Propagation Delay and Short-Circuit Power Dissipation Modeling of the CMOS Inverter

Labros Bisdounis, *Student Member, IEEE,* Spiridon Nikolaidis, *Member, IEEE,* and Odysseas Koufopavlou, *Member, IEEE*

***Abstract*— This paper introduces a new, accurate analytical model for the evaluation of the delay and the short-circuit power dissipation of the CMOS inverter. Following a detailed analysis of the inverter operation, accurate expressions for the output response to an input ramp are derived. Based on this analysis improved analytical formulae for the calculation of the propagation delay and short-circuit power dissipation, are produced. Analytical expressions for all inverter operation regions and input waveform slopes are derived, which take into account the influences of the short-circuit current during switching, and the gate-to-drain coupling capacitance. The effective output transition time of the inverter is determined in order to map the real output voltage waveform to a ramp waveform for the model to be applicable in an inverter chain. The final results are in very good agreement with SPICE simulations.**

I. Introduction

SWITCHING-SPEED is one of the most critical performance parameters in VLSI circuits. Much effort has to be devoted for the extraction of accurate, analytical expressions for timing models of basic circuits. These expressions can be incorporated in switch-level and logic simulators, optimizing the design verification procedure. Much of past research has addressed the development of delay models for CMOS circuits. In some of these models [1]–[3] *RC* circuit approaches were used in order to map the transistors to equivalent resistors. However, switch-level simulators based on *RC* models tend to model the average circuit behavior only, since the nonlinear behavior of the transistors is not well represented by linear and constant resistors.

Macromodeling approaches in order to achieve efficient delay modeling have been proposed. Brocco *et al.* [4] presented an approach based on look-up tables which are generated via precharacterization using SPICE simulations. The methods which are based on tables with presimulation results are time consuming and incorporate interpolation errors. Other approaches derived delay expressions which take into account the slope of the input waveform. However they use approximations for the load currents by assuming mean charge conservation across the CMOS structure [5], or use pseudoempirical factors obtained from SPICE simulations in order to evaluate delays in CMOS structures controlled by slowly varying input signals [6].

Deriving expressions for accurately describing the propagation delay is difficult even though these are for simple gates. One of the goals of this paper is the analytical evaluation of the propagation delay in a CMOS inverter. To do this, analytical expressions of the output waveform are derived, directly from the differential equation describing the temporal evolution of the inverter output. It is important to have an accurate model for the CMOS inverter operation, since several fast methods for reducing a CMOS gate to an equivalent inverter have been proposed [7], [8]. By using these so-called "collapsing" techniques the output response and the propagation delay of a gate can be computed quickly and accurately without the complications associated with trying to generalize the inverter-based model to complex gates.

The first closed-form delay expression based on the output response which was obtained directly from the differential equation describing the CMOS inverter operation was derived in [9] for a step input. Analytical expressions for the output waveform and the propagation delay, including the effect of the input waveform slope, were presented by Hedenstierna and Jeppson [10]. In this the influences of the short-circuit current and the gate-to-drain coupling capacitance were neglected. These output waveform expressions was extended by Kayssi *et al.* [11] for the case of exponential input waveform. More recently in [12], the differential equation describing the discharge of the load capacitor was solved for a rising input ramp considering the current through both transistors and the coupling capacitance. However, in the case where the PMOS device is in the linear region, the quadratic term of the current through the PMOS device was neglected. Moreover, it was not mentioned how the integration constant between the linear and the saturation regions of the PMOS device is calculated for fast inputs. For slow inputs least-square fitting techniques are used. Vemuru and Thorbjornsen [13] derived an expression for the output waveform, which includes the previously mentioned quadratic term of the PMOS current but ignored the influence of the coupling capacitance. A power series was used to approximate the solution of the differential equation. However, only the first five terms of the series were considered, and a recursion form for the calculation of higher order terms in order to obtain better accuracy was not given.

Sakurai and Newton [14], [15] presented closed-form delay expressions for the CMOS inverter, based on the α-power (n-power in [15]) law MOS model which includes the carriers'

Manuscript received March 22, 1996; revised January 29, 1997. This paper was recommended by Associate Editor T. Noll.

L. Bisdounis and O. Koufopavlou are with VLSI Design Laboratory, Department of Electrical and Computer Engineering, University of Patras, 26500 Patras, Greece.

S. Nikolaidis is with the Electronics and Computers Division, Department of Physics, Aristotle University of Thessaloniki, 54006 Thessaloniki, Greece.

Publisher Item Identifier S 1057-7122(98)01410-X.

velocity saturation effect. However, this model requires the extraction of the empirical velocity saturation index (α or n) from the static device characteristics for each transistor width. For the derivation of the output expression in [14], both the short-circuit current and the coupling capacitance are neglected. In [15], in order to approximate the CMOS inverter by an NMOS circuit, a fictitious input ramp is used which is clamped to ground for ramp voltages less than the switching voltage. This approximation is exact only for extreme cases of input ramps. An extension in the delay expression of [14] for the case of very lightly loaded inverter and/or slow input signals is presented in [16]. In this, a table of coefficients produced from SPICE simulations is used but still neglecting the short-circuit current. The delay model presented in [17] uses the α-power MOS model taking into account the short-circuit current of the inverter, but the output voltage and the currents through both transistors are assumed to be piecewise linear.

In this paper analytical expressions for the CMOS inverter output response to an input voltage ramp are derived. The proposed method overcomes the deficiencies of previous works. Based on these derived expressions, accurate analytical formulas for the evaluation of the propagation delay and the output transition time of the inverter for all the cases of input ramp slopes are produced. The derived timing model takes into account the complete expression of the short-circuit current, and the input–output coupling capacitance. This is achieved without using empirical approaches based on simulation results or approximations for the transistor currents as in previous works. The simplified bulk-charge MOS model [18] has been chosen. However, the experience derived from the results using this model could be expanded to more accurate and complex models.

The second goal of this paper is the derivation of an analytical expression for the CMOS short-circuit power dissipation. This is very important because the growing demand for low-power portable systems has made power dissipation a critical parameter in chip design [19]. 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. The first work on the evaluation of the short-circuit power dissipation was presented in [20]. A zero load capacitance and current waveform which is mirror symmetric about a central vertical axis (at the half of the input transition time) were assumed. Also, it is considered that the transistor, which is switched from cutoff to saturation, remains in saturation during the entire time when short-circuit current is conducted. More recently, in [10] and [21] an expression for the short-circuit energy dissipation of the CMOS inverter without the simplifications of [20] was derived. However, as mentioned above the expression of the output waveform, was derived without consideration of the short-circuit current and the gate-to-drain coupling capacitance. A closed-form expression for the evaluation of the short-circuit power dissipation based on an expression for the output waveform which considers the current through both transistors was presented in [22]. Sakurai and Newton [14] presented a formula for the short-circuit energy dissipation during one switching cycle which is a direct extension of the formula presented in [20]. The only difference is the use of the α-power law MOS model instead of the square-law MOS model. Recently, in [23] a substitution of the input transition time as given in [14] into the formula for the short-circuit dissipation also presented in [14], was proposed. This results in an expression for the short-circuit dissipation including the load capacitance which is not agreed with the initial assumption of zero load capacitance. Vemuru and Scheinberg [24] proposed a formula for the evaluation of the short-circuit power dissipation based on the α-power MOS model. In this work, the expression of the output waveform does not include the influences of the short-circuit current and the gate-to-drain capacitive coupling. A formulation of the short-circuit power dissipation through an equivalent short-circuit capacitance is presented in [25], where a rough linear approximation of the output waveform is used. Recently, in [26] the short-circuit current waveform was approximated with a piecewise linear function of time, in order to estimate the short-circuit energy dissipation. However, the energy of the reverse current due to the gate-to-drain coupling capacitance is subtracted from the short-circuit energy dissipation resulting in an underestimation.

In this paper, a formula for the evaluation of the short-circuit power dissipation for the CMOS inverter, based on analytical expressions of the output waveform is derived. It takes into account the currents through both transistors without making simplifying assumptions. In order to achieve better accuracy and to avoid an overestimation of the short-circuit power dissipation, the influence of the gate-to-drain coupling capacitance is considered. The derived expression clearly shows the influences of the inverter design characteristics, the load capacitance and the slope of the input waveform driving the inverter on the short-circuit power dissipation.

The rest of the paper is organized as follows. In Section II, analytical expressions of the CMOS inverter output waveform for all the cases of input voltage ramps, are derived. Also, in this section a detailed analysis of all the inverter operation regions is given. Closed-form expressions, results and comparisons with SPICE simulations and previous works of the propagation delay and the output transition time are given in Sections III and IV, respectively. Our approach for the evaluation of the CMOS short-circuit power dissipation, results and comparison with previous works are presented in Section V. Finally, we conclude in Section VI.

II. Inverter Output Waveform Analysis

The following derivations presented are for a rising input ramp:

$$V_{\text{in}} = \begin{cases} 0, & t \le 0 \\ V_{\text{DD}} \cdot (t/\tau), & 0 < t \le \tau \\ V_{\text{DD}}, & t > \tau \end{cases} \tag{1}$$

where τ is the input rise time. The analysis for a falling input ramp is similar. Taking into account the gate-to-drain capacitive coupling (C_M), the differential equation which describes the discharge of the load capacitance C_L for the CMOS inverter (Fig. 1), is derived from the application of the

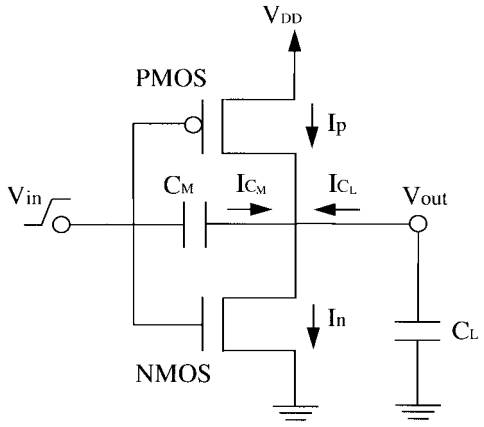

Fig. 1. The CMOS inverter.

Kirchoff's current law at the output node

$$I_{C_L} + I_{C_M} + I_p - I_n = 0$$
$$-C_L\frac{dV_{\text{out}}}{dt} + C_M\left(\frac{dV_{\text{in}}}{dt} - \frac{dV_{\text{out}}}{dt}\right) + I_p - I_n = 0. \quad (2)$$

Consequently, for the rising input ramp of (1), we get

$$\frac{dV_{\text{out}}}{dt} = \begin{cases} \frac{I_p - I_n}{C_L + C_M}, & t \le 0 \text{ or } t > \tau \\ \frac{c_m V_{\text{DD}}}{\tau} + \frac{I_p - I_n}{C_L + C_M}, & 0 < t \le \tau \end{cases} \quad (3)$$

where $c_m = \frac{C_M}{C_L + C_M}$.

The output load consists of the inverter drain junction capacitances, the gate capacitances of fanout gates and the interconnect capacitance. The equivalent gate-to-drain capacitance C_M is the sum of the gate-to-drain capacitances of both transistors

$$C_M = C_{\text{gd-NMOS}} + C_{\text{gd-PMOS}}.$$

The gate-to-drain capacitance of a transistor is the sum of the gate-to-drain overlap capacitance and a part of the gate-to-channel capacitance [27]. The overlap capacitance is voltage independent and is given by

$$C_{\text{gd-overlap}} = WC_{\text{dgo}}$$

where W is the effective width of the transistor and C_{gdo} is the gate-to-drain overlap capacitance per unit channel width which is determined by the process technology. In the cutoff region of the transistor there is no conducting channel and in the saturation region the channel does not extend to the drain. Therefore, the gate-to-drain capacitance due to the channel charge is equal to zero. In the linear region the distributed gate-to-channel capacitance may be viewed as being shared equally between the source and the drain. Thus, in this case

$$C_{\text{gd-channel}} = \frac{1}{2}C_{\text{ox}}WL$$

where C_{ox} is the gate–oxide capacitance per unit area and L is the effective length of the transistor.

Depending on the region of operation the NMOS device current using the simplified bulk-charge MOS model [18] is given by the following equations:

$$I_n = 0, \quad V_{\text{in}} < V_{\text{TN}}, \quad \text{Cutoff region} \quad (4)$$
$$I_n = \frac{\beta_n}{2(1+\delta_n)}(V_{\text{in}} - V_{\text{TN}})^2,$$
$$V_{\text{out}} > V_{\text{satn}}, \quad \text{Saturation region} \quad (5)$$
$$I_n = \beta_n\left[(V_{\text{in}} - V_{\text{TN}})V_{\text{out}} - \frac{(1+\delta_n)}{2}V_{\text{out}}^2\right],$$
$$V_{\text{out}} \le V_{\text{satn}}, \quad \text{Linear region} \quad (6)$$

where β_n is the NMOS device gain factor, V_{TN} is the NMOS device threshold voltage, δ_n is the slope of the first order term at the Taylor series expansion of the NMOS bulk-charge equation and $V_{\text{satn}} = \frac{V_{\text{in}} - V_{\text{TN}}}{1+\delta_n}$ is the NMOS saturation voltage.

The current equations for the PMOS device are

$$I_p = \beta_p\Big[(V_{\text{DD}} - V_{\text{in}} - |V_{\text{TP}}|)(V_{\text{DD}} - V_{\text{out}}) - \frac{(1+\delta_p)}{2}(V_{\text{DD}} - V_{\text{out}})^2\Big],$$
$$V_{\text{out}} \ge V_{\text{satp}}, \quad \text{Linear region} \quad (7)$$
$$I_p = \frac{\beta_p}{2(1+\delta_p)}(V_{\text{DD}} - V_{\text{in}} - |V_{\text{TP}}|)^2,$$
$$V_{\text{out}} < V_{\text{satp}}, \quad \text{Saturation region} \quad (8)$$
$$I_p = 0, \quad V_{\text{in}} > V_{\text{DD}} - |V_{\text{TP}}|, \quad \text{Cutoff region} \quad (9)$$

where β_p is the PMOS device gain factor, V_{TP} is the PMOS device threshold voltage, δ_p is the slope of the first order term at the Taylor series expansion of the PMOS bulk–charge equation and $V_{\text{satp}} = V_{\text{DD}} - \frac{(V_{\text{DD}} - V_{\text{in}} - |V_{\text{TP}}|)}{1+\delta_p}$ is the PMOS saturation voltage. The above expressions are similar to those of the square-law MOS model [28] but have the factor $(1+\delta_{n(p)})$, which reduces the current to a more accurate value.

In order to give a complete analysis, four cases of input ramps are considered. First, the case of *very fast* input ramps where the PMOS device is turned off after its linear region, without entering saturation, is studied. Since the input ramp will reach its final value with the NMOS device either in saturation or in the linear region, two more cases of input ramps are considered. For *fast* input ramps, the NMOS device is still saturated while for *slow* input ramps the NMOS is in its linear region, when the input voltage ramp reaches its final value. Finally, the case of *very slow* input ramps where the PMOS is turned off when the NMOS is in its linear region, is examined.

In the following, normalized voltages with respect to V_{DD}, i.e., $u_{\text{in}} = V_{\text{in}}/V_{\text{DD}}$, $u_{\text{out}} = V_{\text{out}}/V_{\text{DD}}$, $n = V_{\text{TN}}/V_{\text{DD}}$, $p = |V_{\text{TP}}|/V_{\text{DD}}$, $u_{\text{satn}} = V_{\text{satn}}/V_{\text{DD}}$, $u_{\text{satp}} = V_{\text{satp}}/V_{\text{DD}}$, and the variable $x = t/\tau$, are used.

Case A: The first case to be studied is for very fast input ramps such that the PMOS transistor is turned off after its linear region, without entering saturation (Fig. 2). Also, the NMOS transistor is still saturated when the input voltage reaches its final value.

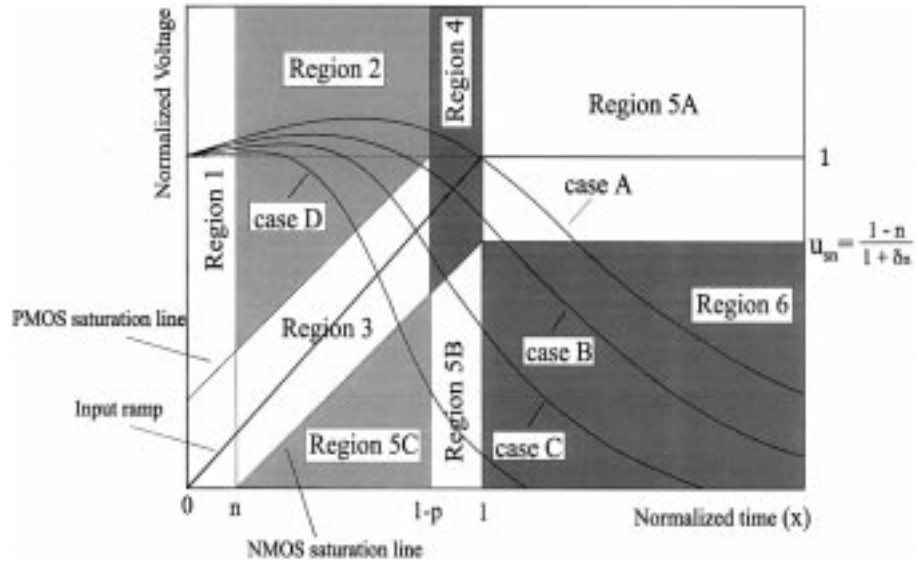

Fig. 2. Operation regions of the inverter for very fast, fast, slow, and very slow input ramps.

Region 1—$0 \le x \le n$: The NMOS transistor is off, and the PMOS transistor is in the linear region. The first term of the right-hand side of (3) (for $0 < t \le \tau$) corresponds to the charging current through the coupling capacitance (C_M) which causes the major influence on the output voltage waveform in this region. Part of the charge from the input which injected through this capacitance causes an overshoot at the early part of the output voltage waveform (Fig. 2). During the overshoot the PMOS device operates in a reversed linear mode because the output voltage is greater than the supply voltage. Thus the PMOS device initially helps to discharge the load capacitance toward the supply voltage. The differential equation (3) using (4), (7) becomes a nonlinear Riccati equation [29] which cannot be solved analytically, if a particular solution is not known. Thus, a power-series expansion method [30], [31] has been used, resulting to the following recursive expression:

$$u_{\text{out}} = 1 - \sum_{k=1}^{\infty} f_k x^k \tag{10}$$

where

$$f_1 = -c_m, \quad f_2 = -\frac{A_p}{2}(1-p)f_1$$

and

$$f_k = \frac{A_p}{k}\left\{[f_{k-2} - (1-p)f_{k-1}] + \frac{(1+\delta_p)}{2}\sum_{i=1}^{k-2} f_i f_{k-i-1}\right\} \quad \text{for } k > 2,\ A_p = \tfrac{\beta_p V_{\text{DD}}\tau}{C_L + C_M}.$$

The second term in the braces for $k > 2$ corresponds to the influence of the quadratic current term of the PMOS device which was neglected in [12]. A satisfactory limit to truncate the above series is obtained for $k = 8$. The error of truncating the series is not more than 0.01%.

Region 2—$n \le x \le 1 - p$: The NMOS transistor is saturated and the PMOS transistor is in the linear region. During the output voltage overshoot the PMOS device still operates in a reversed linear mode. As in region 1 the differential equation (3) using (5), (7) becomes a Riccati equation. In this case, the power-series method results to the following recursive expression:

$$u_{\text{out}} = u_{12} + 1 - \sum_{k=1}^{\infty} g_k (x-n)^k \tag{11}$$

where

$$g_1 = -c_m, \quad g_2 = -\frac{A_p}{2}(1-p-n),$$
$$g_3 = \frac{A_n}{6(1+\delta_n)} + \frac{A_p}{3}\left[g_1 - (1-p-n)g_2 + \frac{(1+\delta_p)}{2}g_1^2\right],$$
$$g_k = \frac{A_p}{k}\left\{[g_{k-2} - (1-p-n)g_{k-1}] + \frac{(1+\delta_p)}{2}\sum_{i=1}^{k-2} g_i g_{k-i-1}\right\} \quad \text{for } k > 3, \quad A_n = \tfrac{\beta_n V_{\text{DD}}\tau}{C_L + C_M}.$$

$u_{12} = -\sum_{k=1}^{\infty} f_k n^k$ is the integration constant which is inserted to ensure continuity with respect to region 1. The influence of the quadratic current term of the PMOS device is inserted to the output waveform at the third coefficient of the series. The truncating error of the above series at $k = 10$ is not more than 0.02%. As shown below, the series output expressions of regions 1 and 2 give waveforms very close to those derived from SPICE simulations, which indicates their validity.

Region 4—$1 - p \le x \le 1$: The NMOS transistor is saturated and the PMOS transistor is off. It can be observed in Fig. 2 that for very fast input ramps (case A), the inverter doesn't pass from region 3 because the PMOS device is not

saturated. The analytical solution of the differential equation (3) in this region is

$$u_{\rm out} = u_{24} + c_m x - \frac{A_n}{6(1+\delta_n)}(x-n)^3. \tag{12}$$

The integration constant u_{24} is inserted to ensure continuity with respect to region 2 and is given by

$$u_{24} = u_{[1-p]} - c_m(1-p) + \frac{A_n}{6(1+\delta_n)}(1-p-n)^3.$$

$u_{[1-p]}$ is the value of the output voltage in which the PMOS device is turned off and is calculated from (11) for $x = 1-p$.

Region 5A— $1 \le x \le x_{\rm sn}$*:* The input ramp has reached its final value with the NMOS device still in saturation and the PMOS device off. $x_{\rm sn}$ is the normalized time value where the NMOS device leaves saturation, i.e., $V_{\rm out} \le V_{\rm satn}$ (see NMOS current equations). The analytical solution of the differential equation (3) (for $t > \tau$) is

$$u_{\rm out} = u_{24} + c_m - \frac{A_n(1-n)^3}{6(1+\delta_n)} - \frac{A_n(1-n)^2}{2(1+\delta_n)}(x-1). \tag{13}$$

Region 6—$x \ge x_{\rm sn}$*:* The NMOS device is entering its linear region and the PMOS is off. The analytical solution of (3) is

$$u_{\rm out} = \frac{2u_{\rm sn}}{1 + e^{A_n(x-x_{\rm sn})(1-n)}} \tag{14}$$

where $u_{\rm sn} = (1-n)/(1+\delta_n)$, and $x_{\rm sn}$ is calculated from (13) for $u_{\rm out} = u_{\rm sn}$.

Case B: The second case studies fast input ramps. The PMOS transistor is entering the saturation after the linear region and the NMOS transistor is still saturated when the input ramp reaches its final value (Fig. 2). The expressions of the output waveform for regions 1 and 2 are the same with those of case A. Note, that the right limit of region 2 in this case is the normalized time value $(x_{\rm sp})$ where the PMOS device is entering the saturation region, i.e., $V_{\rm out} < V_{\rm satp}$ (see PMOS current equations). It is determined by the PMOS saturation condition

$$u_{\rm out} = u_{\rm satp} = 1 - \frac{1-x-p}{1+\delta_p}. \tag{15}$$

Region 3—$x_{\rm sp} \le x \le 1-p$*:* Both transistors are saturated. The analytical solution of the differential equation (3) is

$$u_{\rm out} = u_{23} + c_m x - \frac{A_n}{6(1+\delta_n)}(x-n)^3 - \frac{A_p}{6(1+\delta_p)}(1-x-p)^3. \tag{16}$$

The integration constant u_{23} is inserted to ensure continuity with respect to region 2 and is given by

$$u_{23} = u_{\rm sp} - c_m x_{\rm sp} + \frac{A_n}{6(1+\delta_n)}(x_{\rm sp}-n)^3 + \frac{A_p}{6(1+\delta_p)}(1-x_{\rm sp}-p)^3$$

where $u_{\rm sp}$ is the value of the normalized output voltage when the PMOS device is entering the saturation region. For the derivation of an analytical output waveform expression and the calculation of the propagation delay, the integration constant u_{23} must be determined. To achieve this, the calculation of the values $x_{\rm sp}$ and $u_{\rm sp}$ is required. These values satisfy the PMOS saturation condition, expressed by (15), and they can be calculated by solving the system of (11) and (15). Since, the order of (11) is high, the system of these equations cannot be solved analytically. Hence, in the following an efficient method for the calculation of $x_{\rm sp}$ and $u_{\rm sp}$ is introduced, which is illustrated in Fig. 3.

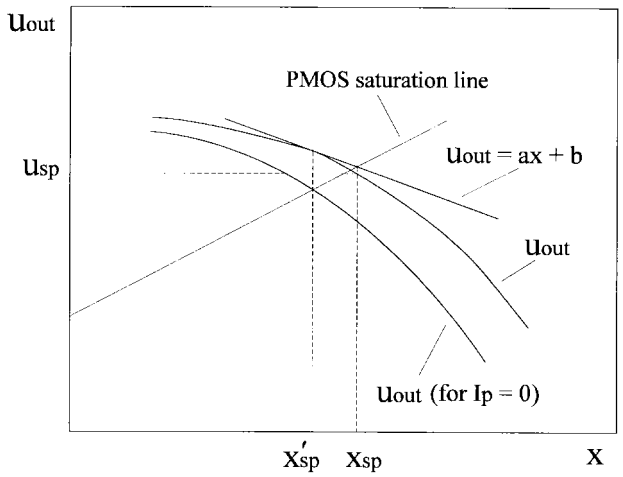

Fig. 3. Approximation of the normalized time $x_{\rm sp}$ when the inverter is entering region 3.

The analytical solution of the differential equation describing the discharge of the load capacitance in region 2, if negligible PMOS current is assumed, is given by

$$u_{\rm out} = u'_{12} + c_m x - \frac{A_n}{6(1+\delta_n)}(x-n)^3 \tag{17}$$

where

$$u'_{12} = 1 - c_m n - \sum_{k=1}^{\infty} f_k n^k.$$

By equating (15) and (17) the normalized time $x'_{\rm sp}$ in which the inverter is entering region 3 with the assumption of negligible PMOS current becomes the root of a cubic equation which belongs to the interval $[n, 1-p]$. The next step is the determination of the tangent of the output waveform expressed by (11), at the point which corresponds to $x'_{\rm sp}$ (Fig. 3). This tangent is expressed by the equation

$$u_{\rm out} = ax + b \tag{18}$$

where

$$a = \left.\frac{du_{\rm out}}{dx}\right|_{x=x'_{\rm sp}} = -\sum_{k=1}^{\infty} k g_k (x'_{\rm sp} - n)^{k-1}$$

and

$$b = 1 + u_{12} - a x'_{\rm sp} - \sum_{k=1}^{\infty} g_k (x'_{\rm sp} - n)^k.$$

From (15) and (18) an accurate approximation for $x_{\rm sp}$ is

$$x_{\rm sp} = \frac{(1+\delta_p)b - \delta_p - p}{1 - a(1+\delta_p)}. \tag{19}$$

By substituting $x_{\rm sp}$ in (11) the normalized output voltage $u_{\rm sp}$ is evaluated. The error which is introduced in the calculation of $x_{\rm sp}$ due to the above method is up to 0.08%. The expressions of the output waveform for the regions 4, 5A, and 6 are the same with those of Case A, if the constant u_{24} is substituted by the constant u_{23}.

Case C: In the third case, slow input ramps are studied (Fig. 2). The NMOS device leaves saturation while the input voltage is still a ramp. The output expressions for the regions 1, 2, and 3 are the same with those of the previous case.

Region 4—$1-p \le x \le x_{\rm sn}$*:* The NMOS transistor is saturated and the PMOS transistor is off. The solution of the differential equation which describes the temporal evolution of the inverter output in this region is given in (12) by substituting the constant u_{24} with u_{23}. The normalized time value $x_{\rm sn}$ is calculated from this equation for $u_{\rm out} = u_{\rm satn} = (x-n)/(1+\delta_n)$, which corresponds to the NMOS saturation line (Fig. 2).

Region 5B—$x_{\rm sn} \le x \le 1$*:* The NMOS transistor is in the linear region and the PMOS transistor is off. Neglecting the charging current through the coupling capacitance an approximated solution of the differential equation (3) is

$$u_{\rm out} = \frac{1}{(1+\delta_n)\sqrt{\frac{A_n}{2}}e^{y^2}\left[\frac{1}{y_{\rm sn}e^{y_{\rm sn}^2}} - \frac{\sqrt{\pi}}{2}(\mathrm{erf}[y] - \mathrm{erf}[y_{\rm sn}])\right]} \tag{20}$$

where $y = \sqrt{\frac{A_n}{2}}(x-n)$, $y_{\rm sn} = \sqrt{\frac{A_n}{2}}(x_{\rm sn}-n)\cdot \mathrm{erf}[y]$ and $\mathrm{erf}[y_{\rm sn}]$ are the error functions of y and $y_{\rm sn}$ respectively. Standard ways for the evaluation of the error function can be found in most mathematical handbooks [32].

Region 6—$x \ge 1$*:* The input ramp has reached its final value, the NMOS device is still in the linear region and the PMOS device is off. The differential equation which describes the operation of the inverter in this region is the same with those of Case A with different initial conditions. Its analytical solution is

$$u_{\rm out} = \frac{2u_{\rm sn}}{1 + \frac{2u_{\rm sn}-u_{[1]}}{u_{[1]}}e^{A_n(x-1)(1-n)}} \tag{21}$$

where $u_{\rm sn} = (1-n)/(1+\delta_n)$ and $u_{[1]}$ is the value of the normalized output voltage when the input ramp reaches its final value. $u_{[1]}$ is calculated if $x=1$ is set in (20).

Case D: In this case, very slow input ramps are studied. The PMOS device is turned off when the NMOS device is in its linear region. After region 3 the inverter enters directly to region 5C where the NMOS device is in its linear region and the PMOS device is saturated. The output waveform expressions for the regions 1, 2 and 3 are the same with those of the previous case.

Regions 5C and 5B— $x_{\rm sn} \le x \le 1$*:* The NMOS device is in its linear region for both regions. In region 5B ($1-p \le x \le 1$), the PMOS device is off, while in region 5C ($x_{\rm sn} \le x \le 1-p$) is saturated. The differential equation which describes the inverter operation, starting from (3), in both regions is

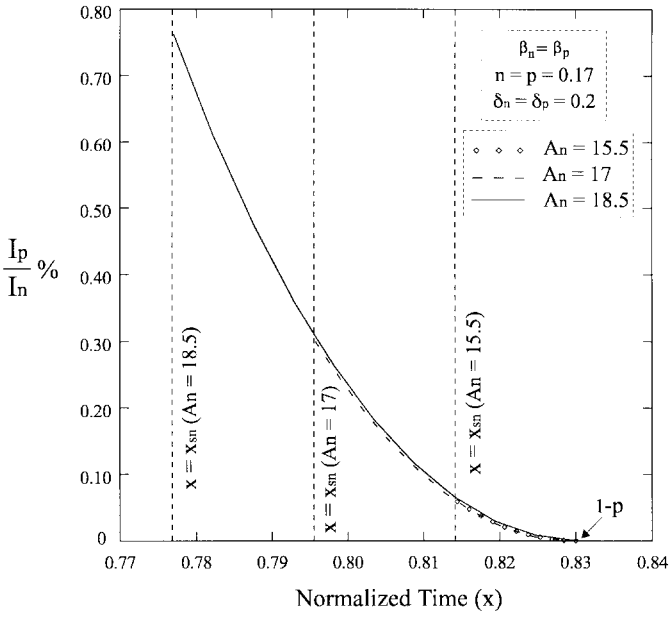

Fig. 4. Comparison of PMOS and NMOS device currents in region 5C (Case D).

$$\frac{du_{\rm out}}{dx} = c_m - A_n\left[(x-n)u_{\rm out} - \frac{(1+\delta_n)}{2}u_{\rm out}^2\right] + \frac{A_p}{2(1+\delta_p)}(1-x-p)^2 H(1-x-p) \tag{22}$$

where $H(1-x-p)$ is the Heaviside's step function ($H(1-x-p)=0$ for $x>1-p$ and $H(1-x-p)=1$ for $x \le 1-p$). In region 5C the PMOS device is poorly conducting, thus its influence can be neglected. In Fig. 4 is shown that for $A_n = A_p = 15.5$ the PMOS current is up to 0.06% of the NMOS current. As in Case C (region 5B) the charging current through the coupling capacitance is neglected and the solution of the above differential equation is given by (20). In this case the normalized time value $x_{\rm sn}$ is calculated from (16), for $u_{\rm out} = u_{\rm satn} = (x-n)/(1+\delta_n)$. The expression for the output waveform in region 6 is given by (21).

Typical inverter output waveforms from the above derived equations, are shown in Fig. 5. The results have been obtained for an inverter with equal NMOS and PMOS gain factors $\beta_n = \beta_p = 0.5$ mA/V^2, $n = p = 0.17$ and $\delta_n = \delta_p = 0.2$, operating at $V_{\rm DD} = 5$ V with an output load of 0.5 pF. In order to give output waveforms for several values of $A_{\rm no}(A_{\rm no} = \beta_n V_{\rm DD}\tau/C_L)$ in the same diagram, the normalized output voltage is plotted as a function of the normalized time ($x = t/\tau$). The output waveforms produced from long-channel level 3 SPICE simulations are added for comparison. It can be observed that the analytical waveforms are very close to those produced from SPICE simulations. The output waveforms for $A_{\rm no} \le 2$ correspond to Case A, the output waveforms for $2 < A_{\rm no} \le 5$ to Case B, the output waveforms for $5 < A_{\rm no} \le 14$ to Case C and those for $A_{\rm no} > 14$ to Case D.

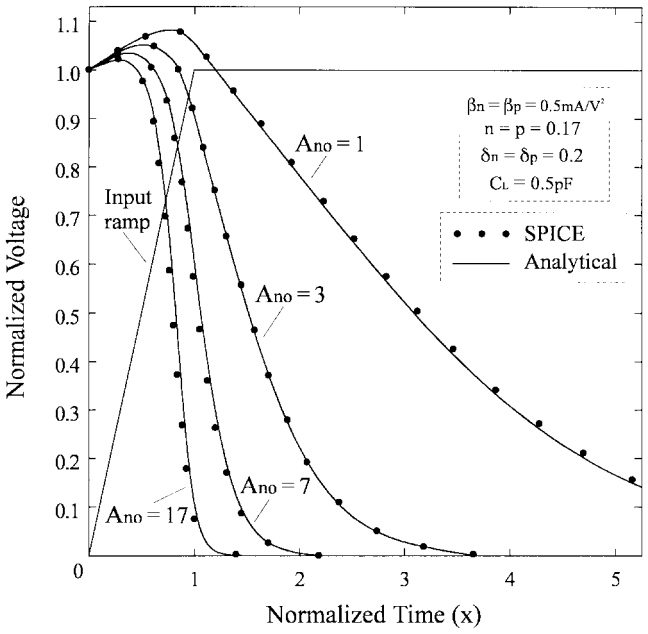

Fig. 5. Inverter output waveforms derived from the proposed analytical method and from SPICE simulations.

III. Propagation Delay

The fall propagation delay at the 50% voltage level is written as

$$t_{\mathrm{DHL}} = t_{0.5} - \frac{\tau}{2} = x_{0.5}\tau - \frac{\tau}{2} \tag{23}$$

where $x_{0.5}$ is the normalized time value when $u_{\mathrm{out}} = 0.5$. Thus, for the evaluation of the propagation delay, the normalized time value $x_{0.5}$ must be determined for the four cases of input ramps.

Cases A and B: In the cases of very fast and fast input ramps the output voltage reaches the 50% voltage level $(u_{\mathrm{out}} = 0.5)$ when the inverter operates in region 6. The normalized time value $x_{0.5}$ is calculated from (14) for $u_{\mathrm{out}} = 0.5$

$$x_{0.5} = x_{\mathrm{sn}} + \frac{\ell n[4u_{\mathrm{sn}} - 1]}{A_n(1-n)}. \tag{24}$$

Case C: In the case of slow input ramps the condition $u_{\mathrm{out}} = 0.5$ occurs in region 6 when $5 < A_{\mathrm{no}} \leq 8$ or in region 5B when $8 < A_{\mathrm{no}} \leq 14$. In the first case the normalized time value $x_{0.5}$ is calculated from (21) for $u_{\mathrm{out}} = 0.5$

$$x_{0.5} = 1 + \frac{\ell n\left[\frac{(4u_{\mathrm{sn}}-1)u_{[1]}}{2u_{\mathrm{sn}}-u_{[1]}}\right]}{A_n(1-n)}. \tag{25}$$

For the evaluation of $x_{0.5}$ in the case where the output voltage reaches the 50% level in region 5B a linear approximation of the output voltage is used in the vicinity of the 50% voltage level, since the expression of the output waveform in region 5B (20) cannot be solved analytically. Then

$$x_{0.5} = x_{\mathrm{sn}} + \frac{0.5 - [(x_{\mathrm{sn}} - n)/(1+\delta_n)]}{d} \tag{26}$$

where $d = \frac{du_{\mathrm{out}}}{dx}|_{x=x_{\mathrm{sn}}} = -\frac{A_n(x_{\mathrm{sn}}-n)^2}{2(1+\delta_n)}$ is the slope of the output waveform.

Case D: In the case of very slow input ramps the condition $u_{\mathrm{out}} = 0.5$ can occur in three possible regions. For $14 < A_{\mathrm{no}} \leq 18$ occurs in region 5B or in region 5C. In this case the normalized time value $x_{0.5}$ is given by (26). For $A_{\mathrm{no}} > 18$, $u_{\mathrm{out}} = 0.5$ occurs in region 3 and $x_{0.5}$ is calculated from (16) for $u_{\mathrm{out}} = 0.5$

$$x_{0.5} = 2\sqrt{Q}\cos\left[\frac{\theta + r\pi}{3}\right] - \frac{k_1}{3} \tag{27}$$

where

$$Q = \frac{k_1^2 - 3k_3}{9}, \quad \theta = \cos^{-1}\left[\frac{R}{\sqrt{Q^3}}\right]$$

$$R = \frac{9k_1k_2 - 27k_3 - 2k_1^3}{54}$$

$$k_1 = \frac{3[nB_n - (1-p)B_p]}{B_p - B_n}$$

$$k_2 = \frac{3[(1-p)^2B_p - n^2B_n + c_m]}{B_p - B_n}$$

$$k_3 = \frac{u_{23} + n^3B_n - (1-p)^3B_p - 0.5}{B_p - B_n}$$

$$B_n = \frac{A_n}{6(1+\delta_n)}, \quad B_p = \frac{A_p}{6(1+\delta_p)}.$$

r is equal to 0 or 4 if $A_n > A_p$ or $A_n < A_p$, respectively. When $A_n = A_p$ and $\delta_n = \delta_p$, $x_{0.5}$ becomes the solution of a simple quadratic equation.

By substituting $x_{0.5}$ from one of the equations (24), (25), (26), and (27) in (23) the fall propagation delay of the inverter for all the cases of input ramps can be evaluated analytically. The error which is introduced in the evaluation of the propagation delay due to the approximation in the calculation of $x_{0.5}$ in regions 5B and 5C [see (26)] is up to 0.6%.

In Fig. 6, the fall propagation delay of the inverter is plotted as a function of A_{no}. The results have been produced for an inverter with equal NMOS and PMOS device gain factors $\beta_n = \beta_p = 0.5$ mA/V^2, $n = p = 0.17$ and $\delta_n = \delta_p = 0.2$, operating at $V_{\mathrm{DD}} = 5$ V, with an output load of 0.5 pF, and input rise time from 0.2 ns $(A_{\mathrm{no}} = 1)$ to 4 ns $(A_{\mathrm{no}} = 20)$. Results using the approaches for the evaluation of the propagation delay presented in [10], [14] and [17], are also given. It can be observed, that the presented analytical method for the evaluation of the inverter propagation delay gives results closer to those derived from long-channel level 3 SPICE simulations (indicated with diamonds) than the other methods.

The fall propagation delay as given by (23), for the case of equal device gain factors and the case of double PMOS device gain factor, is plotted as a function of A_{no} in Fig. 7. For relatively fast input ramps $(A_{\mathrm{no}} \leq 6)$ the delay is slightly shorter for wide PMOS devices than for narrow ones and higher for slower input ramps. This is due to the charging current through the gate-to-drain coupling capacitance, which causes a voltage overshoot at the early part of the output waveform. Initially, the PMOS device helps to discharge the load capacitance toward supply voltage. Then it causes

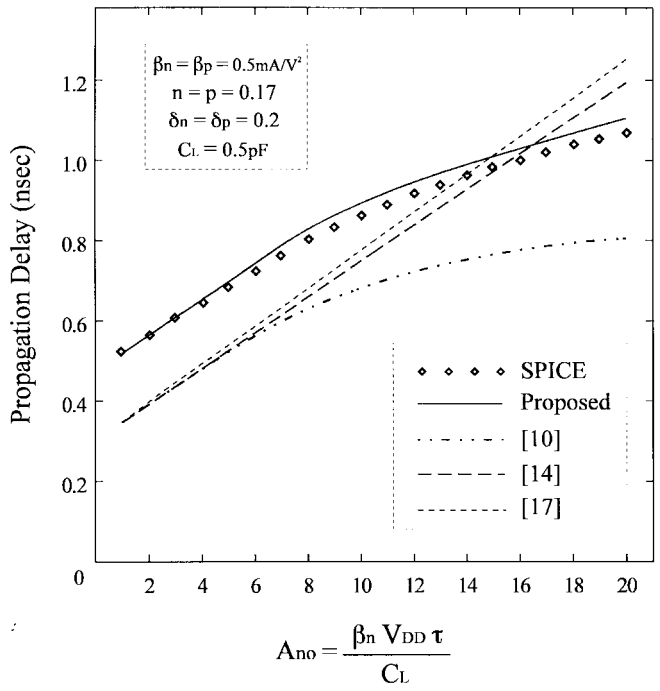

Fig. 6. Inverter fall propagation delay derived from the proposed method, previous models and SPICE simulations.

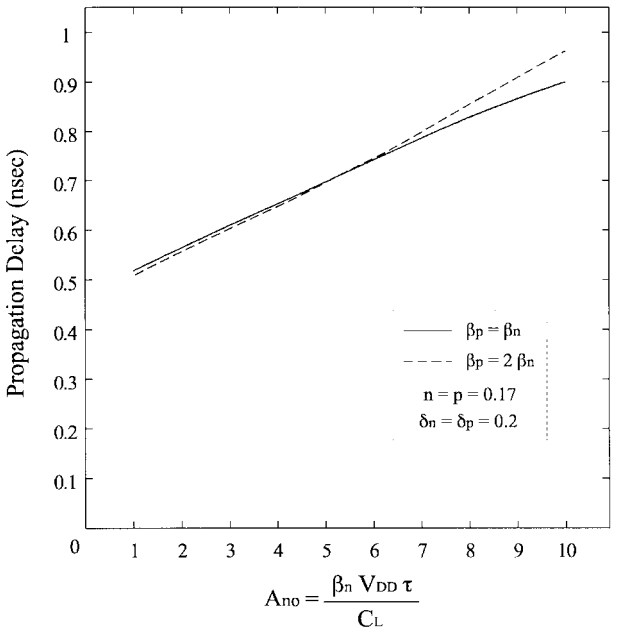

Fig. 7. Fall propagation delay of the CMOS inverter for $\beta_p = \beta_n$ and $\beta_p = 2\beta_n$.

an additional delay in the discharge of the load capacitance through the NMOS device. For fast input ramps the output voltage overshoot is high and the first effect dominates the second, resulting in a reduction of the propagation delay. On the contrary, for slower input ramps $(A_{\rm no} > 6)$ the output voltage overshoot is lower and the reaction in the discharge of the output node caused by the PMOS device current is dominant. This results in an increase of the propagation delay. Corresponding results for the rising output propagation delay are obtained for $\beta_n = 2\beta_p$ as, in this case, the NMOS transistor is the short-circuiting transistor.

IV. Effective Transition Time

In real CMOS datapaths, the input signal of a gate is not a ramp but the output waveform of the preceding gate. In order for the derived ramp delay model to be applicable to real circuits, an approximation of the real input waveform by a ramp waveform is needed, to obtain an effective transition time. According to [10], a good approximation for the evaluation of the effective output transition time $(\tau_{\rm tr})$ of the inverter, is achieved when the output waveform slope is approximated by 70% of its derivative at the point which corresponds to the half supply voltage level (Fig. 8). $\tau_{\rm tr}$ can be used as the τ for the succeeding inverter in the circuit. The effective output transition time may then be written as

$$\tau_{\rm tr} = \frac{V_{\rm DD}}{0.7\left|\frac{dV_{\rm out}}{dt}\right|_{t=t_{0.5}}} = \frac{\tau}{0.7s} \tag{28}$$

where $s = \frac{du_{\rm out}}{dx}\Big|_{x=x_{0.5}}$ is the derivative value of the output waveform at the point $x_{0.5}$. Thus, in order to evaluate the effective output transition time this derivative value must be determined for all the cases of input ramps.

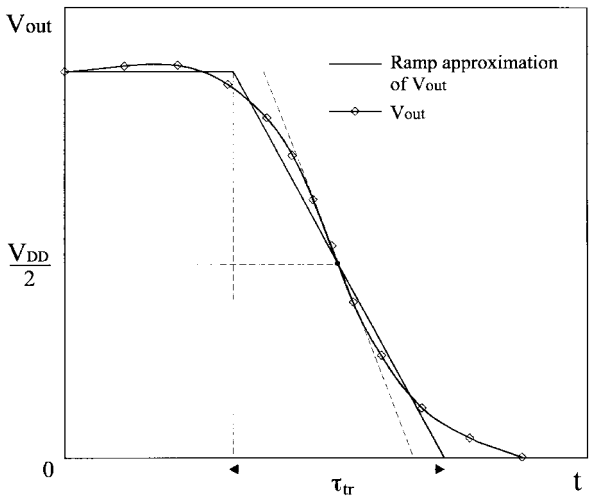

Fig. 8. Determination of the effective output transition time.

Cases A and B: In the cases of very fast and fast input ramps the derivative value of $u_{\rm out}$ at $x_{0.5}$ is evaluated by using (14)

$$s = -\frac{2u_{\rm sn}A_n(1-n)e^{A_n(x_{0.5}-x_{\rm sn})(1-n)}}{\left[1+e^{A_n(x_{0.5}-x_{\rm sn})(1-n)}\right]^2} \tag{29}$$

where $x_{0.5}$ is given by (24).

Case C: In the case of slow input ramps the condition $u_{\rm out} = 0.5$ can occur in region 6 or in region 5B. In the first case $(5 < A_{\rm no} \le 8)$ the derivative value of $u_{\rm out}$ at $x_{0.5}$ is evaluated by using (21)

$$s = -\frac{2u_{\rm sn}u_{[1]}(2u_{\rm sn}-u_{[1]})A_n(1-n)e^{A_n(x_{0.5}-1)(1-n)}}{\left[u_{[1]}+(2u_{\rm sn}-u_{[1]})e^{A_n(x_{0.5}-1)(1-n)}\right]^2} \tag{30}$$

where $x_{0.5}$ is given by (25). When $u_{\rm out} = 0.5$ occurs in region 5B $(8 < A_{\rm no} \le 14)$, the derivative value of $u_{\rm out}$ at $x_{0.5}$ is

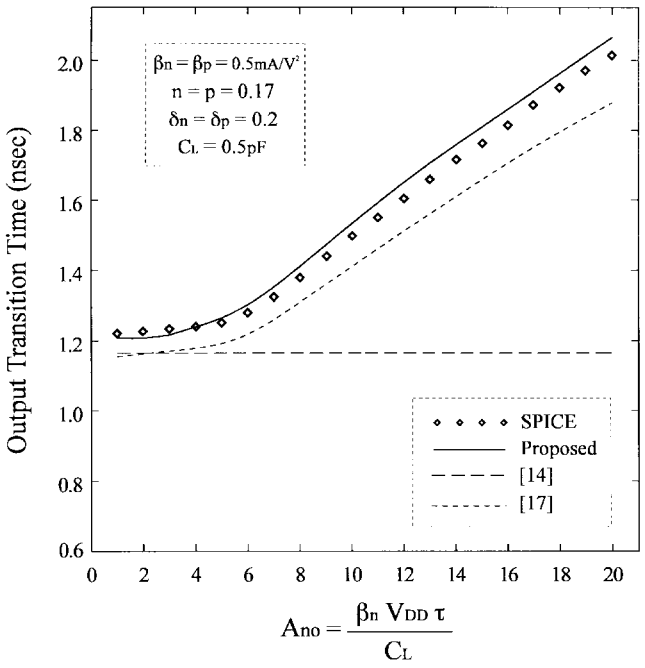

Fig. 9. Inverter output transition time derived from the proposed method, previous models, and SPICE simulations.

calculated by using (20)

$$s = \frac{1 - 2y_{0.5}e^{y_{0.5}^2}D}{(1+\delta_n)e^{2y_{0.5}^2}D^2} \tag{31}$$

where $y_{0.5} = \sqrt{\frac{A_n}{2}}(x_{0.5} - n)$, $D = \frac{1}{y_{\rm sn}e^{y_{\rm sn}^2}} - \frac{\sqrt{\pi}}{2}(\mathrm{erf}[y_{0.5}] - \mathrm{erf}[y_{\rm sn}])$ and $x_{0.5}$ is given by (26).

Case D: In the case of very slow input ramps the condition $u_{\rm out} = 0.5$ can occur in regions 5B, 5C and 3. For $14 < A_{\rm no} \leq 18$ the output voltage reaches the 50% level when the inverter operates in region 5B or in region 5C. In this case the derivative value of $u_{\rm out}$ at $x_{0.5}$ is given by (31). For $A_{\rm no} > 18$, $u_{\rm out} = 0.5$ occurs in region 3 and the derivative value of $u_{\rm out}$ at $x_{0.5}$ is evaluated by using (16)

$$s = c_m - \frac{A_n}{2(1+\delta_n)}(x_{0.5} - n)^2 + \frac{A_p}{2(1+\delta_p)}(1 - x_{0.5} - p)^2 \tag{32}$$

where $x_{0.5}$ is given by (27).

Finally, by substituting the derivative value of $u_{\rm out}$ at $x_{0.5}$ in (28) the effective output transition time of the inverter, for all the cases of input ramps, is evaluated. The effective transition time for a falling output waveform is plotted as a function of $A_{\rm no}$ in Fig. 9. The results have been obtained for an inverter with equal NMOS and PMOS device gain factors $\beta_n = \beta_p = 0.5$ mA/V^2, $n = p = 0.17$ and $\delta_n = \delta_p = 0.2$, operating at $V_{\rm DD} = 5$ V, with an output load of 0.5 pF, and input rise time from 0.2 ns ($A_{\rm no} = 1$) to 4 ns ($A_{\rm no} = 20$). Results using the approaches for the evaluation of the output transition time presented in [14] and [17], are also given. It can be observed, that the presented analytical method for the evaluation of the CMOS inverter output transition time gives results closer to those produced from SPICE simulations (indicated with diamonds) than the other methods. Note, that in the expression presented in [14], the output transition time remains constant with increasing values of input transition time. This approach is valid only for fast input ramps ($A_{\rm no} \leq 5$). This is explained considering that for fast input ramps the discharging current (current through the NMOS transistor) takes its maximum value very early and the PMOS transistor is poorly conducting. As the input transition time (and consequently $A_{\rm no}$) increases the rate of the increase of the discharging current is reduced. In addition the reaction of the PMOS current also increases slowing down the output transition.

TABLE I
PROPAGATION DELAY RESULTS FOR A CMOS INVERTER CHAIN

	Propagation Delay (nsec)	
	Analytical	SPICE
Inverter #1	0.7657	0.8016
Inverter #2	0.9115	0.9390
Inverter #3	1.0329	1.0591
Inverter #4	0.7132	0.7441
Chain	3.4233	3.5438

In the following, the propagation delay evaluation for the CMOS inverter chain of Fig. 10 is examined. The first inverter (the one outside the box) is only used for the derivation of a real input waveform to the inverter chain. The inverters of the chain have different loads and drives. In the delay evaluation the effective output transition time of each inverter is calculated by using (28) and is used as input transition time for the succeeding inverter in the chain. The results derived from the analytical model and those produced from SPICE simulations are presented in Table I. The propagation delay of each inverter and the total delay of the chain are given. The agreement between the simulation and the calculation is very good. The error in the total delay is 3.4%.

V. CMOS SHORT-CIRCUIT POWER DISSIPATION

In a CMOS inverter, short-circuit power is dissipated when a direct path from power supply to ground occurs. For a falling output transition, due to the charging current through the input-to-output coupling capacitance (C_M), an overshoot occurs at the early part of the output voltage waveform (Fig. 11). During the overshoot there is no current from power supply to ground because $V_{\rm out}$ is higher than $V_{\rm DD}$. Thus, as it can be seen in Fig. 11 (shaded region), short-circuit power is dissipated from the end of the output voltage overshoot ($x = x_1$) until the PMOS device is turned off ($x = 1 - p$). The short-circuit power dissipation is given by

$$P_{\rm SC} = (E_{\rm SCF} + E_{\rm SCR})f. \tag{33}$$

$E_{\rm SCF}$ and $E_{\rm SCR}$ are the short-circuit energy dissipation per falling and rising output transition, respectively and f is the switching frequency. In the following, the short-circuit power dissipation for all the cases of input ramps is evaluated. The complete analysis for a falling output transition and final

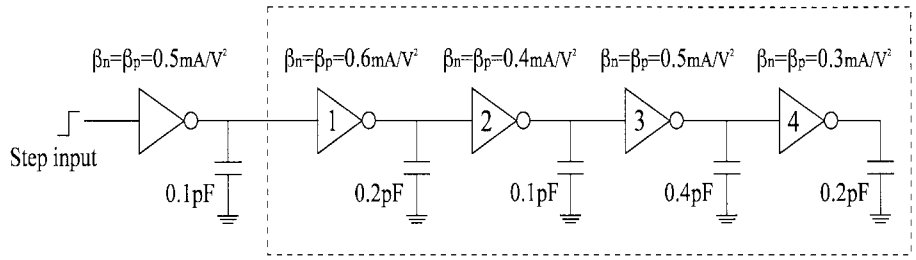

Fig. 10. Example of a CMOS inverter chain.

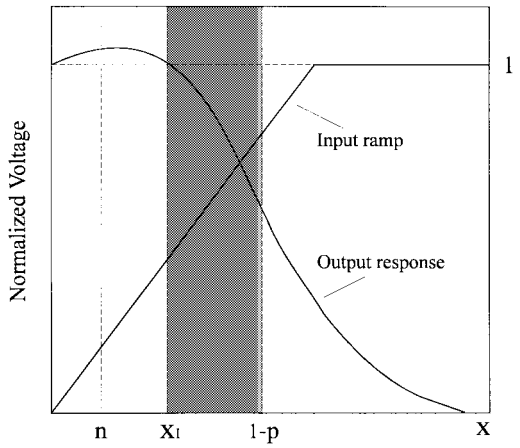

Fig. 11. Region in which short-circuit energy is dissipated $(x_1 \le x \le 1-p)$.

closed-form expressions for the short-circuit energy dissipation during a rising output transition, are given.

Case A: In this case, the PMOS device is turned off before the end of the overshoot at the output voltage waveform (Fig. 2). Thus, there is no short-circuit power dissipation.

Cases B and C: In both cases of input ramps, the NMOS device is still saturated when the input voltage ramp reaches the value $V_{\rm DD}-|V_{\rm TP}|$, i.e., $x=1-p$. The short-circuit energy dissipation during a falling output transition is given by

$$E_{\rm SCF}=V_{\rm DD}\int I_{\rm SC}\,dt=V_{\rm DD}\int_{x_1}^{1-p} I_p\tau\,dx. \tag{34}$$

The application of the Kirchoff's current law at the inverter output node (2) gives

$$I_p=\frac{\beta_n V_{\rm DD}^2}{2(1+\delta_n)}(x-n)^2-\frac{C_M V_{\rm DD}}{\tau}+\frac{(C_L+C_M)V_{\rm DD}}{\tau}\frac{du_{\rm out}}{dx}. \tag{35}$$

By substituting I_p in (34) the integration yields

$$E_{\rm SCF}=\frac{\beta_n V_{\rm DD}^3\tau}{6(1+\delta_n)}\left[3n(1-p-x_1)(n+p-x_1-1)+(1-p)^3-x_1^3\right]-V_{\rm DD}^2C_M(1-p-x_1)+V_{\rm DD}^2(C_L+C_M)(u_{[1-p]}-1) \tag{36}$$

where x_1 is the normalized time value in which the end of the output voltage overshoot occurs. $u_{[1-p]}$ is the value of the normalized output voltage when the PMOS device is turned off and is calculated from (16) for $x=1-p$.

The end of the output voltage overshoot occurs in region 2 due to the fact that the discharge of the output node, which is initially charged at $V_{\rm DD}$, does not start in region 1 since the NMOS device is off. Thus, x_1 must be calculated by (11) for $u_{\rm out}=1$. Since this equation cannot be solved analytically, a method similar with that used for the calculation of $x_{\rm sp}$ (Case B—region 3) is used. Equation (17) for $u_{\rm out}=1$ gives the normalized time value (x_1') in which the end of the overshoot occurs, if negligible PMOS current is assumed. The tangent of the output waveform expressed by (11) at the point which corresponds to x_1', is given by

$$u_{\rm out}=mx+q \tag{37}$$

where

$$m=\left.\frac{du_{\rm out}}{dx}\right|_{x=x_1'}=-\sum_{k=1}^{\infty}kg_k(x_1'-n)^{k-1},$$

and

$$q=1+u_{12}-mx_1'-\sum_{k=1}^{\infty}g_k(x_1'-n)^k.$$

By setting $u_{\rm out}=1$ in (37) an accurate approximation for x_1 is derived

$$x_1=\frac{1-q}{m}. \tag{38}$$

The error which is introduced in the calculation of x_1 due to the above approximation is up to 0.2%. The analysis in order to evaluate the short-circuit energy dissipation during a rising output transition is symmetrical and results to the following formula

$$E_{\rm SCR}=\frac{\beta_p V_{\rm DD}^3\tau}{6(1+\delta_p)}\left[3p(1-n-x_1)(p+n-x_1-1)+(1-n)^3-x_1^3\right]-V_{\rm DD}^2C_M(1-n-x_1)-V_{\rm DD}^2(C_L+C_M)u_{[1-n]}. \tag{39}$$

τ is now the input fall time, x_1 is the normalized time value in which the end of the output voltage undershoot occurs and $u_{[1-n]}$ is the value of the normalized output voltage when the NMOS device is turned off. By substituting $E_{\rm SCF}$ and $E_{\rm SCR}$ from (36) and (39) in (33), the short-circuit power dissipation for the cases of fast and slow input ramps, is evaluated.

Case D: In this case, the NMOS device is entered in its linear region, before the input voltage ramp reaches the value $V_{\rm DD}-|V_{\rm TP}|$. Thus, the short-circuit energy dissipation during the falling output transition is given by

$$E_{\rm SCF}=V_{\rm DD}\int I_p\,dt=V_{\rm DD}\left[\int_{x_1}^{x_{\rm sp}} I_p\tau\,dx+\int_{x_{\rm sp}}^{1-p} I_p\tau\,dx\right] \tag{40}$$

where $x_{\rm sp}$ is the normalized time value in which the PMOS device is entering the saturation region. Using (35) for the PMOS current in the first integral and the current equation of the PMOS device in saturation (8) in the second integral, the short-circuit energy dissipation is given by

$$\begin{aligned}E_{\rm SCF}=&\frac{\beta_n V_{\rm DD}^3\tau}{6(1+\delta_n)}\left[3n(x_{\rm sp}-x_1)(n-x_{\rm sp}-x_1)+x_{\rm sp}^3-x_1^3\right]\\&-V_{\rm DD}^2C_M(x_{\rm sp}-x_1)+V_{\rm DD}^2(C_L+C_M)(u_{\rm sp}-1)\\&+\frac{\beta_p V_{\rm DD}^3\tau}{6(1+\delta_p)}(1-x_{\rm sp}-p)^3.\end{aligned} \tag{41}$$

$u_{\rm sp}$ is the normalized output voltage value in which the PMOS device is entering the saturation region and x_1 is given by (38).

Similarly, the short-circuit energy dissipation during a rising output transition is

$$\begin{aligned}E_{\rm SCR}=&\frac{\beta_p V_{\rm DD}^3\tau}{6(1+\delta_p)}\left[3p(x_{\rm sn}-x_1)(p-x_{\rm sn}-x_1)\right.\\&\left.+x_{\rm sn}^3-x_1^3\right]-V_{\rm DD}^2C_M(x_{\rm sn}-x_1)\\&+V_{\rm DD}^2(C_L+C_M)\left(\frac{1-x_{\rm sn}-n}{1+\delta_n}\right)\\&+\frac{\beta_n V_{\rm DD}^3\tau}{6(1+\delta_n)}(1-x_{\rm sn}-n)^3.\end{aligned} \tag{42}$$

where τ is now the input fall time, x_1 is the normalized time value in which the end of the output voltage undershoot occurs and $x_{\rm sn}$ is the normalized time value in which the NMOS device is entering the saturation region.

In Fig. 12, the short-circuit energy dissipation percentage of the capacitive energy dissipation in one switching cycle, is plotted as a function of $A_{\rm no}$. The results have been derived for an inverter with equal PMOS and NMOS device gain factors $\beta_n=\beta_p=0.5$ mA/V^2, $n=p=0.17$ and $\delta_n=\delta_p=0.2$, operating at $V_{\rm DD}=5$ Volts, with an output load of 0.5 pF, and input rise time from 0.2 ns ($A_{\rm no}=1$) to 4 ns ($A_{\rm no}=20$). As can be observed in Fig. 12, the percentage of the short-circuit energy dissipation increases when the input waveform is slow compared with the output waveform (high values of $A_{\rm no}$). Hence, the contribution of the short-circuit current to the total energy increases when the input transition time is increased and the capacitive load is reduced. SPICE measurements have been obtained by using the powermeter subcircuit proposed in [33] and [34]. Also, results using the approaches for the evaluation of the short-circuit energy dissipation presented in [10], [21] and [24] are given. The proposed approach gives results closer to those derived from SPICE simulations than the other methods. This occurs because our model includes the influences of the short-circuit current and the gate-to-drain coupling capacitance on the expression of the inverter output waveform. Also, a quite accurate method is used for the determination of the time where the short-circuiting transistor changes from the linear region to the saturation. The models for the evaluation of the short-circuit dissipation presented in [14] and [20] give inaccurate results, because zero load capacitance is assumed. For example, in an inverter with identical input and output transition times ($A_{\rm no}=5.9$), the short-circuit energy dissipation in one switching cycle which is evaluated using (36) and (39) is about 9.5% of the value as calculated in [14], and 8.5% of the value as calculated in [20]. The validity of the proposed approach has been also examined for the case of the inverter chain. In the example shown in Fig. 10 the discrepancy between the analytical calculated value and that produced from SPICE measurements is about 7.5%.

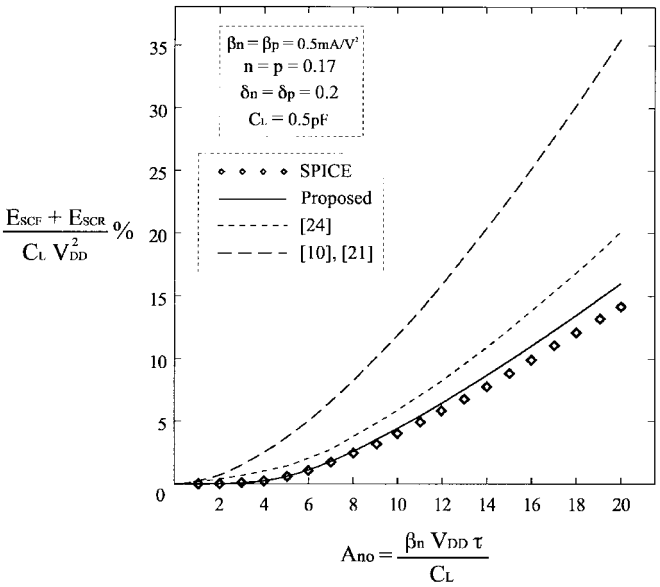

Fig. 12. Short-circuit energy dissipation percentage of the capacitive energy dissipation, derived from SPICE simulations and from analytical expressions.

VI. Conclusion

In this paper an accurate analytical method for the evaluation of the propagation delay and the short-circuit power dissipation in a CMOS inverter, has been presented. In order to achieve that, analytical expressions of the inverter output ramp response for all the cases of input ramps, have been derived. These expressions take into account the influences of the short-circuit current and the gate-to-drain coupling capacitance. In addition, the effective transition time of the inverter is evaluated in order to make the ramp model applicable to real circuit applications.

References

[1] J. Rubinstein, P. Penfield, and M. A. Horowitz, "Signal delay in *RC* tree networks," *IEEE Trans. Computer-Aided Design*, vol. CAD-2, pp. 202–211, July 1983.

[2] J. K. Ousterhout, "A switch level timing verifier for digital MOS VLSI," *IEEE Trans. Computer-Aided Design*, vol. CAD-4, pp. 336–349, July 1985.

[3] A. C. Deng and Y. C. Shiaw, "Generic linear RC delay modeling for digital CMOS circuits," *IEEE Trans. Computer-Aided Design*, vol. 9, pp. 367–376, Apr. 1990.

[4] L. Brocco, S. Mccormic, and J. Allen, "Macromodeling CMOS circuits for timing simulation," *IEEE Trans. Computer-Aided Design*, vol. 7, pp. 1237–1249, Dec. 1988.
[5] D. Deschacht, M. Robert, and D. Auvergne, "Explicit formulation of delays in CMOS data paths," *IEEE J. Solid-State Circuits*, vol. 23, pp. 1257–1264, Oct. 1988.
[6] D. Auvergne, N. Azemard, D. Deschacht, and M. Robert, "Input waveform slope effects in CMOS delays," *IEEE J. Solid-State Circuits*, vol. 25, pp. 1588–1590, Dec. 1990.
[7] H. Y. Chen and S. Dutta, "A timing model for static CMOS gates," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, Nov. 1989, pp. 72–75.
[8] 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.
[9] J. R. Burns, "Switching response of complementary symmetry MOS transistor logic circuits," *RCA Rev.*, vol. 25, pp. 627–661, Dec. 1964.
[10] N. Hedenstierna and K. O. Jeppson, "CMOS circuit speed and buffer optimization," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, pp. 270–281, Mar. 1987.
[11] A. I. Kayssi, K. A. Sakallah, and T. M. Burks, "Analytical transient response of CMOS inverters," *IEEE Trans. Circuits Syst. I*, vol. 39, pp. 42–45, Jan. 1992.
[12] 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.
[13] S. R. Vemuru and A. R. Thorbjornsen, "A model for delay evaluation of a CMOS inverter," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 1990, vol. 1, pp. 89–92.
[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, pp. 584–594, Apr. 1990.
[15] T. Sakurai and A. R. Newton, "A simple MOSFET model for circuit analysis," *IEEE Trans. Electron Devices*, vol. 38, pp. 887–894, Apr. 1991.
[16] 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.
[17] S. H. K. Embabi and R. Damodaran, "Delay models for CMOS, BiCMOS, BiNMOS circuits and their applications for timing simulations," *IEEE Trans. Computer-Aided Design*, vol. 13, pp. 1132–1142, Sept. 1994.
[18] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*, New York: McGraw-Hill, 1988, pp. 123–130.
[19] A. P. Chandrakasan, S. Sheng, and R. W. Brodersen, "Low-power CMOS digital design," *IEEE J. Solid-State Circuits*, vol. 27, pp. 473–484, Apr. 1992.
[20] 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, pp. 468–473, Aug. 1984.
[21] N. Hedenstierna and K. O. Jeppson, "Comments on 'A module generator for optimized CMOS buffers'," *IEEE Trans. Computer-Aided Design*, vol. 12, pp. 180–181, Jan. 1993.
[22] L. Bisdounis, S. Nikolaidis, O. Koufopavlou, and C. E. Goutis, "Modeling the CMOS short-circuit power dissipation," in *Proc. IEEE Int. Symp. on Circuits and Systems*, May 1996, vol. 4, pp. 469–472.
[23] B. S. Cherkauer and E. G. Friedman, "A unified design methodology for CMOS tapered buffer," *IEEE Trans. VLSI Syst.*, vol. 3, pp. 99–111, Mar. 1995.
[24] 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.
[25] S. Turgis, N. Azemard, and D. Auvergne, "Explicit evaluation of short-circuit power dissipation for CMOS logic structures," in *Proc. Int. Symp. on Low-Power Design*, Apr. 1995, pp. 129–134.
[26] 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*, May 1996, vol. 4, pp. 751–754.
[27] N. H. E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design: A Systems Perspective*. New York: McGraw-Hill, 1993. pp. 183–191.
[28] H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," *IEEE J. Solid-State Circuits*, vol. SC-3, pp. 285–289, Sept. 1968.
[29] H. T. Davis, *Introduction to Nonlinear Differential and Integral Equations*, New York: Dover, pp. 57–78, 1962.
[30] M. Shoji, *CMOS Digital Circuit Technology*. Englewood Cliffs, NJ: Prentice-Hall, 1988, pp. 121–123.
[31] A. L. Nelson, K. W. Folley, and M. Coral, *Differential Equations*. Boston, MA: Heath, pp. 180–214, 1964.
[32] M. R. Spiegel, *Mathematical Handbook of Formulas and Tables*. New York: McGraw-Hill, p. 183, 1968.
[33] S. M. Kang, "Accurate simulation of power dissipation in VLSI circuits," *IEEE J. Solid-State Circuits*, vol. SC-21, pp. 889–891, Oct. 1986.
[34] G. Y. Yacoub and W. H. Ku, "An enhanced technique for simulating short-circuit power dissipation," *IEEE J. Solid-State Circuits*, vol. 24, pp. 844–847, June 1989.

Labros Bisdounis (S'95) was born in Agrinio, Greece, in 1970. He received the Diploma degree in electrical engineering from the Department of Electrical and Computer Engineering, University of Patras, Greece, in 1992. He is currently pursuing the Ph.D. degree at the VLSI Design Laboratory of the same department.

His main research interest is on various aspects of CMOS VLSI design such as circuit timing analysis, power dissipation modeling, low power and high speed CMOS digital design.

Spiridon Nikolaidis (S'89–M'93) was born in Kavala, Greece, in 1965. He received the Diploma and Ph.D. degrees in electrical engineering from Patras University, Greece, in 1988 and 1994, respectively.

Since September 1996 he has been with the Department of Physics of the Aristotle University of Thessaloniki, Greece, as a Lecturer in VLSI design. His research interests include CMOS gate propagation delay and power consumption modeling, high speed and low power CMOS circuit techniques, and high speed and low power DSP architectures.

Odysseas Koufopavlou (S'89–M'90) was born in Athienou, Cyprus, in 1959. He received the Diploma of electrical engineering in 1983 and the Ph.D. degree in electrical engineering in 1990, both from University of Patras, Greece.

From 1990 to 1994 he was with the IBM Thomas J. Watson Research Center, Yorktown Heights, NY. Since 1994, he has been an Assistant Professor at the Department of Electrical and Computer Engineering, University of Patras. His research interests include VLSI, low power design, and high performance communication subsystems architecture and implementation. He has several publications and inventions.

Dr. Koufopavlou is a member of Technical Chamber of Greece.