An Analytical Model for Current, Delay, and Power Analysis of Submicron CMOS Logic Circuits

Anas A. Hamoui, *Student Member, IEEE,* and Nicholas C. Rumin, *Senior Member, IEEE*

***Abstract*—An analytical model for computing the supply current, delay, and power of a submicron CMOS inverter is presented. A modified version of the nth power law MOSFET model is proposed and used to relate the terminal voltages to the drain current in submicron transistors. By first computing definable reference points on the output voltage waveform, and then using linear approximations through these points to find the actual points of interest, the desired speed and accuracy of the inverter model are achieved. The most important part of the analysis is a three-step approach for computing the time and output voltage when the short-circuit transistor changes its mode of operation. The time and output voltage when the charging/discharging current reaches its maximum are also calculated and then used to evaluate the propagation delay and characterize the output voltage waveform. The model has been validated for both 0.8 μm (5 V) and 0.25 μm (2.5 V) CMOS technologies, for a wide range of inverter sizes, input transition times, and capacitive loads. It predicts the delay, peak supply current, and power dissipation to within a few percent of HSPICE or ELDO simulations based on accurate physically based MOSFET models, while offering about two orders of magnitude gain in CPU time based on a MATLAB implementation.**

***Index Terms*—Analytical model, CMOS logic currents, delay estimation, inverter model, peak supply currents, power estimation, short-channel MOSFET models, short-circuit currents, short-circuit power dissipation, submicron MOSFETs, switching transition.**

I. Introduction

TO MINIMIZE the logic circuit design time, computer-aided design (CAD) tools must include efficient techniques for the rapid, yet reasonably accurate, estimation of critical path delays, power dissipation, and peak supply currents in digital integrated circuits. The problems of controlling the timing and the power consumption are growing as CMOS technology advances. For reliability design, the peak supply-current values are also needed to properly size the power and ground lines in order to avoid electromigration failures and voltage drop problems [1].

A number of methods for computing the delay and/or power dissipation in CMOS inverters have been recently presented [2]–[9]. The emphasis on modeling the inverter stems from the following. First, the worst-case delay can be simulated by replacing complex CMOS gates with their worst-case equivalent inverters. Second, a number of efficient transistor-level techniques for reducing CMOS logic gates to equivalent inverters are available [10], [11]. Third, and most important, the clock distribution networks and busses in a digital VLSI chip are based on inverters or inverter-like circuits which must be carefully designed and modeled. These circuits account for a large fraction of the total power consumption.

In this paper, an analytical model for computing the supply current, delay, and power of a submicron CMOS inverter is presented. The effect of the Miller capacitance is also modeled. A modified version of the nth power law MOSFET model [12] is proposed and used to relate the terminal voltages to the drain current in submicron transistors.

The outstanding feature of the inverter model proposed in this paper is its comprehensiveness: it computes the maximum currents, in addition to both the delay and power, and the same model is used regardless of whether the input voltage switching transition is fast or slow. Furthermore, by first computing definable reference points on the output voltage waveform and then using linear approximations through these points to find the actual points of interest, the desired speed and accuracy of the inverter model are achieved.

II. Submicron MOSFET Model

A desirable submicron MOSFET model for the fast analysis of CMOS ICs involves a small number of parameters, is reasonably accurate, and does not require computationaly expensive procedures to extract the model parameters. In particular, the nth power law model, proposed by Sakurai and Newton [12], offers a simple, yet accurate enough, empirical model for the MOSFET drain current:

1) $V_{DS} \leq V_{DS\text{sat}}$ (linear region)

$$I_D = B\frac{W_e}{L_e}(V_{GS}-V_t)^n\left(2\frac{V_{DS}}{V_{DS\text{sat}}}-\frac{V_{DS}^2}{V_{DS\text{sat}}^2}\right)\cdot(1+\lambda V_{DS}) \quad (1)$$

2) $V_{DS} \geq V_{DS\text{sat}}$ (saturation region)

$$I_D = B\frac{W_e}{L_e}(V_{GS}-V_t)^n(1+\lambda V_{DS}) \quad (2)$$

where the MOSFET region of operation is determined by the drain–source saturation voltage

$$V_{DS\text{sat}} = K(V_{GS}-V_t)^m. \quad (3)$$

Manuscript received March 26, 2000; revised June 14, 2000. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. This paper was recommended by Associate Editor M. Bayoumi.

A. A. Hamoui was with the Department of Electrical and Computer Engineering, McGill University, Montreal, PQ, H3A 2A7, Canada. He is now with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4, Canada.

N. C. Rumin is with the Department of Electrical and Computer Engineering, McGill University, Montreal, PQ, H3A 2A7, Canada.

Publisher Item Identifier S 1057-7130(00)09319-8.

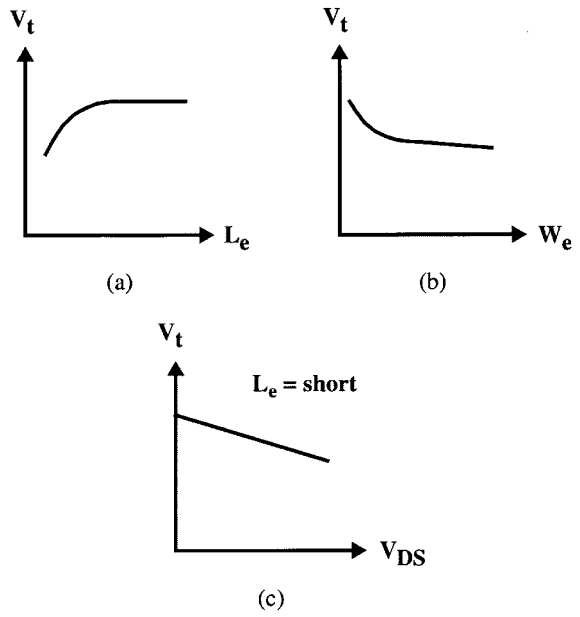


Fig. 1. Sketch of the dependence of the MOSFET threshold voltage V_t on: (a) the effective channel length L_e; (b) the effective channel width W_e; and (c) the drain–source voltage V_{DS} in short-channel devices (drain-induced barrier lowering effect).

The technology-dependent constants $n, m, K,$ and B describe the short-channel effects in an empirical manner, while $\lambda = \lambda_0 + \lambda_1 V_{SB}$ models the channel-length modulation effect. L_e and W_e are, respectively, the effective channel length and width. V_{DS}, V_{GS}, and V_{SB} are the drain–source, gate–source, and source–bulk voltages, respectively, while V_t denotes a threshold voltage.

However, the nth power law model neglects the threshold-voltage variations due to the short length, narrow width, and drain-induced barrier lowering (DIBL) effects, which are significant in submicron MOS technologies (as illustrated in Fig. 1) [13]. To model these variations, the threshold voltage at zero body-bias can be expressed as a linear function of the effective channel length-to-width ratio [14]. Thus, the nth power law MOSFET model [12] has been augmented [14] with the following equation:

$$V_t = V_{tw}\left(1 + f\frac{L_e}{W_e}\right) + \gamma V_{SB}. \tag{4}$$

V_t denotes a threshold voltage and V_{tw} is the corresponding zero body-bias threshold voltage for wide-channel transistors. The empirical factor f describes the dependence of V_t on L_e and W_e, while γ models the body effect.

For a given feature size, a simple one-time procedure is then followed to optimize the set of seven parameters, in the ***modified*** nth power law model [(1)–(4)], for the MOSFET equations to best fit the measured I_D–V_{DS} characteristics (for different V_{GS} values) over the range of nMOS/pMOS channel widths used in circuit design [14]. Note that, although V_t was assumed independent of V_{DS} in (4), the DIBL effect is still implicitly accounted for to some extent by extracting a V_{tw} value optimized over the full range of operating V_{DS} voltages.

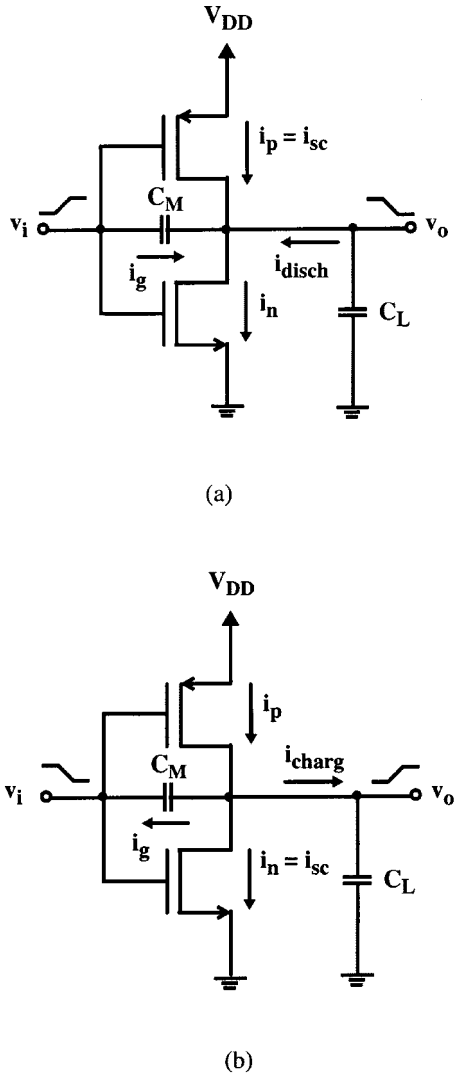


Fig. 2. CMOS inverter circuit. (a) Discharging inverter (rising input). (b) Charging inverter (falling input).

III. The CMOS Inverter

Consider the CMOS inverter circuit in Fig. 2. The effective load C_L includes the drain-bulk junction capacitances of the nMOS and pMOS transistors, the gate-to-bulk and gate-to-source capacitances of the nMOS and pMOS devices of the driven inverters (i.e., the input capacitances of the fanout gates), and the interconnect capacitances. The Miller capacitance C_M consists of the nMOS and pMOS gate-to-drain capacitances. The nonlinear voltage-dependent MOSFET parasitic capacitances are replaced by equivalent constant capacitances. Over each MOSFET mode of operation, the intrinsic gate capacitance is assumed to be a constant fraction of the effective gate-oxide capacitance [15].

For the discharging inverter, the input voltage waveform is assumed to be a rising ramp with transition time T_r

$$v_i(t) = \begin{cases} s_r t, & 0 \le t \le T_r \\ V_{DD}, & t > T_r \end{cases} \tag{5}$$

where $s_r \equiv V_{DD}/T_r$ is the slope of the rising input voltage ramp. This input waveform approximation is widely accepted

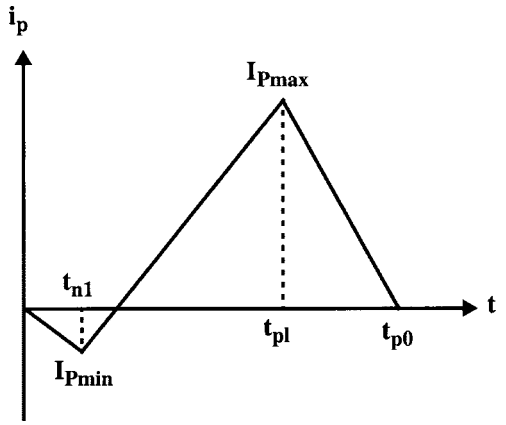


Fig. 3. Piecewise linear approximation of the short-circuit current i_p, used to compute the short-circuit energy dissipation E_{sc} of the discharging inverter (rising input).

because of its simplicity and effectiveness. The differential equation describing the discharging of the CMOS inverter is then given by

$$\frac{dv_o}{dt} = -\frac{i_n - i_p}{C_L + C_M} + \frac{C_M}{C_L + C_M} s_r, \qquad \text{for } 0 \le t \le T_r. \tag{6}$$

In the following analysis, the current, delay, and power are derived for the case of a discharging inverter. The analysis for the charging inverter case is symmetrical.

IV. Power Dissipation

For the CMOS inverter circuit in Fig. 2, the dynamic energy dissipation per switching event (i.e., the full charging and discharging of the output node) is given by

$$E = C_L V_{DD}^2 + 2C_M V_{DD}^2 + E_{sc}. \tag{7}$$

The first two terms represent the energy dissipation due to the charging and discharging of, respectively, the effective load capacitance C_L and the Miller capacitance C_M. The short-circuit energy dissipation E_{sc} is due to the direct-path current from supply to ground when the nMOS and pMOS devices are simultaneously on. Note that, for the discharging (charging) inverter in Fig. 2, the nMOS (pMOS) transistor is the discharging (charging) transistor while the pMOS (nMOS) transistor is referred to as the short-circuiting transistor. As will be shown in Section VI, E_{sc} can account for more than 35% of E. Furthermore, with the ongoing trend toward scaling down the supply voltage and the minimum feature size in CMOS IC's, the contribution of the short-circuit current to the total power dissipation is increasing.

A simple, yet accurate enough, approach for evaluating E_{sc} is to approximate the short-circuit currents (i_n in the charging inverter and i_p in the discharging inverter) by piecewise linear functions of time [8]. This is shown in Fig. 3 for the case of a discharging inverter. Thus, we can express the component of E_{sc} for the discharging inverter as

$$E_{sc}|_{\text{discharge}} = \frac{V_{DD}}{2}[I_{P\max}(t_{p0} - t_{n1}) + I_{P\min} t_{pl}] \tag{8}$$

where $t_{n1} \equiv V_{ten}/s_r$ and $t_{p0} \equiv (V_{DD} - |V_{tep}|)/s_r$ are the times when the nMOS and pMOS devices turn, respectively, on and off. Here, V_{ten} and $|V_{\text{tep}}|$ are the effective nMOS and pMOS threshold voltages, respectively, and are extracted from the I_D–$|V_{GS}|$ characteristics at $|V_{DS}| = 0.01V_{DD}$ [14]. Note that, as discussed in [14], the empirical parameter V_t in equation (4) is significantly larger than the threshold voltage as it is normally defined (i.e., the $|V_{GS}|$ needed to induce a strongly inverted channel under the gate).

Hence, the objective is to determine $I_{P\min}$ and $I_{P\max}$, as well as their times of occurrence. The determination of the former is straightforward and is discussed in [14].

A. Maximum Short-Circuit Current $I_{P\max}$

Let t_{pl} be the time when the short-circuiting pMOS transistor leaves the linear region and enters saturation. Simulation results have shown that, for the purpose of computing E_{sc}, it is valid to assume that the short-circuit current reaches it maximum value $I_{P\max}$ at $t = t_{pl}$. The special case of $t_{pl} = t_{p0}$ corresponds to very fast input ramps where the pMOS device turns off before entering saturation. This occurs if v_i reaches $V_{DD} - |V_{tp}|$ (switching the pMOS transistor off) before the output voltage waveform has completed its overshoot and v_o has dropped below $V_{DD} - |V_{tp}|$.

Let t_{ns} be the time when the nMOS device leaves saturation and enters the linear region. Since at $v_i = v_o$ both the nMOS and pMOS transistors must be in saturation, the pMOS device must enter saturation before the nMOS device leaves it. Therefore, we have $t_{pl} \le t_{nm1}$, where $t_{nm1} \equiv \min(t_{ns}, t_{p0})$.

During the time interval $t_{n1} \le t \le t_{nm1}$, the pMOS transistor operates in its linear region until time t_{pl}, when it saturates. The nMOS device, on the other hand, remains saturated over the entire time interval. A three-step approach is used to evaluate t_{pl}. First, the short-circuit current i_p is neglected and an approximation to t_{pl} is computed. Second, this approximate time is corrected for the short-circuit current (neglected in the first step), yielding a point on the inverter's switching trajectory close to t_{pl}. Finally, the tangent to the output voltage waveform at this point is used to compute t_{pl} and $v_o(t_{pl})$.

Step 1) Assume $i_p = 0$, i.e., neglect the short-circuit current. Furthermore, since t_{pl} occurs at the early stage of the falling output voltage and $V_{DSn} = v_o$, the channel modulation effect can be approximated by $1 + \lambda_n V_{DSn} \approx 1 + \lambda_n V_{DD}$ in the nMOS drain current equation.

Equation (6), with i_n expressed in terms of its terminal voltages using the *modified* nth power law equations, is now solved to get an approximation for the output voltage waveform during the time interval $t_{n1} \le t \le t_{nm1}$. If $v_o(t_{p0}) > V_{DD} - |V_{tp}|$, then the pMOS device will turn off before entering saturation. In this case, $t_{pl} = t_{p0}$ is used and the following steps are skipped. Otherwise, t_{pl} is solved for, noting that the output voltage when the pMOS device changes its mode of operation is given by

$$\begin{aligned} v_o(t_{pl}) &= V_{DD} - V_{SDp_{\text{sat}}} \\ &= V_{DD} - K_p(V_{DD} - |V_{tp}| - s_r t_{pl})^{m_p}. \end{aligned} \tag{9}$$

Step 2) Since the short-circuit current i_p was neglected in Step 1, the computed values of t_{pl} and $v_o(t_{pl})$, denoted $\tilde{t}_q$ and V_{Oq}, respectively, are only *approximations* to the true values. The effective current available to discharge the load is actually only $i_n - i_p$ because the pMOS transistor is on during the time

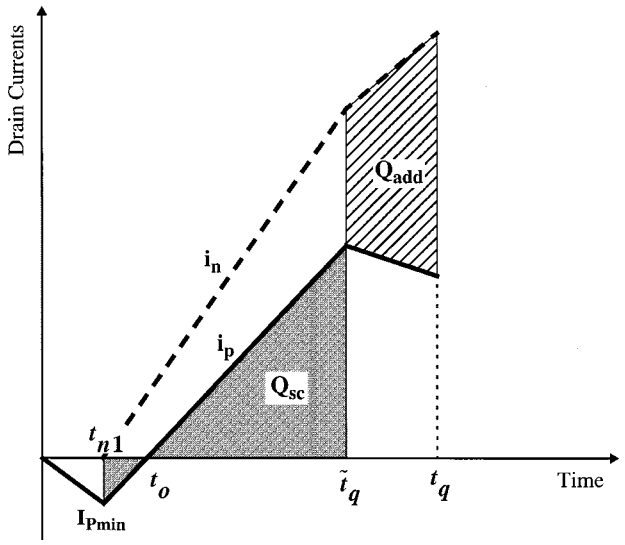


Fig. 4. Piecewise-linear approximations of the discharging current i_n ($- - -$) and the short-circuit current i_p (——), used in Step 2 of the derivation of the maximum short-circuit current $I_{P\max}$ for the discharging inverter (rising input).

interval $t_{n1} \le t \le t_{p0}$. Hence, for the output voltage to actually drop to V_{Oq}, the output node must be discharged by

$$Q_q = \int_{t_{n1}}^{\tilde{t}_q} i_n\, dt = \int_{t_{n1}}^{t_q} (i_n - i_p)\, dt \tag{10}$$

where t_q is the *actual* time required for the output voltage to drop to V_{Oq}. Hence, defining

$$Q_{sc} \equiv \int_{t_{n1}}^{\tilde{t}_q} i_p\, dt \quad \text{and} \quad Q_{\text{add}} \equiv \int_{\tilde{t}_q}^{t_q} (i_n - i_p)\, dt \tag{11}$$

it follows from (10) that $Q_{\text{add}} = Q_{sc}$.

Here, Q_{sc} represents the amount of charge which leaked from the power supply through the short-circuiting pMOS transistor during the time interval $t_{n1} \le t \le \tilde{t}_q$. To compensate for Q_{sc}, the output node must be discharged, during the time interval $\tilde{t}_q \le t \le t_q$, by a net additional charge Q_{add} to allow the output voltage to actually drop to V_{Oq}.

To compute Q_{sc} and Q_{add}, the drain currents $i_n(t)$ and $i_p(t)$ are represented by piecewise linear functions of time, as shown in Fig. 4. The current values at $t = \tilde{t}_q$ are calculated from the nMOS and pMOS drain current equations based on their respective terminal voltages, with the approximation $v_o(\tilde{t}_q) \approx V_{Oq}$. For $t \ge \tilde{t}_q$, the drain currents are described by linear functions of time with rates of change equal to those at $t = \tilde{t}_q$. Thus, equating Q_{sc} and Q_{add} yields

$$t_q = \tilde{t}_q + \frac{-\Delta I + \sqrt{\Delta I^2 + 2Q_{sc}\Delta di}}{\Delta di} \tag{12}$$

with $\Delta I \equiv i_n(\tilde{t}_q) - i_p(\tilde{t}_q)$ and $\Delta di \equiv d/dt(i_n - i_p)|_{t=\tilde{t}_q}$. From Fig. 4

$$Q_{sc} = \frac{1}{2} I_{P\min}(t_o - t_{n1}) + \frac{1}{2} i_p(\tilde{t}_q)(\tilde{t}_q - t_o) \tag{13}$$

where $t_o \equiv (i_p(\tilde{t}_q)t_{n1} - I_{P\min}\tilde{t}_q)/(i_p(\tilde{t}_p) - I_{P\min})$.

Note that in [2], Embabi and Damodaran use an expression derived by Sakurai and Newton in [16] to compute $\tilde{t}_{50}$, the approximate time corresponding to $v_o = V_{DD}/2$. They then consider an approach similar to ours to improve on this estimate of $\tilde{t}_{50}$. However, in their approach, the output voltage, whose corresponding time $\tilde{t}_{50}$ is to be improved, is preset to a value of $V_{DD}/2$. Therefore, the times when the nMOS and pMOS transistors change their mode of operation (i.e., t_{pl} and t_{ns}) must be first determined to be able to compute Q_{sc}. To find t_{pl} and t_{ns}, the output voltage waveform was assumed by them to fall linearly from V_{DD} at $t = t_{n1}$ to $V_{DD}/2$ at $t = \tilde{t}_{50}$, which is a large voltage excursion. As a result, the nonlinear behavior of the transistors was not accurately accounted for.

In our approach, on the other hand, a reference point in time $\tilde{t}_q$ is to be improved on. Since $\tilde{t}_q$ was derived to correspond to t_{pl} for the case of $i_p = 0$ (i.e., neglecting the short-circuit current), it is by definition smaller than the actual t_{pl}. Therefore, the mode of operation of both devices over the entire time interval $t_{n1} \le t \le \tilde{t}_q$ is known: the nMOS is in saturation and the pMOS is in the linear region. Hence, Q_{sc} can be simply computed (as described above), regardless of how the output voltage is changing.

Step 3) Now, $[t_q,\ V_{Oq}]$ represents an actual point on the output voltage waveform very close to the desired point $[t_{pl},\ v_o(t_{pl})]$. Therefore, the output voltage waveform near t_{pl} can be approximated by the tangent line at $[t_q,\ V_{Oq}]$, whose slope is readily obtained from (6). Using this linear approximation to solve (9) yields an improved value of t_{pl}, which takes into account the short-circuit current. The corresponding output voltage $v_o(t_{pl})$ can now be determined, and the maximum value $I_{P\max}$ of the short-circuit current is computed with $V_{SGp} = V_{DD} - s_r t_{pl}$ and $V_{SDp} = V_{DD} - v_o(t_{pl})$.

V. Propagation Delay and Maximum Discharging Current

In a CMOS inverter circuit driving a capacitive load, the output voltage transition can be properly characterized by the tangent line to the output voltage waveform at the time when the charging/discharging current reaches its maximum. This is shown in Fig. 5 for the case of a discharging inverter. The derivation of the delay time using this approach is straightforward and is discussed in [14]. However, to evaluate the delay, the time t_{nm} and output voltage $v_o(t_{nm})$ when the discharging current reaches its maximum $\mathbf{I_{N\max}}$ must first be computed, as described below.

A. Maximum Discharging Current $I_{N\max}$

The discharging current i_n reaches its maximum when the nMOS transistor leaves saturation and enters the linear region (at $t = t_{ns}$), but not later than the time when V_{GSn} attains its maximum value of V_{DD} (at $t = T_r$). Defining $t_{nm1} \equiv \min(t_{ns},\ t_{p0})$ and $t_{nm2} \equiv \min(t_{ns},\ T_r)$, it follows that the time t_{nm} (when $i_n = I_{N\max}$) must occur within one of the following two intervals.

1) *Time Interval 1:* $t_{pl} \le t \le t_{nm1}$

Both the pMOS and nMOS devices are saturated. For $t \le t_{nm}$, $V_{DSn} = v_o(t)$ is larger than $V_{SDp} = V_{DD} - V_{DSn}$ because $v_o(t)$ is a falling signal and $v_o(t_{nm})$ is

close to $V_{DD}/2$. Thus, a possible simplifying assumption, to be used in the drain current equations for the nMOS and pMOS devices, is: $1+\lambda_n V_{DSn} \approx 1+\lambda_n V_{DD}$ and $1+\lambda_p V_{SDp} \approx 1$.

An expression for the output voltage waveform during the time interval $t_{pl} \le t \le t_{nm1}$ can be obtained by solving equation (6), with i_n and i_p expressed in terms of their respective terminal voltages using the *modified* nth power law equations and with initial condition $v_o(t_{pl})$ (computed in Section IV). This, combined with the value of the output voltage when the nMOS device changes its mode of operation

$$v_o(t_{ns}) = V_{DSn_{\rm sat}} = K_n(s_r t_{ns} - V_{tn})^{m_n} \qquad (14)$$

yields t_{ns}. If $t_{ns} \le t_{p0}$, then $t_{nm} = t_{ns}$. Otherwise, time interval 2 must be used to compute t_{ns}.

2) *Time Interval 2:* $t_{p0} \le t \le t_{nm2}$

The pMOS device is off ($i_p = 0$), while the nMOS device is saturated. Steps similar to those above yield the output voltage waveform during the time interval $t_{p0} \le t \le t_{nm2}$ and the time t_{ns}. If $t_{ns} \le T_r$, then $t_{nm} = t_{ns}$. Otherwise, $t_{nm} = T_r$.

The maximum value $I_{N\max}$ of the discharging current is finally computed with $V_{GSn} = s_r t_{nm}$ and $V_{DSn} = v_o(t_{nm})$.

VI. Results

The proposed analytical model, implemented in MATLAB, has been tested with a wide range of inverters designed in both a 5-V 0.8-μm BiCMOS process and a 2.5-V 0.25-μm CMOS technology. To validate the model, the delay, peak supply current, and power dissipation were compared with the "exact" values obtained by simulating the circuits in the ELDO simulator using Nortel's MISNAN MOSFET model [17] for the 0.8-μm process, and in HSPICE using the BSIM3(V3.1) model for the 0.25-μm technology. Most of the results presented here are for the 0.25-μm technology. References [14] and [18] contain results for the older technology.[1]

A. Delay

In Fig. 6(a), the delays computed using the proposed model are compared with those produced by HSPICE for a very large inverter ($W_n = W_p = 12$ μm). Note the agreement even for a load of 250 fF, which is small for such a large inverter. Then, to verify the validity of the model over a wide range of inverter sizes, the delays in the case of a minimum size inverter ($W_n = W_p = 0.8$ μm) are compared in Fig. 6(b). The accuracy of the model over a wide range of switching conditions is demonstrated in Fig. 7, where the input transition time T_i is varied over the range $0.1 \rightarrow 1$ ns, and the corresponding delay is plotted for several values of W_p/W_n. Note that $T_i = 0.1$ ns corresponds approximately to the transition time of the char-

[1]All results in the following are quoted for the 0.25-μm CMOS technology unless it is stated otherwise.

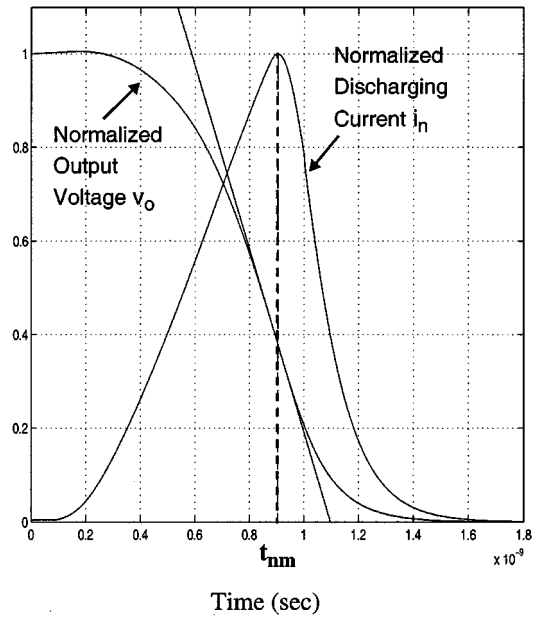
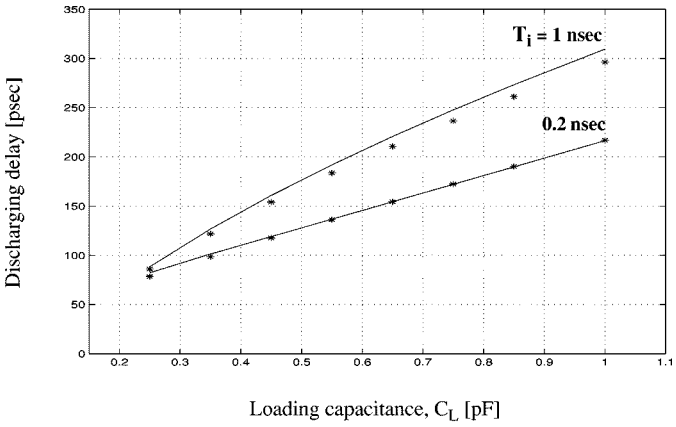
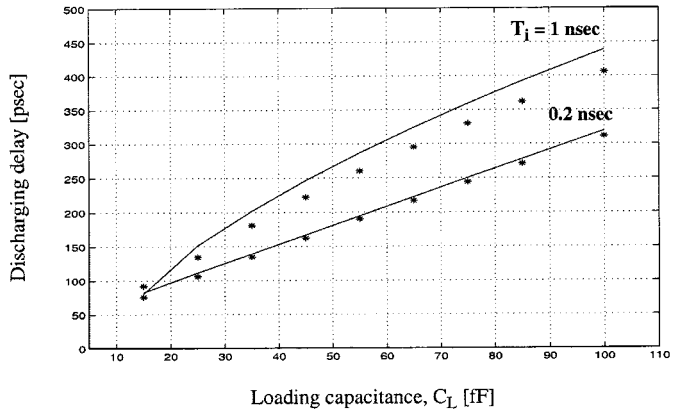


Fig. 5. Linear approximation of the output voltage waveform for a discharging inverter (rising input). The output voltage transition can be fully characterized by the tangent line to the output voltage waveform at time t_{nm}, when the discharging current i_n reaches its maximum.



(a)



(b)

Fig. 6. Dependence of the delay time on the loading capacitance for inverters (in 0.25-μm CMOS technology) with fast and slow input transition times, T_i. (a) Very large CMOS inverter ($W_n = W_p = 12$ μm). (b) Minimum size CMOS inverter ($W_n = W_p = 0.8$ μm). Proposed model: (***). HSPICE simulation: (—).

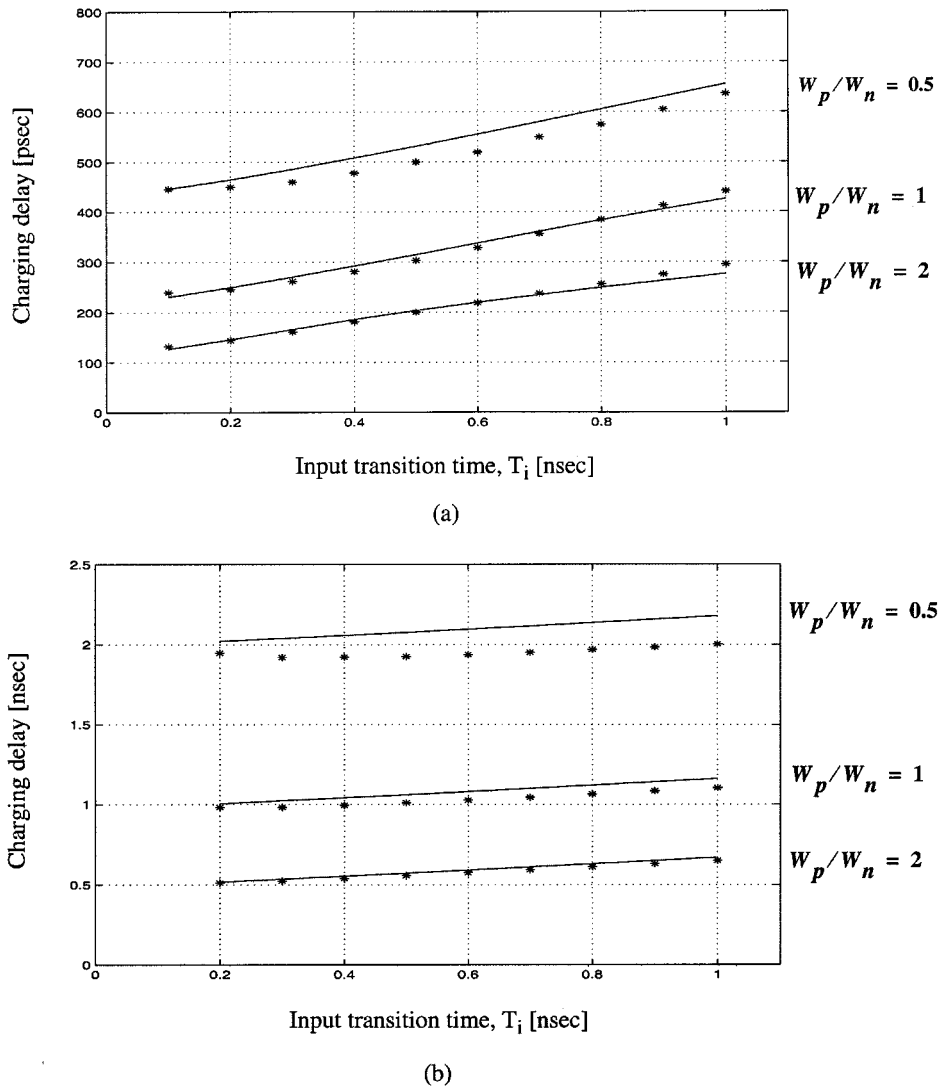


Fig. 7. Dependence of the delay time on the input transition time for inverters (in 0.25-μm CMOS technology) with different transistor-size ratios W_p/W_n ($W_n = 4$ μm), and for small and large loading capacitances C_L: a) $C_L = 150$ fF and b) $C_L = 750$ fF. Proposed model (***). HSPICE simulation: (—).

acteristic waveform[2] for the 0.25-μm technology. Results from HSPICE simulations are also given for comparison. For completeness, this test is performed using a medium size inverter with $W_n = 4$ μm. The maximum error in the delay, computed using the proposed model, is less than 8% compared to HSPICE simulation results.

B. Peak Power-Supply Current

In Fig. 8(a), the input transition time T_i is varied and the corresponding peak supply current is plotted for several values of the inverter aspect ratio, W_p/W_n. Note how the peak supply currents computed using the proposed model follow precisely the nonlinear change (with input transition time) of the "exact" peak-supply-current curves produced by HSPICE. To verify the validity of the model over various loading conditions, the peak supply current for $C_L = 0.15 \rightarrow 1$ pF is displayed in Fig. 8(b) for several values of W_p/W_n. Since the peak supply current is largest for fast input transitions, the test was performed for $T_i = 0.5$ ns. HSPICE simulation results are also plotted for comparison. The maximum error observed for the peak supply current is 2.5%.

C. Power Dissipation

Table I compares the energy dissipations per switching event computed using the proposed inverter model for the 0.8-μm CMOS technology with those obtained ELDO simulations. Both the short-circuit energy E_{sc} and the total dynamic energy E are given. Inverters of different sizes were tested under diverse switching conditions of input transition time and capacitive load. The maximum error in E, calculated using the analytical model, is 5%, while the corresponding ratio of E_{sc}/E is 38.8%. Therefore, in addition to proving the validity of the model, this confirms that the short-circuit power dissipation can no longer be neglected in submicron CMOS circuits, even with 0.8-μm channel lengths.

[2]The characteristic waveform is defined as the definite waveform toward which the voltage waveform converges in a series of identical inverters.

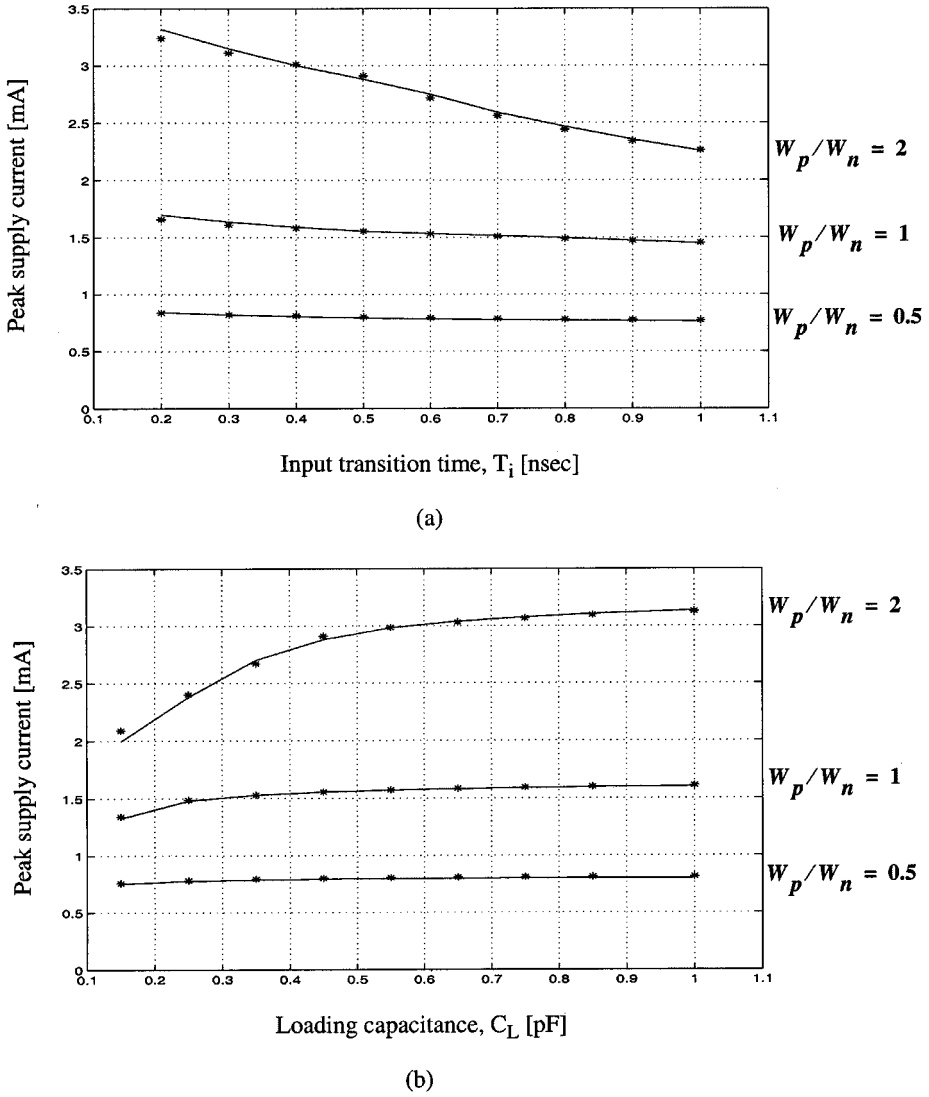


Fig. 8. Variation of the peak power-supply current with: (a) input transition time T_i ($C_L = 450$ fF) and (b) loading capacitance C_L ($T_i = 0.5$ nsec), for inverters (in 0.25 μm CMOS technology) with different transistor-size ratios W_p/W_n ($W_n = 6$ μm). Proposed model: (***). HSPICE simulation: (—).

D. CPU Time

Since the CPU time required to simulate a circuit using HSPICE or ELDO depends strongly on the time step and the duration of the transient analysis (or stop time), the following precautions were taken in running the HSPICE or ELDO simulations.

1) The *time step* was set to the largest step which still allows the capture of the delay and the time of the peak supply current to the nearest 2% T_C, where T_C is the transition time of the characteristic waveform for the technology (0.5 ns for the 0.8-μm and 0.1 ns for the 0.25-μm CMOS technology).
2) The *stop time* was selected to correspond to the shortest duration of the transient analysis which yields the supply energy dissipation to within 5%.

Several inverters of different sizes and various loads were simulated for various input transition times. In each case, an initial simulation was carried out to determine the required time-step and stop-time settings in the simulator. Then, the simulator was rerun several times to find the average CPU time required to simulate the inverter circuit. Results show that the inverter model, run in MATLAB, offers about two orders of magnitude improvements in CPU time over HSPICE or ELDO. It can therefore be expected to be significantly faster if coded in C.

VII. Conclusion

An analytical model for computing the supply current, delay, and power of a submicron CMOS inverter has been presented. It is based on a modified version of the nth power law MOSFET model. The inverter model's accuracy is achieved by computing reference points on the output voltage waveform, which are defined in terms of the states of the transistors, and then using linear approximations through these points to find the actual points of interest. The most important part of the analysis is a three-step process for computing the time and output voltage when the transistor carrying the "short-circuit" current changes its mode of operation. The time and output voltage when the

TABLE I
SHORT-CIRCUIT ENERGY DISSIPATION AND TOTAL DYNAMIC ENERGY DISSIPATION PER SWITCHING EVENT. RESULTS COMPUTED USING THE PROPOSED MODEL ARE COMPARED WITH ELDO SIMULATION RESULTS FOR DIFFERENT CMOS INVERTERS (WITH TRANSISTOR WIDTHS W_n AND W_p, IN 0.8-μm CMOS TECHNOLOGY) UNDER VARIOUS CONDITIONS OF INPUT TRANSITION TIME T_i AND LOADING CAPACITANCE C_L

T_i (ns)	W_n (μm)	W_p (μm)	C_L (fF)	Short-Circuit Energy E_{sc} (pJ)		Total Dynamic Energy E (pJ)		$\% \frac{E_{sc}}{E}$	% Error in E
				Analytical Model	ELDO Simulator	Analytical Model	ELDO Simulator		
2	6	6	100	0.951	1.29	4.45	4.51	28.6	1.3
			250	0.837	0.900	8.07	7.88	11.4	2.4
		12	100	1.65	2.10	5.63	5.67	37.0	0.70
			250	1.50	1.55	9.22	8.87	17.4	3.9
	12	12	200	1.94	2.65	8.878	9.079	29.1	2.2
			350	1.88	2.17	12.56	12.35	17.5	1.7
		36	200	3.32	4.29	11.24	11.40	37.6	1.4
			350	3.26	3.63	14.92	14.49	25.0	2.9
4	6	6	250	1.99	2.74	9.23	9.72	28.1	5.0
		12		3.42	4.52	11.2	11.80	38.3	5.0

charging/discharging current reaches its maximum are also calculated and then used to evaluate the propagation delay and to characterize the output voltage waveform. The model has been validated on the basis of accurate, physically-based, submicron MOSFET models using both 0.8-μm (5 V) and 0.25-μm (2.5 V) CMOS process parameters. Results have been presented for a wide range of inverter sizes, input transition times, and capacitive loads. They demonstrate that the proposed analytical model can predict the delay, peak supply current, and power dissipation to within 8% of HSPICE or ELDO simulation results, while offering about two orders of magnitude gains in CPU time. Since the tested model was implemented in MATLAB, one can expect a significant increase in performance from a fully coded implementation.

ACKNOWLEDGMENT

The Canadian Microelectronic Corporation (CMC) provided the equipment, CAD tools, and technical support, and facilitated the access to NORTEL's 0.8-μm and TSMC's 0.25-μm technologies. The authors are very grateful to M. Oulmane for testing the model in the 0.25-μm CMOS technology.

REFERENCES

[1] W. S. Song and L. A. Glasser, "Power distribution techniques for VLSI circuits," *IEEE J. Solid-State Circuits*, vol. 21, pp. 150–156, Feb. 1986.
[2] S. H. K. Embabi and R. Damodaran, "Delay models for CMOS, BiCMOS and BinMOS circuits and their applications for timing simulations," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 1132–1142, Sept. 1994.
[3] S. Dutta, S. S. Mahant Shetti, and S. L. Lusky, "A comprehensive delay model for CMOS inverters," *IEEE J. Solid-State Circuits*, vol. 30, pp. 864–871, Aug. 1995.
[4] K. O. Jeppson, "Modeling the influence of the transistor gain ratio and the input-to-output coupling capacitance on the CMOS inverter delay," *IEEE J. Solid-State Circuits*, vol. 29, pp. 646–654, June 1994.
[5] L. Bisdounis, S. Nikolaidis, and O. Koufopavlou, "Analytical transient response and propagation delay evaluation of the CMOS inverter for short-channel devices," *IEEE J. Solid-State Circuits*, vol. 33, pp. 302–306, Feb. 1998.
[6] H-J Park and M. Soma, "Analytical model for switching transitions of submicron CMOS logics," *IEEE J. Solid- State Circuits*, vol. 32, pp. 880–889, June 1997.
[7] S. R. Vemuru and N. Scheinberg, "Short-circuit power dissipation estimation for CMOS logic gates," *IEEE Trans. Circuits Syst. I*, vol. 41, pp. 762–765, Nov. 1994.
[8] A. Hirata, H. Onodera, and K. Tamaru, "Estimation of short-circuit power dissipation for static CMOS gates," *IEICE Trans. Fundamentals*, vol. E79-A, pp. 304–311, Mar. 1996.
[9] S. Turgis and D. Auvergne, "A novel macromodel for power estimation in CMOS structures," *IEEE Trans. Computer-Aided Design*, vol. 17, pp. 1090–1098, Nov. 1998.
[10] A. Nabavi-Lishi and N. C. Rumin, "Inverter models of CMOS gates for supply current and delay evaluation," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 1271–1279, Oct. 1994.
[11] A. Chatzigeorgiou, S. Nikolaidis, and I. Tsoukalas, "A modeling technique for CMOS gates," *IEEE Trans. Computer-Aided Design*, vol. 18, pp. 557–575, May 1999.
[12] T. Sakurai and A. R. Newton, "A simple MOSFET model for circuit analysis," *IEEE Trans. Electron Devices*, vol. 38, pp. 887–894, Apr. 1991.
[13] K. Lee, M. Shur, T. A. Fjeldly, and T. Ytterdal, *Semiconductor Device Modeling for VLSI*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[14] A. Hamoui, "Current, delay, and power analysis of submicron CMOS circuits," M.Eng., McGill Univ., Montreal, Canada, 1998.
[15] N. H. E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design: A Systems Perspective*, 2nd ed. Reading, MA: Addison-Wesley, 1993.
[16] 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, pp. 584–594, Apr. 1990.

[17] A. R. Boothroyd, S. W. Tarasewicz, and C. Slaby, "MISNAN—A physically-based continuous MOSFET model for CAD applications," *IEEE Trans. Computer-Aided Design*, vol. 10, pp. 1512–1529, Dec. 1991.

[18] A. A. Hamoui and N. C. Rumin, "An analytical current, delay, and power model for the submicron CMOS inverter," in *Proc. 6th IEEE Int. Conf. Electronics, Circuits and Systems*, Pafos, Cyprus, Sept. 1999, pp. 1547–1551.

Anas A. Hamoui (S'95) was born in Damascus, Syria, in 1973. He received the M.Eng. degree in electrical engineering from McGill University, Montreal, Canada, in 1998. He is currently pursuing the Ph.D. degree in electrical engineering at the University of Toronto, Toronto, Canada, supported by a postgraduate scholarship from the Natural Sciences and Engineering Research Council of Canada.

From 1996 to 1998, he was a Research Assistant at the Microelectronics and Computer Systems Laboratory, McGill University, working in the area of timing and power analysis of submicron CMOS circuits. Since September 1998, he has been with the Electronics Group, University of Toronto. His current research interests include mixed-signal integrated-circuit design for wireless data-communication systems, and high-speed high-resolution A/D data conversion.

Nicholas C. Rumin (S'60–M'65–SM'78) received the B.Eng., M. Sc., and Ph.D. degrees in 1957, 1961, and 1966, respectively, all from McGill University, Montreal, Canada.

From 1957 to 1959, he was a Design Engineer with the Canadian Marconi Company, and between 1961–1963, he was a Research Associate with the Montreal Neurological Institute, Montreal, Canada. After spending two years with RCA Research Laboratories, Montreal, Canada, he joined the Department of Electrical and Computer Engineering, McGill University, in 1967, where he is currently a Professor, previously having served as Chairman of the Department between 1988 and 1998. In 1975, he spent a sabbatical leave with the Bell Northern Research Laboratories in Ottawa, Ontario, Canada. His current research interests include the modeling of integrated circuits for delay, power and current analysis.

Dr. Rumin is a member of the Ordre des Ingenieurs de Quebec.