacitance	2.85 fF	4.09 fF	4.57 fF	5.30 fF	6.38 fF

Table 4.4: The value of equivalent lumped capacitance of basic logic gates.

Basic logic gates input/output voltage and output/transistor power dissipation simulation results due to lumped capacitance are shown in Figure 4.9 to Figure 4.13.

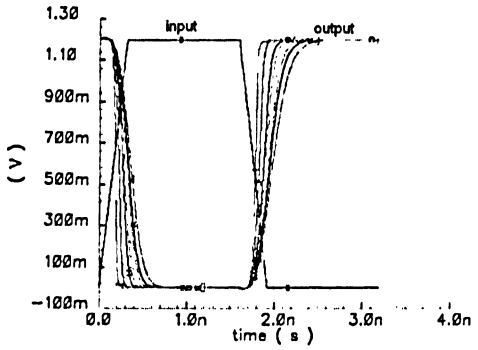
As illustrated in these Figures, the output power dissipation of basic logic gates is in rang of μW and increase with increasing the fanout and they have a first-order relation with fanout. The maximum value of power dissipation and duration of power dissipation increase with fanout and complexity of basic logic gates. The output power dissipation of basic logic gates are also twice of transistor power dissipation since the output power dissipation are equal the sum of PMOS and NMOS transistors power dissipations. Moreover, power dissipation of basic logic gates increase with complexity of gates and 3-input nor gate has the highest power dissipation for each fanout. The delay of basic logic gates increase with the fanout and complexity of basic logic gates.

Comparison of simulation results show that transistor power dissipation of basic logic gates that are loaded by actual gates are less than transistor power dissipation of basic logic gates that are loaded by lumped capacitance. Also, the delay of basic logic gates that are loaded by actual gates are less than the delay of basic logic gates that are loaded by lumped capacitance.



(a) Output power dissipation

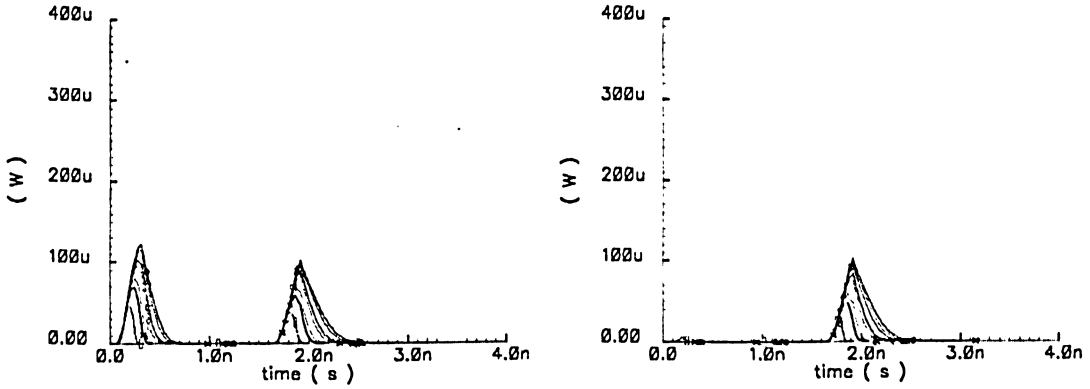
(b) Transistor power dissipation



(c) Input and output voltage

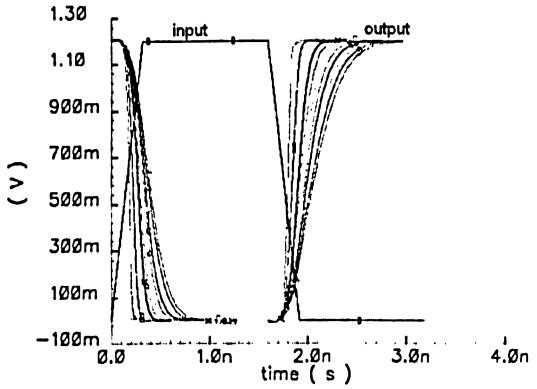
Figure 4.9: Inverter input/output voltage and transistor/output power dissipation due to lumped capacitance.

Figure 4.9 illustrates the inverter gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. The maximum value of output power dissipation of inverter gate is almost same as its transistor power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



(a) Output power dissipation

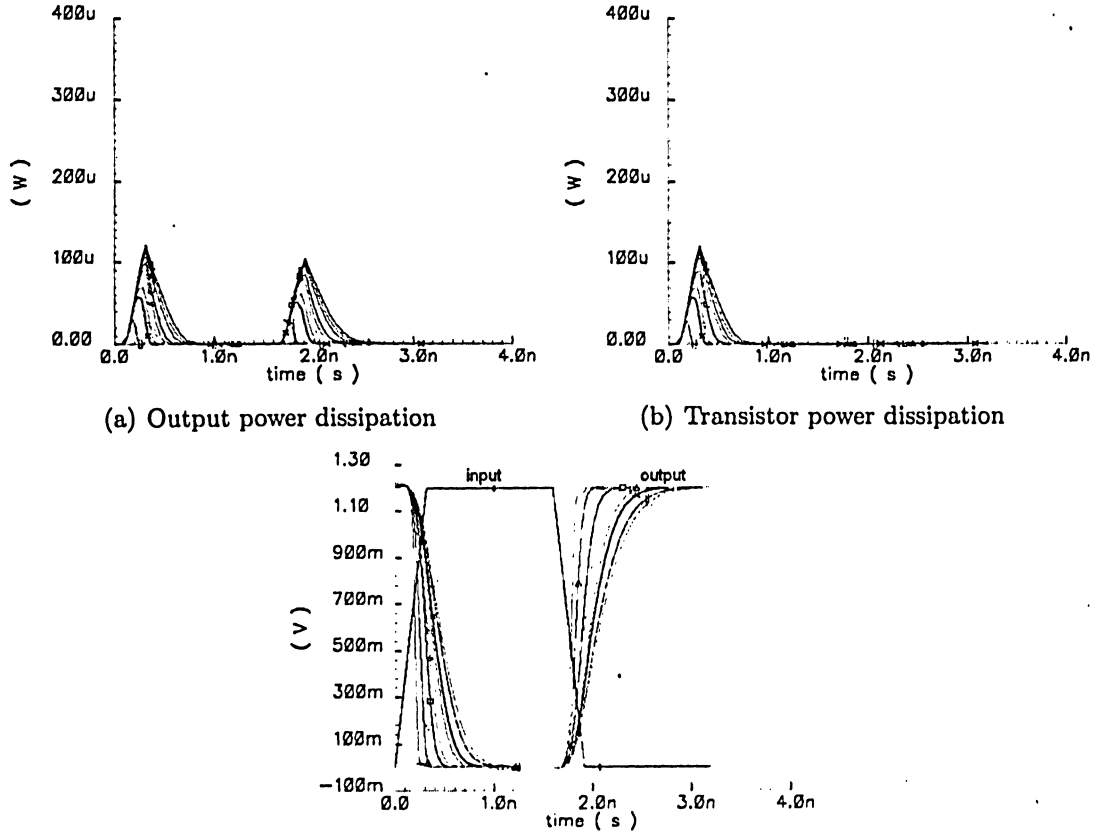
(b) Transistor power dissipation



(c) Input and output voltage

Figure 4.10: Nand2 input/output voltage and output/transistor power dissipation due to lumped capacitance ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.52\mu m$).

Figure 4.10 illustrates the 2-input nand gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. The maximum value of output power dissipation of 2-input nand gate is almost same as its transistor power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



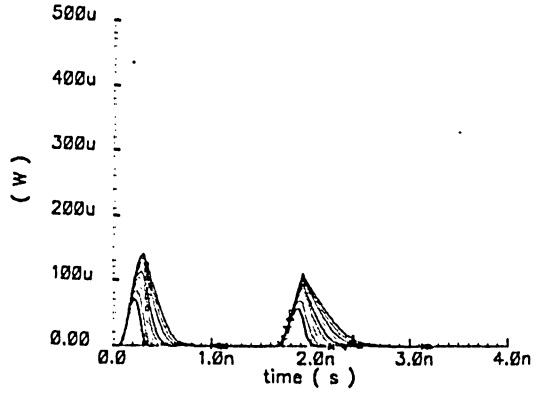
(a) Output power dissipation

(b) Transistor power dissipation

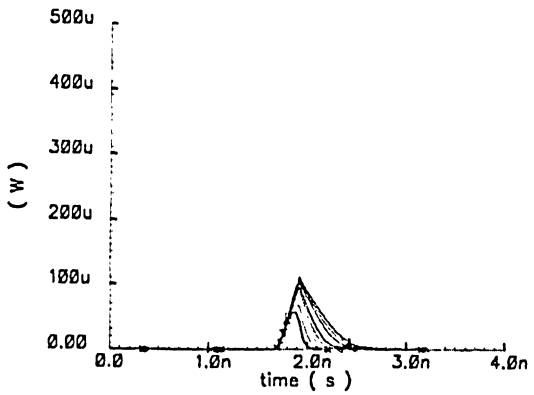
(c) Input and output voltage

Figure 4.11: Nor2 input/output voltage and output/transistor power dissipation due to lumped capacitance($W_{PMOS} = 1.04\mu m$ and $W_{NMOS} = 0.26\mu m$).

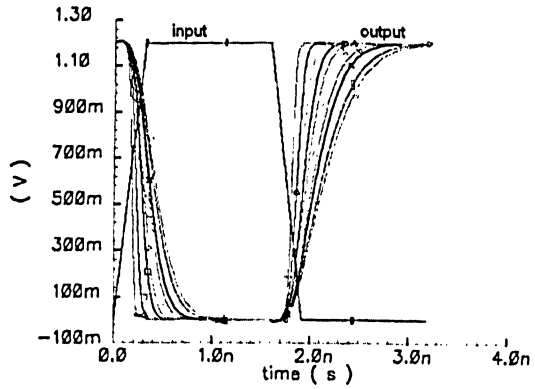
Figure 4.11 illustrates the 2-input nor gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. The maximum value of output power dissipation of 2-input nor gate is almost same as its transistor power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



(a) Output power dissipation



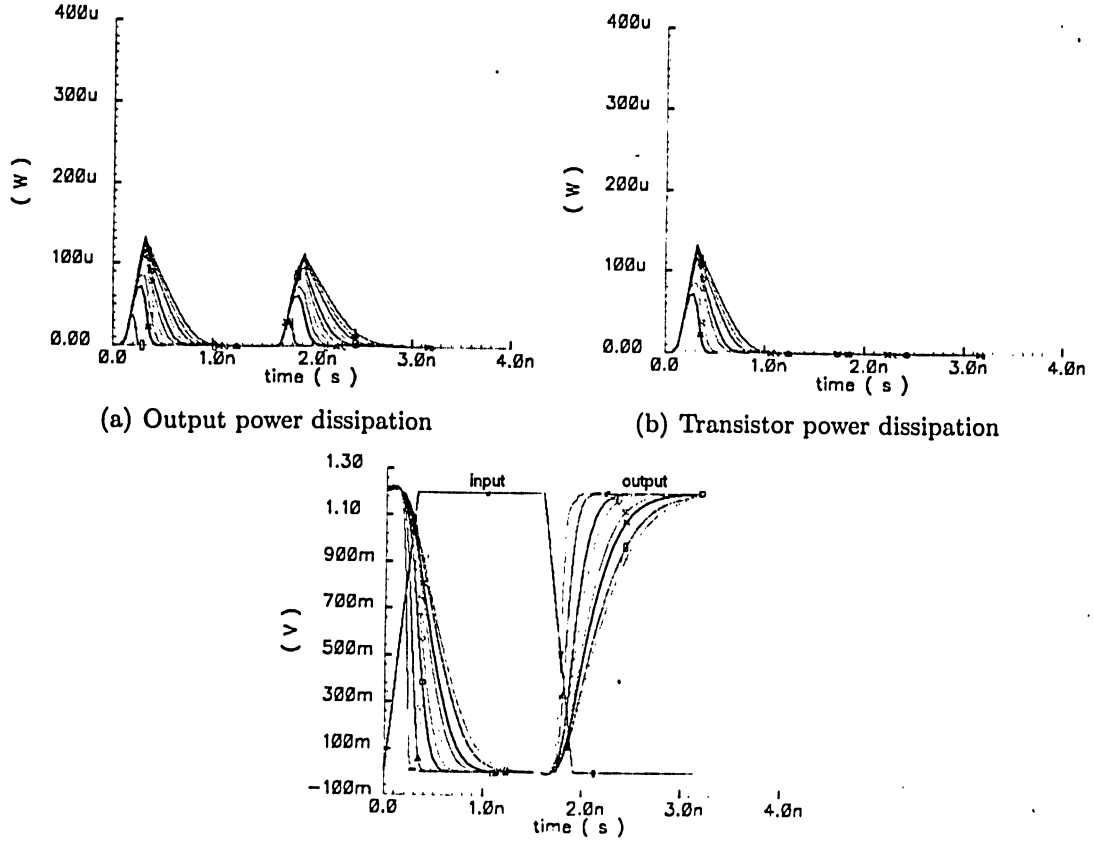
(b) Transistor power dissipation



(c) Input and output voltage

Figure 4.12: Nand3 input/output voltage and output/transistor power dissipation due to lumped capacitance ($W_{PMOS} = 0.52\mu m$ and $W_{NMOS} = 0.78\mu m$).

Figure 4.12 illustrates the 3-input nand gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. The maximum value of output power dissipation of 3-input nand gate is almost same as its transistor power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.



(a) Output power dissipation

(b) Transistor power dissipation

(c) Input and output voltage

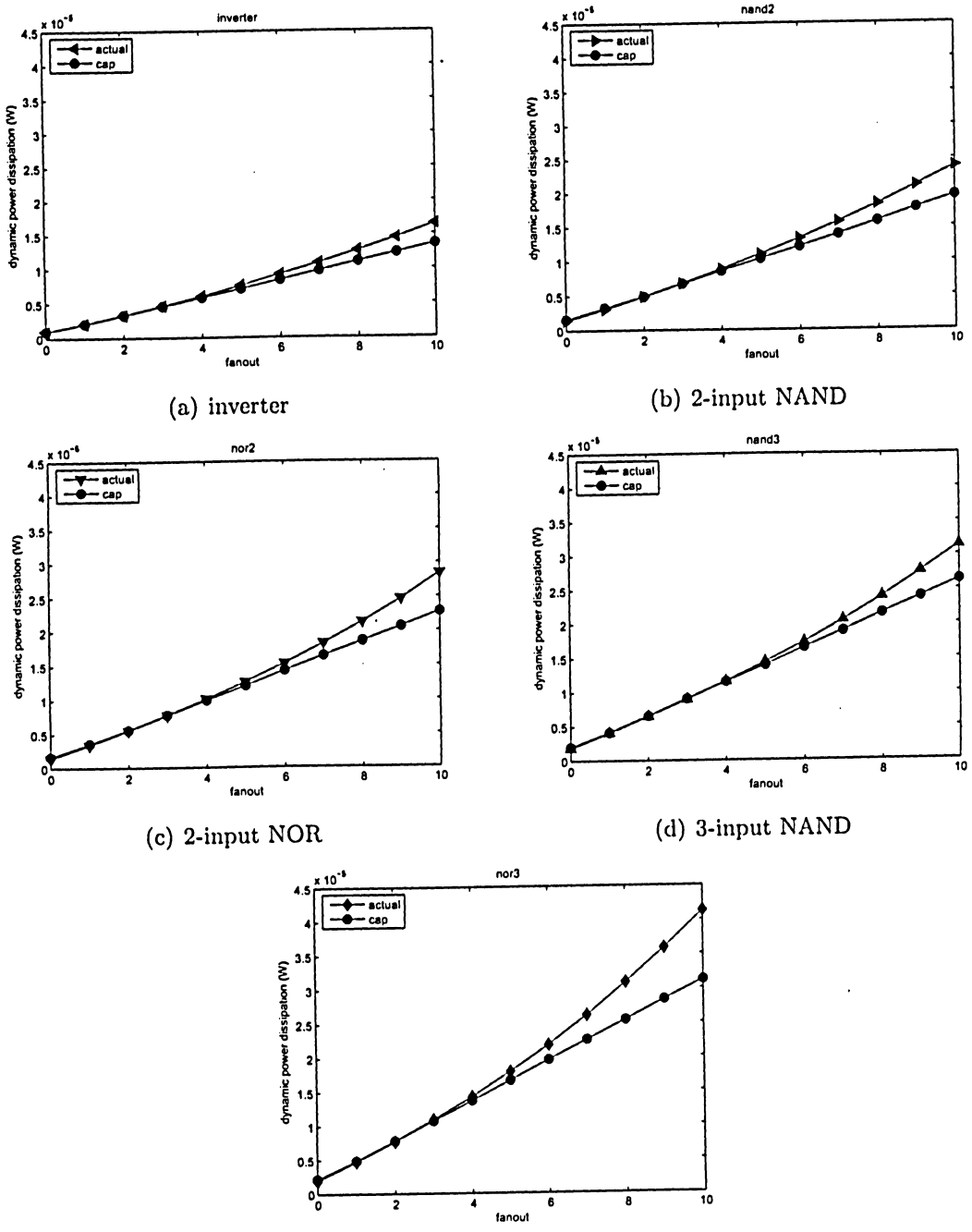
Figure 4.13: Nor3 input/output voltage and output/transistor power dissipation due to lumped capacitance ($W_{PMOS} = 1.56\mu m$ and $W_{NMOS} = 0.26\mu m$).

Figure 4.13 illustrates the 3-input nor gate output voltage and transistor/output power dissipation waveform of fanout-of-0 to fanout-of-10. As shown in this figure the transistor/output power dissipation maximum value and duration of power dissipation increase with increasing fanout. The maximum value of output power dissipation of 2-input nor gate is almost same as its transistor power dissipation. The output voltage waveform shows the over-shoot effect by exceeding V_{dd} at the beginning of transition (for rising input) and same for falling input the output waveform goes below zero.

The power dissipation of the minimum size basic logic gates due to equivalent lump capacitance (P_{cap}) were simulated and compared with P_{act} of basic logic gates under different loading conditions and shown in Table 4.5 and Figure 4.14. As illustrated in this table and figure, power dissipation of minimum size basic logic gates are in the range of μW and increase with increasing of fanout. The basic logic gates P_{act} have a second-order relation with fanout but P_{cap} have a first-order relation with fanout. The value of P_{act} and P_{cap} are close at low fanout while differences between those increase with increasing the fanout. The P_{cap} is lower than P_{act} at low fanout and is higher than P_{act} at high fanout because P_{act} has a first order relation with fanout while P_{act} has a second order relation with fanout.

	Inverter			NAND2			NOR2			NAND3			NOR3		
fanout/xCin	P_{cap}	P_{act}	$E\%$	P_{cap}	P_{act}	$E\%$	P_{cap}	P_{act}	$E\%$	P_{cap}	P_{act}	$E\%$	P_{cap}	P_{act}	$E\%$
0	0.98	0.98	-0.01	1.58	1.46	-8.22	1.66	1.46	-13.70	2.00	1.80	-11.11	2.20	1.94	-13.40
1	2.07	2.07	-0.01	3.28	3.08	-6.49	3.60	3.40	-5.88	4.15	4.10	-1.22	4.82	4.75	-1.47
2	3.33	3.30	-0.91	4.90	4.88	-0.41	5.56	5.50	-1.13	6.58	6.50	-1.30	7.78	7.70	-1.07
3	4.57	4.64	1.51	6.78	6.80	0.29	7.76	7.72	-0.47	9.12	9.00	-1.30	10.7	10.9	1.74
4	5.87	6.09	3.61	8.59	8.86	3.05	9.89	1.01	2.20	11.5	11.6	0.45	13.7	14.3	3.94
5	7.12	7.62	6.56	10.4	11.0	5.64	12.1	12.6	4.36	14.0	14.4	2.92	16.6	17.9	7.12
6	8.40	9.20	8.70	12.1	13.3	9.02	14.2	15.3	7.10	16.5	17.3	4.56	19.6	21.8	10.17
7	9.70	10.9	10.60	13.9	15.7	11.46	16.4	18.2	10.2	18.9	20.5	7.59	22.5	26.1	13.79
8	11.0	12.6	12.66	15.8	18.3	13.66	18.5	21.3	12.96	21.5	23.9	10.14	25.4	30.9	17.81
9	12.3	14.5	14.94	17.7	21.0	15.71	20.6	24.7	16.60	23.9	27.6	13.38	28.4	35.9	21.01
10	13.6	16.5	17.58	19.5	23.8	18.07	22.8	28.4	19.85	26.3	31.4	16.12	31.2	41.3	24.43
	Maximum $E\%$		32.3												
	Average $E\%$		9.2												

Table 4.5: Comparison of basic logic gates power dissipation due to actual gate (P_{act}) and equivalent lump capacitance (P_{cap}) under different load conditions (power values are in μW and $E\% = \frac{P_{act}-P_{cap}}{P_{act}} \times 100$).



(a) inverter

(b) 2-input NAND

(c) 2-input NOR

(d) 3-input NAND

(e) 3-input NOR

Figure 4.14: Comparison of basic logic gates power dissipation due to actual gate (P_{act}) and equivalent lump capacitance (P_{cap}) under different load conditions .

4.3 Dynamic Power Dissipation and Transistor Width

The simulation was repeated while the transistor width of the basic logic gates under the test were increased 2, 3, 5, 10, 15 and 30 times the minimum width. Table 4.6 to 4.10 and Figure 4.15 to 4.19 show comparison of proposed model (P_{model}) and basic logic gates P_{act} with different transistor width under different loading condition. The highest error between basic logic gates P_{act} and P_{model} happen at fanout-of-zero. The reason of high discrepancy at fanout-of-zero is that the output of basic logic gates is floating and it makes estimating of $C_{gate,par}$ difficult. As illustrated in these figures, P_{model} have a good agreement with P_{act} of basic logic gates for all different transistor sizes.

	$Z=1$			$Z=2$			$Z=3$			$Z=5$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	0.98	0.95	3.20	0.21	0.2	5.14	0.32	0.3	6.53	0.55	0.52	6.12
1	0.21	0.21	-1.20	0.41	0.41	-1.25	0.62	0.62	-1.28	10.36	10.51	-1.45
2	3.30	3.33	-0.94	6.40	6.47	-1.14	9.57	9.75	-1.86	16.02	16.32	-1.87
3	4.64	4.66	-0.38	8.90	9.01	-1.25	13.40	13.60	-1.12	22.44	22.63	-0.84
4	6.09	6.07	0.26	11.76	11.70	0.22	17.60	17.60	-0.19	29.40	29.44	-0.13
5	7.62	7.58	0.51	14.70	14.60	0.38	22.00	22.00	0.01	36.86	36.75	0.30
6	9.200	9.18	0.24	17.80	17.70	0.34	27.00	26.60	1.30	44.83	44.56	0.60
7	10.85	10.90	-0.14	21.07	21.00	0.24	31.60	31.60	0.07	53.05	52.87	0.34
8	12.64	12.60	-0.02	24.48	24.50	-0.03	36.70	36.80	-0.25	61.54	61.68	-0.23
9	14.46	14.50	-0.35	28.09	28.10	-0.17	42.30	42.30	-0.03	71.02	70.99	0.04
10	16.50	16.50	0.19	31.97	32.00	-0.02	48.00	48.10	-0.13	80.69	80.80	-0.14

	$Z=10$			$Z=15$			$Z=30$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	11.50	10.71	6.87	16.30	17.40	6.58	35.30	32.86	6.90
1	20.90	21.26	-1.73	32.10	31.70	-1.38	63.70	64.50	-1.25
2	32.20	32.85	-2.03	49.60	48.40	-2.44	97.00	99.33	-2.40
3	45.00	45.48	-1.06	68.60	68.30	-0.42	136.00	137.40	-0.99
4	59.10	59.14	-0.07	89.10	88.80	-0.39	178.00	178.60	-0.32
5	74.10	73.84	0.35	111.00	112.00	0.20	223.60	223.00	0.27
6	90.00	89.58	0.47	135.00	136.00	0.47	272.40	270.60	0.65
7	106.80	106.40	0.42	160.00	161.00	0.49	323.00	321.40	0.48
8	124.10	124.20	-0.05	187.00	187.00	-0.01	375.50	375.50	0.01
9	143.00	143.00	-0.01	215.00	215.00	0.01	432.20	432.70	-0.11
10	162.60	162.90	-0.19	245.00	245.00	-0.22	492.20	493.10	-0.18
	Maximum $E\%$		6.9						
	Average $E\%$		1.1						

Table 4.6: Comparison of proposed model (P_{model}) and inverter P_{act} with respect to transistor width under different load conditions (power values are in μW).

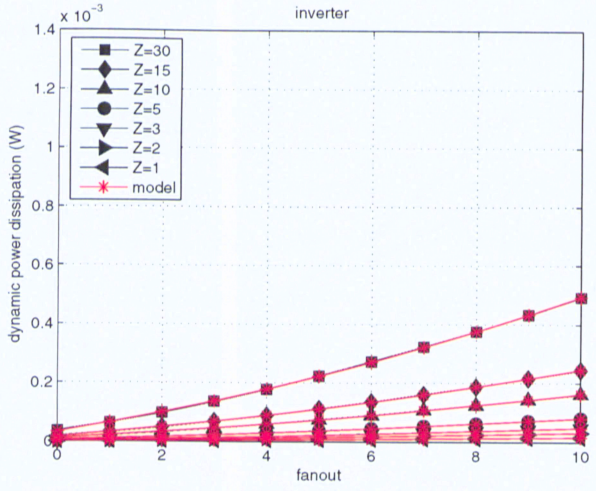


Figure 4.15: Comparison of proposed model (P_{model}) and inverter P_{act} with respect to transistor width under different load conditions.

	$Z=1$			$Z=2$			$Z=3$			$Z=5$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	1.46	1.45	0.6	2.96	3.03	2.3	4.53	4.52	0.3	7.70	7.32	4.9
1	3.08	3.10	0.7	6.07	6.17	1.6	9.16	9.20	0.5	1.53	1.53	0.4
2	4.88	4.88	0.0	9.80	9.58	2.2	14.30	14.30	0.0	23.83	24.01	0.8
3	6.800	6.79	0.1	13.30	13.30	0.2	19.80	19.80	0.1	33.05	33.34	0.9
4	8.860	8.83	0.3	17.30	17.20	0.4	25.80	25.70	0.2	43.09	43.34	0.6
5	11.00	11.00	0.0	21.40	21.50	0.3	32.15	32.10	0.2	53.72	54.00	0.5
6	13.30	13.30	0.0	26.00	26.00	0.1	38.80	38.80	0.1	64.80	65.32	0.8
7	15.70	15.70	0.2	30.60	30.80	0.5	46.00	46.00	0.0	79.62	77.30	2.9
8	18.30	18.30	0.0	35.60	35.80	0.6	53.40	53.60	0.3	89.26	89.94	0.8
9	21.00	21.00	0.1	41.40	41.10	0.7	61.60	61.60	0.1	102.9	103.2	0.3
10	23.80	23.80	0.0	46.70	46.70	0.1	70.00	70.00	0.1	117.1	117.2	0.1
	$Z=10$			$Z=15$			$Z=30$					
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$			
0	14.92	14.72	1.4	23.30	23.70	1.8	47.80	47.300	1.1			
1	29.18	29.48	1.0	46.60	46.20	0.9	92.70	93.680	1.1			
2	45.44	45.60	0.4	72.10	71.80	0.4	144.00	144.40	0.3			
3	63.12	63.07	0.1	99.70	99.30	0.4	200.00	199.50	0.2			
4	82.04	81.89	0.2	129.00	130.00	0.2	259.50	259.00	0.2			
5	102.30	102.10	0.2	161.00	162.00	0.1	323.40	322.80	0.2			
6	123.80	123.60	0.2	195.00	195.00	0.1	391.20	391.00	0.1			
7	146.50	146.50	0.0	231.00	232.00	0.3	463.40	463.50	0.0			
8	170.30	170.70	0.2	270.00	270.00	0.1	539.10	540.40	0.2			
9	195.90	196.30	0.2	310.00	309.00	0.5	621.10	621.70	0.1			
10	223.60	223.20	0.2	353.00	354.00	0.2	708.30	707.30	0.1			
	Maximum $E\%$		4.9									
	Average $E\%$		0.5									

Table 4.7: Comparison of proposed model (P_{model}) and 2-input nand P_{act} with respect to transistor width under different load conditions (power values are in μW).

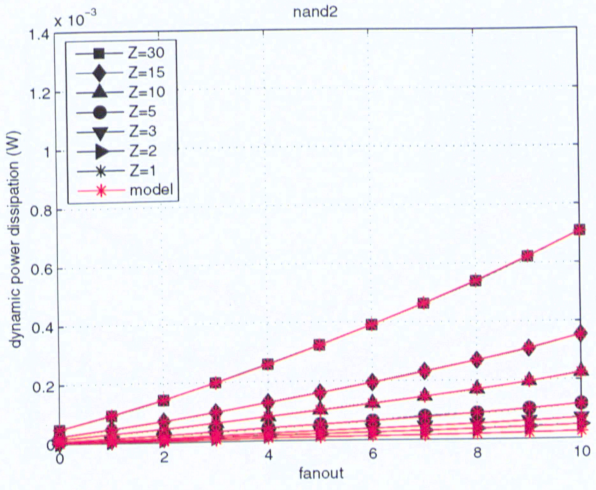


Figure 4.16: Comparison of proposed model (P_{model}) and 2-input nand P_{act} with respect to transistor width under different load conditions.

	$Z=1$			$Z=2$			$Z=3$			$Z=5$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	1.46	1.57	7.4	3.01	3.14	4.5	4.60	4.73	2.7	7.91	8.12	2.6
1	3.40	3.40	0.1	6.67	6.69	0.4	10.00	10.10	0.7	16.75	16.89	0.8
2	5.50	5.42	1.4	10.70	10.6	0.7	16.10	16.00	0.6	26.77	26.67	0.4
3	7.72	7.63	1.2	15.10	14.90	1.0	22.60	22.50	0.4	37.86	37.46	1.1
4	10.11	10.00	0.9	19.80	19.70	0.7	29.80	29.60	0.7	49.64	49.26	0.8
5	12.60	12.60	0.1	24.80	24.70	0.2	37.50	37.20	0.7	62.13	62.08	0.1
6	15.30	15.40	0.3	30.12	30.20	0.3	45.20	45.50	0.6	75.66	75.91	0.3
7	18.20	18.30	0.5	36.00	36.10	0.2	54.00	54.30	0.5	90.36	90.75	0.4
8	21.30	21.40	0.6	42.00	42.30	0.7	63.50	63.70	0.3	10.62	10.66	0.4
9	24.70	24.70	0.2	48.90	48.90	0.1	73.60	73.60	0.0	123.6	123.5	0.1
10	28.40	28.20	0.5	56.20	56.00	0.4	84.40	84.20	0.3	141.6	141.3	0.2

	$Z=10$			$Z=15$			$Z=30$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	15.74	16.03	1.8	24.70	24.30	1.5	49.03	49.830	1.6
1	33.03	33.26	0.7	51.00	50.50	0.9	101.30	102.10	0.8
2	52.63	52.48	0.3	80.00	80.60	0.2	161.60	160.70	0.5
3	74.19	73.70	0.7	113.00	114.00	1.0	226.70	225.60	0.5
4	97.51	96.90	0.6	148.00	149.00	0.4	298.00	296.90	0.4
5	122.30	122.1	0.2	187.00	187.00	0.1	375.00	374.40	0.2
6	149.00	149.3	0.2	229.00	228.00	0.3	457.50	458.30	0.2
7	177.90	178.5	0.3	274.00	273.00	0.3	549.60	548.40	0.2
8	208.90	209.6	0.3	321.00	321.00	0.1	641.10	644.90	0.6
9	242.80	242.8	0.0	372.00	372.00	0.1	745.90	747.70	0.2
10	278.50	277.9	0.2	426.00	427.00	0.2	860.00	856.80	0.4
	Maximum $E\%$		7.4						
	Average $E\%$		0.7						

Table 4.8: Comparison of proposed model (P_{model}) and 2-input nor P_{act} with respect to transistor width under different load conditions (power values are in μW).

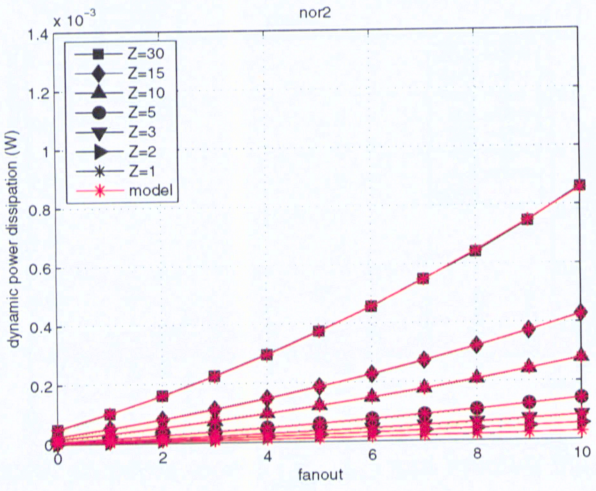


Figure 4.17: Comparison of proposed model (P_{model}) and 2-input nor P_{act} with respect to transistor width under different load conditions.

	$Z=1$			$Z=2$			$Z=3$			$Z=5$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	1.80	1.95	8.1	3.75	3.87	3.1	5.70	5.83	2.3	9.61	9.85	2.5
1	4.10	4.08	0.6	8.00	8.02	0.2	12.02	12.00	0.1	20.05	20.09	0.2
2	6.50	6.39	1.8	12.60	12.50	0.7	18.80	18.80	0.3	31.44	31.24	0.6
3	9.00	8.87	1.4	17.50	17.40	0.8	26.20	26.00	0.7	43.54	43.29	0.6
4	11.60	11.50	0.5	22.70	22.60	0.6	33.90	33.8	0.3	56.42	56.24	0.3
5	14.40	14.40	0.1	28.10	28.10	0.1	42.10	42.10	0.0	70.26	70.09	0.2
6	17.30	17.40	0.7	33.90	34.00	0.4	50.80	50.90	0.3	84.61	84.84	0.3
7	20.50	20.60	0.6	40.20	40.30	0.2	60.20	60.30	0.1	100.0	100.5	0.5
8	23.90	24.00	0.4	46.80	46.90	0.2	70.00	70.20	0.3	117.0	117.0	0.0
9	27.60	27.6	0.1	53.80	53.80	0.1	80.60	80.6	0.0	13.45	13.45	0.0
10	31.4	31.3	0.3	61.30	61.20	0.2	91.70	91.50	0.2	15.31	15.29	0.2
	$Z=10$			$Z=15$			$Z=30$					
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$			
0	19.11	19.55	2.3	29.90	29.40	1.7	59.00	60.44	2.4			
1	39.40	39.55	0.4	60.50	60.40	0.2	121.40	121.50	0.1			
2	61.74	61.36	0.6	93.90	94.00	0.1	188.60	188.10	0.2			
3	85.55	84.97	0.7	130.00	131.00	0.6	262.40	260.40	0.8			
4	110.8	110.4	0.4	169.00	170.00	0.4	340.00	338.30	0.5			
5	137.7	137.6	0.1	211.00	211.00	0.1	422.60	421.80	0.2			
6	166.1	166.6	0.3	255.00	255.00	0.2	509.00	510.90	0.4			
7	197.0	197.4	0.2	302.00	302.00	0.3	603.70	605.70	0.3			
8	229.5	230.1	0.2	352.00	352.00	0.2	704.10	706.00	0.3			
9	264.6	264.5	0.0	405.00	405.00	0.0	813.30	812.00	0.2			
10	301.3	300.7	0.2	461.00	461.00	0.2	924.80	923.60	0.1			
	Maximum $E\%$		8.1									
	Average $E\%$		0.6									

Table 4.9: Comparison of proposed model (P_{model}) and 3-input nand P_{act} with respect to transistor width under different load conditions (power values are in μW).

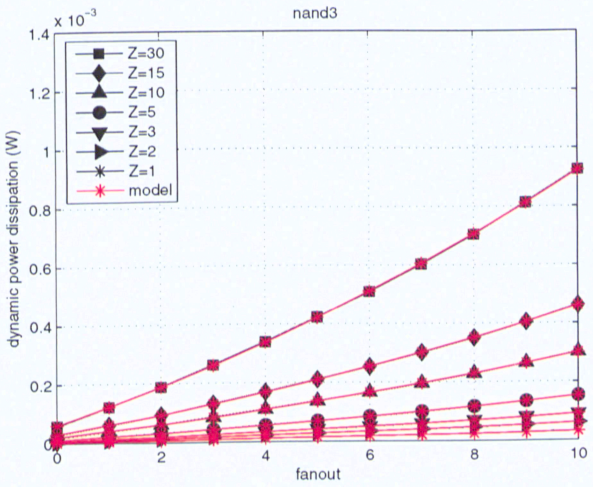


Figure 4.18: Comparison of proposed model (P_{model}) and 3-input nand P_{act} with respect to transistor width under different load conditions.

	$Z=1$			$Z=2$			$Z=3$			$Z=5$		
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$
0	1.94	2.15	10.9	3.95	4.38	10.8	6.03	6.61	9.6	1.02	1.10	8.3
1	4.75	4.70	1.0	9.33	9.29	0.5	14.00	13.90	0.4	23.31	23.27	0.2
2	7.70	7.56	1.9	15.20	14.80	2.4	22.70	22.30	2.0	37.85	37.15	1.9
3	10.90	10.70	1.8	21.30	21.0	1.3	32.10	31.5	1.7	53.51	52.66	1.6
4	14.26	14.2	0.7	28.10	27.80	0.9	42.10	41.80	0.7	70.14	69.81	0.5
5	17.90	17.90	0.0	35.30	35.30	0.0	53.00	53.10	0.1	88.50	88.59	0.1
6	21.80	22.00	0.7	43.20	43.40	0.5	65.00	65.30	0.4	108.7	109.0	0.3
7	26.10	26.3	0.8	51.80	52.10	0.6	77090	78.50	0.7	130.2	131.1	0.7
8	30.90	31.00	0.2	61.10	61.50	0.7	92.20	92.70	0.5	154.4	154.7	0.2
9	35.90	35.90	0.0	71.50	71.50	0.0	108.0	108.0	0.2	179.6	180.1	0.3
10	41.30	41.10	0.4	82.60	82.20	0.5	124.4	124.0	0.4	208.0	207.0	0.5
	$Z=10$			$Z=15$			$Z=30$					
fanout	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$	P_{act}	P_{model}	$E\%$			
0	20.19	21.64	7.2	33.40	31.10	7.6	62.40	67.68	8.5			
1	45.67	45.64	0.1	69.90	70.10	0.2	140.2	140.00	0.2			
2	74.09	72.88	1.6	112.00	114.00	1.8	226.5	222.60	1.7			
3	104.90	103.4	1.4	158.00	161.00	1.6	320.30	315.50	1.5			
4	137.80	137.1	0.5	210.00	211.00	0.4	421.3	418.70	0.6			
5	173.90	174.10	0.1	266.00	265.00	0.4	534.10	532.20	0.4			
6	213.40	214.30	0.4	328.00	327.00	0.3	653.60	656.00	0.4			
7	256.30	257.70	0.6	395.00	393.00	0.3	784.40	790.10	0.7			
8	304.00	304.40	0.1	466.00	465.00	0.4	928.10	934.50	0.7			
9	353.60	354.30	0.2	543.00	543.00	0.0	109.00	108.90	0.1			
10	409.20	407.50	0.4	625.00	627.00	0.3	126.00	125.40	0.5			
	Maximum $E\%$		10.9									
	Average $E\%$		1.4									

Table 4.10: Comparison of proposed model (P_{model}) and 3-input nor P_{act} with respect to transistor width under different load conditions (power values are in μW).

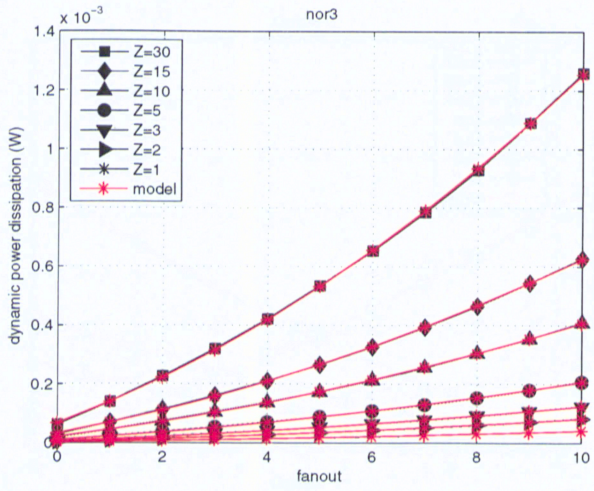


Figure 4.19: Comparison of proposed model (P_{model}) and 3-input nor P_{act} with respect to transistor width under different load conditions.

Chapter 5

Conclusions and Future Works

With the growing use of portable and wireless electronic systems, reduction in power dissipation has become an important topic in today's VLSI circuit and system designs. In order to investigate the power dissipation, all the components which are contributing in the total power dissipation must be considered. In this thesis project, the source of dynamic power dissipation is studied. Previous works on modeling dynamic power dissipation components including switching, overshoot and short circuit power dissipation are reviewed. A model is proposed to estimate overshoot power dissipation P_{ov} which has a second-second degree relation with fanout. Also the existing model of short circuit power dissipation that is based on equivalent short circuit capacitance was modified in order to increase the accuracy of the model. Therefore, the existing dynamic power dissipation proposed model is modified in order to increase its accuracy. Using these two proposed model a model was proposed to accurately estimate dynamic power dissipation. The proposed model of dynamic power dissipation gives design engineers an accurate model of dynamic power dissipation. Therefore, design engineers do not have to refer to Cadence to estimate dynamic power dissipation. This model which is based on input capacitance ($C_{gate,in}$), parasitic capacitance ($C_{gate,par}$) and input-to-output coupling capacitance (C_M) of basic logic gate has a second degree relation with fanout. The value of $C_{gate,in}$, $C_{gate,par}$ and C_M are extracted from minimum size basic logic gates simulation results. Moreover dynamic power dissipation result of actual gate and lumped capacitance are not same and it is not practical to replace actual gate

with lumped capacitance. Also, the proposed model takes into account transistor width. Using 0.13μm CMOS technology and netlist (Cadence) the proposed model is validated by comparing with dynamic power dissipation simulation results. The model exhibits good accuracy when compared with Cadence created simulation results of basic logic gates loaded with identical basic logic gates. The Proposed model of inverter and Cadence simulated power dissipation have a maximum error of 3.20% and an average error of 0.13%. Model was extended for nand and nor gates as well and it exhibited maximum error of 3.20% and an average error of -0.28%.

Since the dynamic power dissipation is proportional to the basic logic gates capacitances, it is possible to reduce the dynamic power dissipation by decreasing basic logic gates transistors size. However, decreasing the size of the transistors also decreases their current drive, and thus makes the gates slower. This trade off between power and delay can be used to size transistors in a way to optimize the dynamic power dissipation and delay in future works.

Bibliography

[1] T. Kuroda, "Low Power, High Speed CMOS VLSI Design," *in Proceedings of IEEE VLSI on Computers and Processors*, pp. 310-315, 2002.

[2] S. Turgis, D. Auvergne, "A Novel Macromodel for Power Estimation in CMOS Structures," *in Proceedings of IEEE on Computer-Aided Design of Integrated Circuits and Systems*, vol. 17, pp. 1090-1098, 1998.

[3] N.H.E. Weste, D. Harris, *CMOS and VLSI Design, A Circuits and Systems Perspective.* 3rd ed., Pearson Addison Wesley Publication, pp. 190-196, 2005.

[4] A.P. Chandrakasan, R.W. Brodersen, "Minimizing Power Consumption in Digital CMOS Circuits," *in Proceedings of IEEE*, vol. 83, pp. 498-523, 1995.

[5] S.M.S. Kang, "Elements of low power design for integrated systems," *in Proceedings of International Symposium on Low Power Electronics and Design*, pp. 205-210, 2003.

[6] N.S. Kim, T. Austin, D. Blaauw, T. Mudge, K. Flautner, J.S. Hu, M.J. Irwin, M. Kandemir, V. Narayanan, "Leakage Current: Moore's Law Meets Static Power," *in Proceedings of IEEE on Computer Society*, vol. 36, pp. 68-75, 2003.

[7] C.V. Schimpfle, S. Simon, J.A. Nossek, "High-Level Circuit Modeling for Power Estimation," *in Proceedings of International Conference on Electronics, Circuits and Systems*, vol. 2, pp. 807-810, 1999.

[8] R.X. Gu, M.I. Elmasry, "Power dissipation analysis and optimization of deep submicron CMOS digital circuits," *IEEE Journal of Solid-State Circuits*, vol. 31, pp.707-713, 1996.

[9] S. Sharifi, J. Jaffari, M. Hosseinabady, A. Afzali-Kusha, Z. Navabi, "Simultaneous reduction of dynamic and static power in scan structures," *in proceedings of the Design, Automation and Test in Europe Conference and Exhibition*, vol. 2, pp.190-196, 2005.

[10] B. Labouygues, J. Schindler, S. Engels, P. Maurine, N. Azemard, D. Auvergne, "Continuous Representation of the Performance of a CMOS Library," *in Proceedings of the European Solid-State Circuits Conference*, pp. 595-598, 2003.

[11] I. Sutherland, B. Sproull, D. Harris, *Logical Effort: Designing Fast CMOS Circuits.* Morgan Kaufmann Publication, 1999.

[12] Z. Huang, A. Kurokawa, Y. Inoue, "Modeling the Overshooting Effect in the Submicron CMOS Inverters," *in Proceedings of International conference on Communications, Circuits and Systems*, vol. 2, pp. 1191-1195, 2005.

[13] J.L. Rossello, J. Segura, "An analytical charge-based compact delay model for submicrometer CMOS inverters," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 51, pp. 1301-1311, 2004.

[14] H.J.M. Veendrick, "Short-Circuit Power Dissipation of Static CMOS Circuitry and Its Impact on the Design of Buffer Circuits," *in Proceeding of IEEE J. Solid-state Circuits*, vol. SC-19, pp. 468-473, 1984.

[15] E. Acar, R. Arunachalam, S.R. Nassif, "Predicting Short Circuit Power from Timing Models," *in Proceedings of Asia and South Pacific Design Automation Conference*, pp. 277-282, 2003.

[16] S. Ulman, "Macromodel for Short Circuit Power Dissipation of Submicron CMOS Inverters and Its Application to Design CMOS Buffers," *in Proceedings of International Symposium on Circuits and Systems*, vol. 5, pp. 269-272, 2003.

[17] U. Ko, P.T. Balsara, "Short-circuit power driven gate sizing technique for reducing power dissipation," *IEEE Transactions on VLSI Systems*, vol. 3, pp.450-455, 1995.