Macromodel for Short Circuit Power and Propagation Delay Estimation of CMOS Circuits

Seung-Ho Jung, Jong-Humn Baek, Seok-Yoon Kim
Department of Computing, Soongsil University
1-1 Sangdo 5 dong, Dongjak-Ku, Seoul, Korea
Phone: +82-2-813-0682 Fax: +82-2-822-9872
E-mail: ikaros@ic.soongsil.ac.kr, jhbaek@ic.soongsil.ac.kr, ksy@comp.ssu.soongsil.ac.kr

Abstract

This paper presents a simple method to estimate short-circuit power dissipation and propagation delay for static CMOS logic circuits. Short-circuit current expression is derived by accurately interpolating peak points of actual current curves which is influenced by the gate-to-drain coupling capacitance. The macromodel and its expressions estimating the delay of CMOS circuits, which is based on the current modeling expression, are also proposed after investigating the voltage waveforms at transistor output modes. It is shown through simulations that the proposed technique yields better accuracy than previous methods when signal transition time and/or load capacitance decreases, which is a characteristic of the present technological evolution.

1. Introduction

As clock speeds increases and function block sizes scale down in today's technology trend, the estimation for power and gate delay need to be fine-tuned. In circuit design flows, the accurate estimation of the power dissipation and gate delay of the CMOS gate has become on essential factor for reliability and performance of the chips.

In general, power dissipation in static CMOS circuits are composed of dynamic power dissipation, static power dissipation and short-circuit power dissipation. During the output transition in a static CMOS structure, a direct path from power supply to ground is created, resulting in a short circuit power dissipation that can be more than 20 percentage of the total power dissipation since the number of transitions increases due to clock speed up. Also, the short-circuit current has an effect on the evaluation of propagation delay of CMOS gate. Therefore, estimating the short-circuit current is important to design high-speed VLSI's.

Much effort has been devoted for the extraction of accurate, analytical expressions for timing models of basic circuits. However, previous works in the short circuit power estimation area have not entirely or exactly reflected gate-to-drain coupling capacitance and the short channel effect [2][3][4][5]. For example, a recent work [6] overlooked the gate-to-drain capacitance effect. And required a lot of computational iteration due to the introduction of power series. Thus, we present a macromodel based on previous work, which enables detailed analysis of the transient behavior of CMOS inverters. We model the short-circuit current by an appropriate linear or nonlinear function, in order to achieve better accuracy and to avoid an overestimation of short-circuit power dissipation, with taking the influence of the gate-to-drain coupling capacitance into account. Also, we derive propagation delay equations from short-circuit current equations.

2. Analysis of Short Circuit Current in CMOS inverter

Figure1 illustrate a rising ramp input(eq.(1)) and CMOS inverter model:

$$V_{in} = \frac{V_{DD}}{t_T} t, \quad 0 \le t \le t_T \tag{1}$$

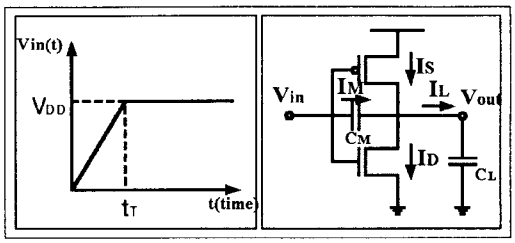


Fig. 1. The ramp input and the CMOS inverter

In Fig. 1, the output load consists of the inverter drain junction capacitance, the gate capacitance of fan-out gates and the interconnect capacitance. The gate-to-drain coupling capacitance C_M is

$$C_M = C_{gdo} + C_g \tag{2}$$

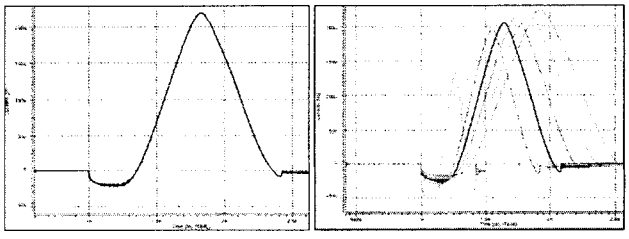
and

$$C_g = C_{gs} + C_{gd} \tag{3}$$

where C_{gdo} is gate-to-drain overlap capacitance, C_g is gate-to-channel capacitance, C_{gs} is gate-to-source channel capacitance, C_{gd} is gate-to-drain channel capacitance and C_{ox} is gate-to-oxide capacitance except overlap.

$$C_{gd} = \frac{1}{2} C_{ox} WL \tag{4}$$

Figure 2 shows the influence of gate-drain coupling capacitance in short-circuit current waveforms. Note that as transition time decreases the negative portion in Fig. 2 is increasing compared with the positive portion. We model the short-circuit current based on Fig. 2 in the following section.



(a) (b)

Fig. 2. HSPICE simulation result of short-circuit current waveforms (a) case (t_T=1.7ns, C_L=30fF), (b) varying input

transition time (C_L=30fF, t_T=0.5ns~1.9ns)

3. Modeling the Short Circuit Current

As shown in Fig. 3, we model the short-circuit current waveform from two choices: the piecewise nonlinear modeling and the piecewise linear modeling.

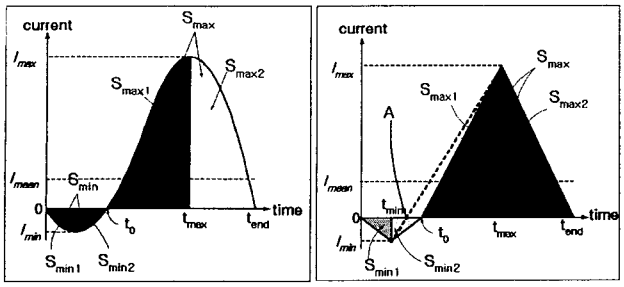


Fig. 3. Current Modeling Choices (a) Piecewise nonlinear (b) Piecewise linear

The short-circuit power dissipation caused by a short circuit current I_S during the rising transition of input signal can be estimated as :

$$P_{SC} = I_{mean} * V_{DD} \quad (5)$$

I_{mean} is the average current and is given by

$$I_{mean} = 1 / t_r(S_{min}+S_{max}) = S_{total}/t_r \quad (6)$$

where S_{max} and S_{min} are marked in Fig. 3.

In order to obtain S_{min} and S_{max} values, we model short-circuit current, I_S according to three methods as follows :

• Method A : Current waveform I_S is modeled by

$$I_s(t) = \alpha_1 t \cdot \beta_1^{-\gamma_1 t} \quad (7)$$

where α_1 , β_1 , γ_1 can be obtained by using the peak points, and are given by

$$\alpha_1 = \frac{I_{min}\beta_1^{\frac{1}{\log \beta_1}}}{t_{min}}, \quad \beta_1 > 0 \ (\beta \neq 1), \quad \gamma_1 = \frac{1}{t_{min}\log \beta_1}$$

Then S_{total} is

$$S_{total} = (e\text{-}2)\{t_0 I_{min} + (t_{end}\text{-}t_0) I_{max}\} \quad (8)$$

• Method B : I_S is modeled by

$$I_s(t) = \alpha_2 t \cdot \beta_2^{\frac{t^2}{\gamma_2}} \quad (9)$$

where α_2 , β_2 , γ_2 can be obtained using the similar works in Method A, and are given by

$$\alpha_2 = \frac{I_{min}\beta_2^{\frac{1}{2\log \beta_2}}}{t_{min}}, \quad \beta_2 > 0 \ (\beta \neq 1), \quad \gamma_2 = \frac{1}{t_{min}\log \beta_2}$$

Then S_{total} is

$$S_{total} = (\sqrt{e}\text{-}1)\{t_0 I_{min} + (t_{end}\text{-}t_0) I_{max}\} \quad (10)$$

• Method C : I_S is modeled by

$$I_s(t) = \alpha_3 t \quad (11)$$

where α_3 is

$$\alpha_3 = \frac{I_{min}}{t_{min}}$$

Then S_{total} is given by using triangular area formula :

$$S_{total} = 0.5\{t_0 I_{min} + (t_{end}\text{-}t_0) I_{max}\} \quad (12)$$

Since we use four linear equations, Method C is more accurate than [4]

Consequently, for the accurate estimation of short-circuit power, unknown values t_{min}, I_{min}, t_{max}, I_{max} have to be calculated as exactly as possible. In this derivation process, we use the simplified bulk-charge MOS model[8] for the operational region of the MOS devices.

3. 1 Derivation of I_{min}

To obtain I_{min}, we should investigate the behavior of MOS transistors. At this time point, the NMOS transistor is off, and the PMOS transistor is in the linear region. Therefore, t_{min} is equal to t_{thn} and I_{min} is

$$I_{min} = \beta_p\left[\left(V_{DD} - \frac{V_{DD}}{t_T}t_{thn} - |V_{thp}|\right)(V_{DD} - V_{out}(t_{thn})) - \frac{1+\delta_p}{2}(V_{DD} - V_{out}(t_{thn}))^2\right] \quad (13)$$

Also, eq. (14) is derived from the application of the Kirchoff's current law at the output node

$$I_s = \frac{dV_{out}}{dt}(C_L + C_M) - \frac{V_{DD}}{t_T}C_M \quad (14)$$

Substituting equations (7), (9), (11) for eq. (14), $V_{out}(t_{thn})$ is expressed as linear function of I_{min} :

$$V_{out}(t_{min}) = V_{DD} + \frac{C_M V_{DD} t_{thn}}{(C_L + C_M)t_T} + \frac{S_{min1}}{(C_L + C_M)} \quad (15)$$

where $V_{out}(t_{thn})$ is the output voltage when the input voltage reaches V_{thn}. Solving eq.(13) and eq.(15), we can obtain I_{min} as follows:

$$I_{min} = \frac{-a_2 - \sqrt{a_2^2 - 4a_1a_3}}{2a_1} \quad (16)$$

Each parameter in methods A to C is specified in Appendix A and we assume that the value of inner root is always positive.

3. 2 Derivation of I_{max}

To calculate I_{max}, we consider two cases as follow (Case A and Case B). We divided two cases using criteria defined in [6].

• Case A : In this case, PMOS transistor is in linear region at t_{max} , NMOS transistor is in saturation region. Thus, I_{max} is given by

$$I_{max} = \beta_p\left[\left(V_{DD} - \frac{V_{DD}}{t_T}t_{thn} - |V_{thp}|\right)(V_{DD} - V_{out}(t_{max})) - \frac{1+\delta_p}{2}(V_{DD} - V_{out}(t_{max}))^2\right] \quad (17)$$

To calculate t_0, t_{max} in eq. (17) can be obtained by using the fact that short circuit current is almost symmetric in this case. By applying KCL at ouput node, $V_{out}(t_{thn})$ is derived by linear a equation as follows.

$$V_{out}(t_{max}) = V_{DD} - \frac{\beta_n t_T (V_{DD}t_{max} - V_{thn})^3}{6V_{DD}(C_M + C_L)(1+\delta_n)}\left(\frac{V_{DD}t_{max}}{t_T}\right)^3 + \frac{C_M V_{DD} t_{max}}{(C_L + C_M)t_T} + \frac{S_{min} + S_{max1}}{(C_L + C_M)} \quad (18)$$

Substituting eq. (18) for eq. (17), I_{max} is obtained as solutions of 2nd-order polynomial as follows:

$$I_{max} = \frac{-b_1 - \sqrt{b_2^2 - 4b_1b_3}}{2b_1}, \quad b_2^2 - 4b_1b_3 > 0 \quad (19)$$

and parameters described in eq. (19) are specified in Appendix B

• Case B : In this case, both PMOS and NMOS transistors are in saturation region at t_{max}. By solving the differential equation obtained at output node, $V_{out}(t)$ and I_{max} are represented by

$$V_{out}(t) = V_{DD} - \frac{\beta_n t_T (V_{in}(t) - V_{thn})^3}{6V_{DD}(C_M + C_L)(1+\delta_n)} + \frac{(I_{max} - I_{min})(t - t_{thn})^2}{2(t_{max} - t_{min})(C_M + C_L)} + \frac{I_{min}(t - t_{min})}{C_M + C_L} + \frac{C_M V_{in}(t)}{C_M + C_L} + \frac{S_{min1}}{C_M + C_L} \quad (20)$$

$$I_{max} = \frac{\beta_p}{2(1+\delta_p)}(V_{DD}-V_{in}(t_{max})-|V_{thp}|)^2 \quad (21)$$

We can obtain eq. (22), because t_{max} is time at logic threshold voltage [14].

$$V_{out}(t_{max}) = V_{in}(t_{max}) = V_{INV} \quad (22)$$

Solving eq. (20), (21) and (22), we get eq. (23) :

$$c_1 V_{INV}^3 + c_2 V_{INV}^2 + c_3 V_{INV} + c_4 = 0 \quad (23)$$

and coefficients are specified in Appendix C. Thus, solving 3rd-order polynomial eq. (23), we can derive I_{max}.

4. Propagation Delay Estimation

The propagation delay of CMOS gates is the difference between time point at 50% of final values of the input waveform and output waveform. The time $t_{50\%}$, at 50% of final value of the output waveform can be derived from using the output waveform equations derived in the previous section. The time point, $t_{50\%}$ may exist in five regions of Fig.4. In order to give a complete analysis, three cases of input ramp are considered and we divided three cases using criteria ($k=t_T\beta V_{DD}/C_L$) defined in [6]:

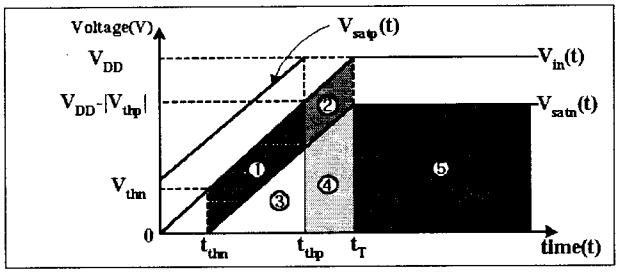


Fig. 4. Region diagram that $t_{50\%}$ may fall into

• Case A ($k \le 5$) : In the case of fast input ramps, $t_{50\%}$ value occurs in region ⑤. In this case, PMOS is off and NMOS is in linear region. Thus, The $t_{50\%}$ value is expressed as

$$V_{out}(t) = \frac{2(V_{DD}-V_{thn})}{(1+\delta_n)[\exp\left\{\frac{(\beta_n(V_{thn}-V_{DD})(t_{satn}-t)}{C_L+C_M}\right\}+1} \quad (24)$$

where t_{satn} is the time point that output waveform crosses $V_{satn}(t)$ in Fig. 4 and using the boundary condition, t_{satn} is

$$t_{satn} = \frac{2(1+\delta_n)(S_{total}+V_{DD}(C_L+2C_M))}{\beta_n(V_{DD}-V_{thn})^2} + \frac{2(C_M+C_L)}{\beta_n(V_{DD}-V_{thn})} + \frac{2t_T+t_n}{3} \quad (25)$$

Therefore, $t_{50\%}$ is

$$t_{50\%} = t_{satn} - \frac{C_L+C_M}{\beta_n(V_{thn}-V_{DD})}\ln\left[\frac{4(V_{DD}-V_{thn})}{V_{DD}(1+\delta_n)}-1\right] \quad (26)$$

• Case B ($5 < k \le 14$) : In the case of slow input ramps, $t_{50\%}$ value occurs in region ⑤, ($5 < k \le 8$), or in region ④ ($8 < k \le 14$). In the first case, the $t_{50\%}$ value is calculated by (parameters are extracted same as the Case A)

$$t_{50\%} = -\frac{1}{d_1}\ln\left[\frac{1}{d_2}\left(\frac{2}{V_{DD}}+\frac{d_3}{d_1}\right)\right] \quad (27)$$

In next case we used Newton-Raphson method and Bisection method [9] for the evaluation of t50%, since t50% cannot be expressed analytically

• Case C ($14 < k$) : In the case of very slow input ramps, $t_{50\%}$ value occurs in three possible regions. For $14 < k \le 18$ occurs in region ④ or in region ⑤. In the former case the $t_{50\%}$ value is given by (24). For $k > 18$, $t_{50\%}$ occurs in region ③ and $t_{50\%}$ is calculated by

$$V_{in}(t_{satp})^3 + e_1 V_{in}(t_{satp})^2 + e_2 V_{in}(t_{satp}) + e_3 = 0 \quad (28)$$

each coefficients are specified in Appendix. D. therefore, $t_{50\%}$ value becomes the solution of 3rd-order polynomial.

5. Experimental Results

We examine the accuracy of the short-circuit power and propagation delay expressions. Figure 5 shows the short-circuit power values calculated by method C presented in this paper, those calculated by the formulas proposed by Vemuru[2] and Veendrick[11], and those simulated using HSPICE. We compareour estimation results of the short-circuit power dissipation per one clock cycle with others results as a function of output load capacitance under various input transition times. It is seen that the error of out expressions is smaller than that of other works. Note that error ratio is increasing as scaling down output load capacitance. Average relative error ratio is 6.5% in our experiments. Figure 6 shows estimation results of the propagation are compared to those of SPICE, average relative error ration is about 6.2% in our experiments.

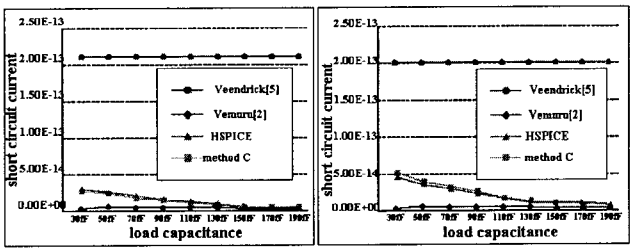


Figure. 5. Compared to simulation result of other work
(a) transition time : 0.9ns (b) transition time : 1.1ns

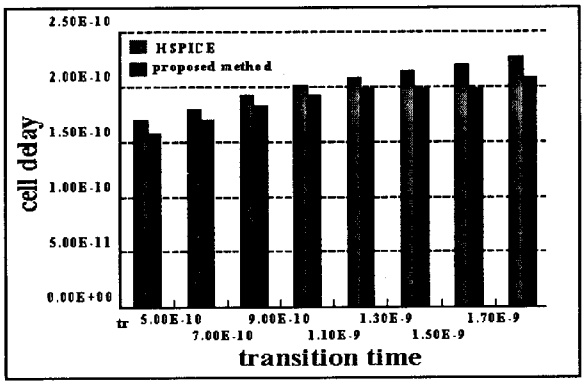


Figure 6. Compared to propagation delay of the SPICE

Simulation environment is L=0.5μm, W=3.5μm, $|V_{thp}|$=826.92mV, β_p=2.0mA/V², δ_p=0.24 in PMOS transistor and L=0.5μm, W=2 μm, V_{thn} =697.82mV, β_n=2.0mA/V², δ_n=0.326 in NMOS transistor.

6. Conclusion

In this paper, we proposed a analytical expression to estimate short circuit power dissipation and propagation delay in static CMOS inverter. In order to achieve the goal, analytical expressions at the output node of CMOS inverter have been derived for three cases of input ramps. In this process, we accurately taken account the influence

of the short-circuit current and the gate-to-drain coupling capacitance. Therefore, our method yield better accuracy and efficiency computation time than previous works. Especially, when transition time of a signal and load capacitance of a CMOS gates decrease, it is shown that the estimation results are in good agreement with SPICE simulations.

7. Appendix

Appendix. A

In Method A, parameters are

$$a_1 = \frac{(1+\delta_p)(e-2)^2 t_{thn}^{\ 2}}{2(C_L+C_M)},$$

$$a_2 = \frac{1}{\beta_p} + \frac{t_{thn}(e\text{-}2)}{(C_L+C_M)}\left\{V_{DD} - V_{thn} - |V_{thp}| + \frac{C_M V_{thn}(1+\delta_p)}{C_L+C_M}\right\}$$

$$a_3 = \frac{C_M V_{thn}}{C_L+C_M}\left\{V_{DD} - V_{thn} - |V_{thp}| + \frac{C_M V_{thn}(1+\delta)}{2(C_L+C_M)}\right\} \quad \text{(A - 1)}$$

In Method B, parameters are

$$a_1 = \frac{(1+\delta_p)(\sqrt{e}-2)^2 t_{thn}^{\ 2}}{2(C_L+C_M)},$$

$$a_2 = \frac{1}{\beta_p} + \frac{t_{thn}(\sqrt{e}\text{-}2)}{(C_L+C_M)}\left\{V_{DD} - V_{thn} - |V_{thp}| + \frac{C_M V_{thn}(1+\delta_p)}{C_L+C_M}\right\}$$

a_3 = same as a_3 in method A (A - 2)

In Method C, parameters are

$$a_1 = \frac{(1+\delta_p) t_{thn}^{\ 2}}{8(C_L+C_M)},$$

$$a_2 = \frac{1}{\beta_p} + \frac{t_{thn}}{2(C_L+C_M)}\left\{V_{DD} - V_{thn} - |V_{thp}| + \frac{C_M V_{thn}(1+\delta_p)}{C_L+C_M}\right\}$$

a_3 = same as a_3 in method A (A - 3)

Appendix. B

In Method C, parameters are

$$b_1 = \frac{t_1(1+\delta_p)^2}{4} \quad \text{(B - 1)}$$

$$b_2 = t_1 t_2(1+\delta_p) + \frac{1}{\beta_p} + t_1\left(V_{DD} - \frac{V_{DD} t_{max}}{t_T} - |V_{thp}|\right) \quad \text{(B - 2)}$$

$$b_3 = t_2\left\{\frac{(1+\delta_p)^2}{4} + V_{DD} - \frac{V_{DD} t_{max}}{t_T} - |V_{thp}|\right\} \quad \text{(B - 3)}$$

Appendix. C

Coefficient of eq.(23) are

$$c_1 = \frac{t_T}{2V_{DD}}\left\{\frac{\beta_n}{3(1+\delta_n)} - \frac{\beta_p}{2(1+\delta_p)}\right\} \quad \text{(C - 1)}$$

$$c_2 = -\frac{\beta_n t_{thn}}{2(1+\delta_n)} + \frac{\beta_p}{2(1+\delta_p)}\left\{\frac{t_T(V_{DD} - |V_{thp}|)}{V_{DD}} + \frac{t_{thn}}{2}\right\} \quad \text{(C - 2)}$$

$$c_3 = \frac{\beta_n t_T V_{thn}^2}{2V_{DD}(1+\delta_n)} - \frac{I_{\min} t_T}{2V_{DD}} + C_L$$
$$- \frac{\beta_p(V_{DD} - |V_{thp}|)}{4(1+\delta_p)}\left\{\frac{t_T(V_{DD} - |V_{thp}|)}{V_{DD}} + 2t_{thn}\right\} \quad \text{(C - 3)}$$

Appendix. D

Coefficient of eq. (28) are

$$e_1 = -3V_{thn} - \frac{3I_{max} t_T(1+\delta_n)}{\beta_n V_{DD}(t_{max} - t_0)} \quad \text{(D - 1)}$$

$$e_2 = 3V_{thn}^2 + \frac{6(1+\delta_n)}{\beta_n}$$
$$\left\{\frac{I_{max} t_0}{t_{max} - t_0} - \frac{V_{DD}}{t_T}\left(CM - \frac{C_L + C_M}{1+\delta_p}\right)\right\} \quad \text{(D - 2)}$$

$$e_3 = -V_{thn}^3 - \frac{3V_{DD}(1+\delta_n)}{\beta_n t_T}$$
$$\left\{\frac{I_{max} t_0}{t_{max} - t_0} - \left(\frac{2(V_{DD} - |V_{thp}|)(C_L + C_M}{1+\delta_p}\right)\right\} \quad \text{(D - 3)}$$

References

[1] Mark C. Johnson, Dinesh Somasekhar, and Kaushik Roy, "Models and algorithm for bounds on leakage in CMOS circuits," IEEE Trans. Computer-Aided Design, vol. 18, no. 6, June 1999, pp. 714-725.

[2] S. R. Vemuru and N. Scheinberg, "Short-circuit power dissipation estimation for CMOS logic gates," IEEE Trans. Circuits Syst. I, vol. 41, Nov. 1994, pp. 762-765.

[3] N. Hedenstierna and K. O. Jeppson, "CMOS circuit speed and buffer optimization," IEEE Trans. Computer-Aided Design, vol. CAD-6, Mar. 1987, pp. 270-281.

[4] N. Hedenstierna and K. O. Jeppson, "Comments on 'A module generator for optimized CMOS buffers' ," IEEE Trans. VLSI Syst., vol. 3, no. 3, Mar. 1995, pp. 99-111.

[5] A. Hirata, H. Onodera, and K. Tamaru, "Estimation of short-circuit power dissipation and its influence on propagation delay for static CMOS gates," in Proc. IEEE Int. Symp. on Circuits and Systems, vol. 4, May 1996, pp. 751-754.

[6] L. Bisdounis, S. Nikolaidis, and O. Koufopavlou, "Propagation delay and short-circuit power dissipation modeling of the CMOS inverter," IEEE Trans. Circuits Syst. I, vol. 45, no. 3, Mar. 1998, pp. 259-270.

[7] N. H. E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design A Systems Perspective*. New York: McGraw-Hill, 1993. pp. 183-191.

[8] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, New York: McGraw-Hill, 1988. pp. 123-130.

[9] William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, 1990. pp. 156-157.

[10] A. Hirata, H. Onodera, and K. Tamaru, "Estimation of propagation delay considering short-circuit current for static CMOS gates," IEEE Trans. Circuits Syst. I, vol. 45, no. 11, Nov. 1998, pp. 1194-1198.

[11] H. J. M. Veendrick, "Short-circuit dissipation of static CMOS circuitry and its impact on the design of buffer circuits," IEEE J. Solid-State Circuits, vol. SC-19, Aug. 1984, pp. 468-473.

[12] A. Papoulis, *Probability, Random Variables, and Stochastic Processes.*, International Student Edition, McGraw-hill, pp. 102-105.

[13] S. Turgis and D. Auvergne, "A novel macromodel for power estimation in CMOS structures," IEEE Trans. Computer-Aided Design, vol. 17, no. 11, Nov. 1998, pp. 1090-1098.

[14] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," IEEE J. Solid-State Circuits, vol. 25, Apr. 1990, pp. 584-594.