An Efficient Timing Model for CMOS Combinational Logic Gates

CHUNG-YU WU, JEN-SHENG HWANG, CHIH CHANG, AND CHING-CHU CHANG

Abstract—**A new general timing model for CMOS combinational logic gates is proposed. In this model, the linearized large-signal equivalent circuit of a gate is first constructed. Then applying the dominant-pole–dominant-zero (DPDZ) method, the dominant pole of the equivalent circuit is calculated. Using this pole, the signal timing can be explicitly expressed. Comparisons between calculation results and simulation results are made and error analyses are performed. The worst-case error in characteristic-waveform timing can be confined to be within 35 percent for CMOS inverters, multi-input NOR gates or multi-input NAND gates with different device dimensions, capacitive loads, and device parameters. Better accuracy can be obtained for logic gates with commonly-used channel dimension or large capacitive load. For internal waveforms not deviating much from the characteristic waveforms, the worst-case error in signal timing is not substantially increased. Applying the proposed timing model in an experimental timing simulator, the signal timing can be analyzed accurately and efficiently with reduced CPU time and memory.**

List of Symbols

- B Mobility correction parameter (SPICE device parameter).
- $C_{\text{bdn(p)}}$ Bulk–drain p-n junction capacitance of an n-channel (p-channel) MOSFET.
- $C_{\text{bsn(p)}}$ Bulk–source p-n junction capacitance of an n-channel (p-channel) MOSFET.
- $C_{\text{gdovn(p)}}$ Gate–drain overlap capacitance of an n-channel (p-channel) MOSFET.
- $C_{\text{gsovn(p)}}$ Gate–source overlap capacitance of an n-channel (p-channel) MOSFET.
- C_L Fixed load capacitance of a logic gate.
- $C_{on(p)}$ Channel oxide capacitance of an n-channel (p-channel) MOSFET.
- DELTA Channel width factor (SPICE device parameter).
- GAMMA Bulk threshold parameter in SPICE, which represents the proportionality factor relating the change in threshold voltage to backgate bias.
- $I_{dn(p)}$ Drain current of an n-channel (p-channel) MOSFET in large-signal model.
- I_{DS} DC drain current of an MOSFET.
- L_{eff} Effective channel length of an MOSFET.
- NSUB Substrate Doping.
- $P_{f(r)}$ Effective dominant pole in the fall (rise) characteristic waveform case.
- $P_{fd}(Z_{\text{fd}})$ Dominant pole (zero) in the fall characteristic waveform case.
- $P_{rd}(Z_{rd})$ Dominant pole (zero) in the rise characteristic waveform case.
- q Magnitude of electronic charge.
- $t_{df(r)}$ Initial fall (rise) delay time.
- $T_{F(\text{R})}$ Fall (rise) time which is the time interval within which the output voltage lowers (raises) from 0.9 V_{DD} (0.1 V_{DD}) to 0.1 V_{DD} (0.9 V_{DD}).
- T_{OX} Channel oxide thickness.
- T_P Pair delay time which is the sum of fall delay time and rise delay time.
- $T_{\text{PHL(LH)}}$ Fall (rise) delay time which is the time interval between $v_i = \frac{1}{2} V_{\text{DD}}$ to $v_0 = \frac{1}{2} V_{\text{DD}}$.
- UCRIT Critical field for mobility degradation (SPICE device parameter).
- UEXP Critical field exponent in mobility degradation (SPICE device parameter).
- UEXPL(W) Correction factor of UEXP for short channel (narrow channel) case (SPICE device parameter).
- UO Surface mobility (SPICE device parameter).
- UOL(W) Correction factor of UO for short channel (narrow channel) case (SPICE device parameter).
- $v_{i(o)}$ Input (output) large-signal voltage of a logic gate.
- $v_{02(3)}$ Large-signal voltage at internal nodes of a 3-input CMOS NOR gate.
- $V_{\text{BS}(D)}$ Bulk–source (drain) reverse bias of an MOSFET.
- V_{DD} Power supply voltage.
- $V_{\text{DS(GS)}}$ Drain (gate) source voltage of an MOSFET.
- V_{max} Maximum drift velocity of carriers.
- V_T Threshold voltage of an MOSFET under backgate bias.
- V_{T0} Zero-bias threshold voltage of the MOSFET (SPICE device parameter).
- W_{eff} Effective channel width of an MOSFET.

Manuscript received October 8, 1984; revised January 8, 1985. This work was supported by ERSO, ITRI, Republic of China.

C.-Y. Wu was with the Institute of Electronics, National Chiao-Tung University, Hsin-Chu, Taiwan, Republic of China. He is now with the Department of Electrical Engineering, School of Engineering and Applied Science, Portland State University, Portland, OR 92707.

Y.-H. Yang is with the Institute of Electronics, National Chiao-Tung University, Hsin-Chu, Taiwan, Republic of China.

C. Chang and C.-C. Chang are with the Electronic Research and Service Organization, Industrial Technology Research Institute, Hsin-Chu, Taiwan, Republic of China.

0278-0070/85/1000-0636$01.00

$W_{n(p)}$ Channel width (as drawn) of an n-channel (p-channel) MOSFET.
XJ Metallurgical junction depth of an MOSFET (SPICE device parameter).
$\epsilon_{SI(SIO_2)}$ Permittivity of Si semiconductor (silicon dioxide).
μ_s Surface mobility of carriers.
ϕ_F Fermi potential.
$\phi_{Fn(p)}$ Fermi potential of an n-type (p-type) silicon.

I. Introduction

WITH THE enhanced demands for chip performance and complexity in MOS LSI/VLSI design, signal timing correctness has become more and more important in order to avoid speed degradation, signal glitches, or logic faults. Therefore, new circuit simulators or timing simulators such as MOTIS [1], SPLICE [2], and RELAX [3] have been constructed to analyze a complex VLSI circuit efficiently with acceptable accuracy and less CPU time and storage.

Meanwhile, increasing efforts have concentrated on the constructions of analytic and efficient timing models or calculation methods for logic gate delay [4]–[9] and wiring delay [10], [11]. Generally, the purposes of such models or methods are:

(1) to implement timing simulators [12] or logic simulators [9] to efficiently analyze or verify the signal timing of LSI/VLSI chips;

(2) to obtain a deeper physical insight into the speed characteristics of digital IC's and to optimize their performance [8], [13];

(3) to quickly perform the so-called "timing synthesis" [14] which determines optimum MOSFET channel dimensions from given timing specifications through the use of analytic timing equations.

Thus it will be quite useful in VLSI design if an accurate and analytic timing model can be developed.

Among all the timing models [4]–[9] explored so far, Glasser [7] developed closed form expressions for the upper and lower bounds on propagation delay of a logic gate. In his model, Thevenin equivalent circuits and *RC* ladder networks were adopted to replace logic gates and interconnection lines, respectively. Thus digital circuits can be treated as analog circuits. All other models [4]–[6], [8], [9] constructed timing models for inverters [4]–[6], [8], [9] or transmission gates [9] based upon the concept of charging and discharging the capacitive load.

In this paper, a new general timing model for CMOS combinational logic gates is proposed. In this model, the logic gate under consideration is first replaced by its large-signal equivalent circuit. Then nonlinear currents and capacitances in the equivalent circuit are linearized and the resultant equivalent circuit becomes an analog one. After the dominant pole is found, a closed-form expression for the rise or fall times can be obtained. Based upon the derived rise and fall times, the gate delay can be explicitly expressed. Applying these model formulation procedures, signal timing for CMOS inverters, multi-input NOR gates, and multi-input NAND gates have been characterized. As an illustrating example in model formulation, only a 3-input CMOS NOR gate is considered in Section II of this paper.

Comparisons between exact computer simulation results and theoretical calculation results have been extensively made. Based on these comparisons, error analysis is performed and the versatility of the proposed timing model is investigated in Section III. It is found that for logic gates with voltage waveforms not severely different from characteristic waveforms [4], [8], and for logic gates with different MOS channel dimensions, different capacitive loads and different device parameters, the general agreement between calculation results and simulation results is satisfactory.

As a preliminary observation, the derived timing equations have also been used in experimental timing simulator. Simulation results of some circuit examples are investigated and analyzed. It is found that the CPU time and storage requirement in timing simulations using the developed timing model is quite small.

II. Model Formulation

It is found in the design of CMOS multi-stage tapered buffer that the minimum total delay time can be achieved when the delay time in each stage is forced to be the same [8]. In this optimized case, the waveforms at the output nodes of all the intermediate stages are the same, being independent of the input excitation [8]. Such waveforms are called the characteristic waveforms [4].

Based upon this concept, it is expected that the total delay time of a complex IC can be reduced if the deviations of waveforms at internal nodes are minimized despite of different capacitive loads at these nodes. Using this design technique, the internal waveforms within that IC will be very close to the characteristic waveforms.

Therefore, in the proposed new timing model for CMOS logic gates, the case of characteristic waveform is considered. However, as will be shown later, the results of our timing model can be applied for waveforms not deviating much from characteristic waveforms, as in some actual design cases. This versatility increases the applicability of our timing model.

The first step in the model formulation is to determine the operating region of each MOSFET in the logic gate under consideration from its characteristic waveform. As a demonstrating example, consider a CMOS 3-input NOR gate in the worst-case timing condition in which only the input farthest from the output node is excited, as shown in Fig. 1. Its typical fall characteristic waveforms are shown in Fig. 2. Such waveforms can be obtained from the output node of an intermediate stage in a string of NOR gates with the same capacitive load C_L and fanout number N as shown in Fig. 1.

To determine the operating region of an MOSFET, the

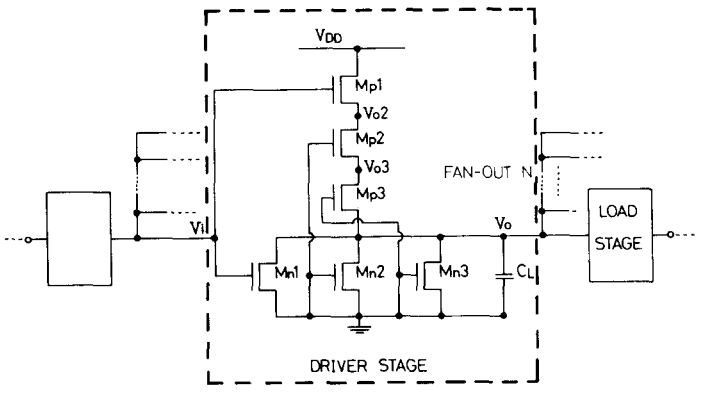

Fig. 1. A chain of identical CMOS 3-input NOR gates.

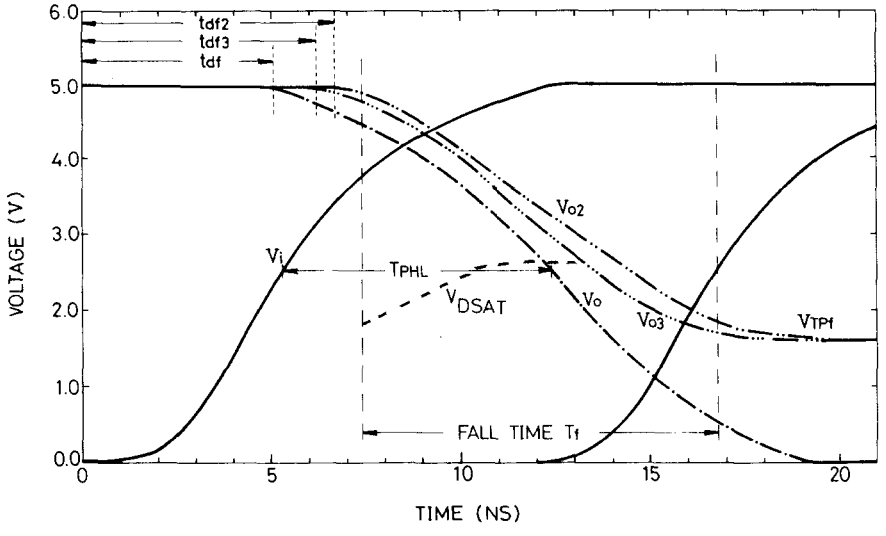

Fig. 2. Typical fall characteristic waveforms of a CMOS 3-input NOR gate.

drain–source saturation voltage V_{DSAT} of the device under various values of V_{GS} and V_{DS} as determined from the waveforms in Fig. 2, are calculated by computer simulations which consider the velocity saturation effect. Typical V_{DSAT} curve for the MOSFET M_{n1} with effect channel length 1.9 μm has been plotted along with the falling waveforms in Fig. 2. It can be seen that during nearly half of the fall-time period defined as the time interval within which the output voltage v_0 lowers from 0.9 V_{DD} to 0.1 V_{DD}, the MOSFET M_{n1} is operated in the linear region. For simplicity, the MOSFET is considered to be operated in the linear region during the fall-time period. This assumption does not introduce large error in the timing calculations as will be shown later. However, when the device channel length is further scaled down, the MOSFET M_{n1} will mostly be operated in the saturation region during the fall-time period due to the substantially reduced saturation voltage. In this case, the above assumption will cause a larger error.

Based upon the similar consideration, the MOSFET's M_{p2} and M_{p3} are found to be operated in their linear regions. The MOSFET M_{p1} is mostly operated in the off region since its gate–source voltage (V_{DD}–V_i) mostly remains below its threshold voltage. In the load stage, the MOSFET M_{n1} is operated in the linear region during the fall-time period whereas the MOSFET M_{p1} in the load stage is found to be operated in its saturation region.

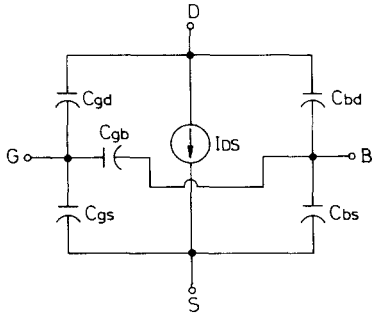

Fig. 3. Four-terminal large-signal equivalent circuit of an MOSFET.

After the operating region of each MOSFET has been determined, the second step is to find the large-signal equivalent circuit of the logic gate. In general, the large-signal equivalent circuit of an MOSFET can be drawn as in Fig. 3 where reverse currents across the drain–bulk and source–bulk junctions have been neglected. The capaci-

tance values of the four capacitors in this circuit are all voltage-dependent [15], whereas the expressions of the drain current I_{DS} are different for different operating regions. In order to conveniently perform comparisons to computer simulation results obtained from the modified SPICE 2 Program called ESPICE, the current model adopted here is the same as that used in ESPICE. Therefore, in the linear region, I_{DS} can be expressed as [16].

$$I_{DS} = \beta\left\{\left(V_{GS} - V_{BIN} - \frac{\eta}{2} V_{DS}\right) V_{DS} - \frac{2}{3}\gamma_s[(2\phi_F + V_{DS} - V_{BS})^{3/2} - (2\phi_F - V_{BS})^{3/2}]\right\}. \quad (1)$$

All the parameters involved in (1) are described in Table I.

Note that any fixed voltage (e.g., V_{DD}) at the node of the equivalent circuit only affects the drain current and has no effect on the capacitor transient currents. Therefore, it can be grounded without affecting the nodal current equations of the equivalent circuit. Based upon this concept, power supply V_{DD} can be grounded. Furthermore, since the output voltage of the load stage remains nearly unchanged during the fall-time period, it can also be effectively grounded. Thus all the output stages of the load stage are shielded out. This characteristic has been found in all the CMOS logic gates.

Directly replacing each MOSFET in the logic gate by its equivalent circuit and grounding the fixed voltages, the resultant large-signal equivalent circuit of the CMOS 3-input NOR gate during the fall-time period is shown in Fig. 4. Since the output node of the load stage has been grounded, it only contributes capacitive load to the driver stage. This capacitive load has been included in C_3.

Because the transistors M_{n1}, M_{p2}, and M_{p3} are operated in the linear region during the fall-time period, their drain currents I_{dn1}, I_{dp2} and I_{dp3} in the equivalent circuit of Fig. 4 can be formulated from (1) with suitable expressions for V_{GS} and V_{DS}. All the constant terms involved in these drain currents must be discarded. The reason will be described later. On the other hand, I_{dp1} is set to zero since the transistor M_{p1} is off. Note that all the currents and capacitances in the equivalent circuit are all voltage-dependent. Therefore, this circuit is a highly nonlinear one and v_0 can not be analytically solved without any linearizations and approximations.

The third step in the model formulation is to linearize the nonlinear capacitances and currents so that the equivalent circuit becomes a linear one and can be treated as an analog circuit. First, using the commonly known approximations [15], the gate–source and gate–drain capacitances are fixed according to the operating region of the MOSFET. Furthermore, all the voltage-dependent drain–bulk and source–bulk junction capacitances are evaluated at fixed biases obtained at the logic crossover point of the logic gate. This approximation on the junction capacitance is just the same as that adopted in MOTIS [1]. During the fall-time period of CMOS 3-input NOR gate considered above, the logic crossover point is defined as the point at which the output voltage falls to $V_{DD}/2$. At this point, the input voltage is nearly equal to $4V_{DD}/5$ as may be seen from the waveforms shown in Fig. 2. The voltages V_{02} and V_{03} at node 2 and node 3, respectively, are approximately $3V_{DD}/5$. Therefore the reverse bias across the capacitances C_{bdn1} and C_{bdp3} is $V_{DD}/2$ whereas the biases across C_{bdp2} (C_{bsp3}) and C_{bdp1} (C_{bsp2}) are $2V_{DD}/5$. In eval-

TABLE I
MOSFET CURRENT EQUATION USED IN THE MODIFIED SPICE2 PROGRAM

$$I_{DS} = \beta\{(V_{GS}-V_{BIN} - \frac{\eta}{2}V_{DS})V_{DS} - \frac{2}{3}\gamma_S[(2\phi_F+V_{DS}-V_{BS})^{3/2} - (2\phi_F-V_{BS})^{3/2}]\}$$

where

$$\beta = \frac{W_{eff}}{L_{eff}} \cdot \mu_S \cdot \frac{\epsilon_{SIO_2}}{T_{OX}}$$

$$XUO = UO \cdot \left(1 - \frac{UOL}{L_{eff}} - \frac{UOW}{W_{eff}}\right)$$

$$XUEXP = UEXP + \frac{UEXPL}{L_{eff}} - \frac{UEXPW}{W_{eff}}$$

$$XUO' = XUO \cdot \left(\frac{UCRIT \cdot \epsilon_{SI}}{\frac{\epsilon_{SIO_2}}{T_{OX}}[V_{GS}-V_T-UTRA \cdot \min(V_{DS}, 2\phi_F)]}\right)^{XUEXP}$$

$$\mu_S = XUO' \cdot \frac{1}{\{1 + [\frac{XUO' \cdot \min(V_{DS}, 2\phi_F)}{VMAX \cdot L_{eff}}]^B\}^{1/B}}$$

$$\eta = 1 + \frac{\pi}{4} \cdot \frac{DELTA \cdot \epsilon_{SI}}{\frac{\epsilon_{SIO_2}}{T_{OX}} \cdot W_{eff}}$$

$$V_{BIN} = V_{BI} + (\eta - 1) \cdot (2\phi_F - V_{BS})$$

$$V_{BI} = V_{TO} - GAMMA \cdot \sqrt{2\phi_F}$$

$$V'_{BIN} = V_{BI} + (\eta - 1) \cdot 2\phi_F$$

$$\gamma_S = GAMMA \cdot (1-\alpha_S-\alpha_D)$$

$$\alpha_S = \frac{1}{2} \cdot \frac{XJ}{L_{eff}} \cdot \left(\sqrt{1+2 \cdot \frac{WS}{XJ}} - 1\right)$$

$$\alpha_D = \frac{1}{2} \cdot \frac{XJ}{L_{eff}} \cdot \left(\sqrt{1 + 2 \cdot \frac{WD}{XJ}} - 1\right)$$

$$W_S = XD \cdot \sqrt{2\phi_F - V_{BS}}$$

$$W_D = XD \cdot \sqrt{2\phi_F - V_{BD}}$$

$$XD = \sqrt{\frac{2\epsilon_{SI}}{q \cdot N_{SUB}}}$$

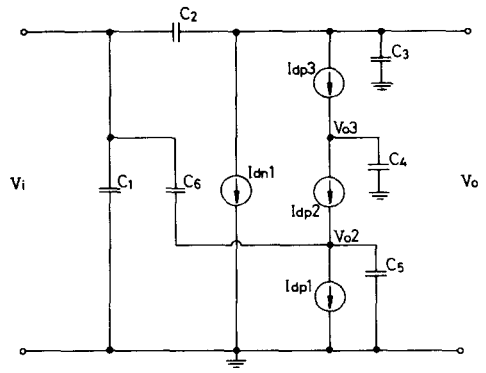

Fig. 4. Large-signal equivalent circuit of a CMOS 3-input NOR gate during the fall-time period.

uating the capacitance value, the formula of junction capacitance used in the SPICE 2 Program [16] is applied.

After the linearization cited above, the capacitances in the equivalent circuit of Fig. 4 becomes voltage-independent and can be expressed in terms of various junction capacitances and MOS capacitances.

To linearize the drain current in (1), first the V_{DS}- and V_{GS}-dependent carrier mobility μ_s as expressed in Table I is evaluated at the logic crossover point. In the case of CMOS 3-input NOR gate, the mobilities μ_{sn1} and μ_{sp2} are evaluated at $V_{DS} = V_{DD}/2$, $V_{GS} = 4V_{DD}/5$ and $V_{DS} = V_{DD}/6$, $V_{GS} = 3V_{DD}/5$, respectively. The mobility μ_{sp3} is evaluated at $V_{DS} = V_{DD}/6$ and $V_{GS} = 3V_{DD}/5$.

Further linearization of the drain current can be performed by first determining the time-domain function of each node voltage from the characteristic waveforms. Here we approximate the node voltage function by a single-pole response. This approximation will simplify the calculations and will lead to a closed-form expression for signal timing. Moreover, the error of this approximation is small as compared with the simulated waveforms, as will be justified in the next section. For the 3-input NOR gate, each node voltage as a function of time t has been determined from Fig. 2 as

$$v_i(t) = [V_{DD} - V_{DD} \exp(-p_r t)]\, u(\cdot) \tag{2}$$

$$v_0(t) = V_{DD}u(t) - V_{DD}u(t - t_{df}) + V_{DD} \exp[-p_f(t - t_{df})]\, u(t - t_{df}) \tag{3}$$

$$v_{03}(t) = V_{DD}u(t) - (V_{DD} - V_{Tpf}) \cdot u(t - t_{df3}) + (V_{DD} - V_{Tpf}) \cdot \exp[-p_{f3}(t - t_{df3})]\, u(t - t_{df3}) \tag{4}$$

$$v_{02}(t) = V_{DD}u(t) - (V_{DD} - V_{Tpf}) \cdot u(t - t_{df2}) + (V_{DD} - V_{Tpf}) \cdot \exp[-p_{f2}(t - t_{df2})]\, u(t - t_{df2}) \tag{5}$$

where the initial delay times t_{df}, t_{df3} and t_{df2} which are defined in Fig. 2 have the relation $t_{df} < t_{df3} < t_{df2}$. The voltage V_{Tpf} is the minimum residual voltage at the source node of a PMOS after complete discharging. The voltage V_{Tpf} can be solved as

$$V_{Tpf} = -\{2[\gamma_{sp}(2\phi_{Fn})^{1/2} - V_{Tp}] + \gamma_{sp}^2\}/2 + \{[2\gamma_{sp}(2\phi_{Fn})^{1/2} - 2V_{Tp} + \gamma_{sp}^2]^2 + 4[2\gamma_{sp}V_{Tp}(2\phi_{Fn})^{1/2} + \gamma_{sp}^2 V_{DD} - V_{Tp}^2]\}^{1/2}/2 \tag{6a}$$

where

$$V_{Tp} \equiv V_{BINp} + \gamma_{sp}(2\phi_{Fn})^{1/2}. \tag{6b}$$

Since it has been found that the approximation $p_{f3} \simeq p_{f2} \simeq p_f$ introduces small error, this approximation will be used thereafter. The unique pole p_f is the one to be calculated from the equivalent circuit.

Further linearization of the drain current in (1) starts from the linearization of the terms $(2\phi_F + V_{DS} - V_{BS})^{3/2}$ and $(2\phi_F - V_{BS})^{3/2}$. The linearization can be described by the following equation:

$$(2\phi_F + V_{DS} - V_{BS})^{3/2} \rightarrow (2\phi_F + V_{DS} - V_{BS})(2\phi_F + V_{DS}|_{t=t_e} - V_{BS}|_{t=t_e})^{1/2}$$

where $t = t_e$ is the time at the logic crossover point. Using (3) and the relation $v_0(t = t_e) = V_{DD}/2$, the time t_e can be solved as

$$t_e = t_{df} + (\ln 2)/pf. \tag{7}$$

Finally, there are nonlinear voltage terms in the drain current equations to be linearized. Applying the linearization procedures as described in Appendix, the resultant expressions for those nonlinear terms after discarding the constant terms are

$$v_i v_0(t) = F_0 v_0(t) \tag{8}$$

$$v_0^2(t) = \tfrac{1}{2} V_{DD} v_0(t) \tag{9}$$

$$v_{03}^2(t) = B_0 v_{03}(t) \tag{10}$$

$$v_{02}^2(t) = C_0 v_{02}(t). \tag{11}$$

The constants F_0, B_0, and C_0 have already been expressed in Appendix.

Using (8)-(11), the linearized drain currents can be written as

$$I_{dn1} = \beta_{n1D} v_0 \tag{12}$$

$$I_{dp3} = \beta_{p3D} v_0 - \beta_{p3S} v_{03} \tag{13}$$

$$I_{dp2} = \beta_{p2D} v_{03} - \beta_{p2S} v_{02} \tag{14}$$

where

$$\beta_{n1D} \equiv \beta_{n1}\left[F_0 - V'_{BINn1} - \frac{\eta_{n1}}{4} V_{DD} - \frac{2}{3}\gamma_{sn1}(2\phi_{FP} + \frac{1}{2} V_{DD})^{1/2}\right] \tag{15a}$$

$$\beta_{p3D} \equiv \beta_{p3}\left[\frac{\eta_{p3}}{4} V_{DD} - V'_{BINp3} - (\eta_{p3} - 1) V_{DD} - \frac{2}{3}\gamma_{sp3}\left(2\phi_{Fn} + \frac{1}{2} V_{DD}\right)^{1/2}\right] \quad (15b)$$

$$\beta_{p3S} \equiv \beta_{p3}\left[\frac{\eta_{p3}}{2} B_0 - V'_{BINp3} - (\eta_{p3} - 1) V_{DD} - \frac{2}{3}\gamma_{sp3}(2\phi_{Fn} + V_{DD} - F_{F2})^{1/2}\right] \quad (15c)$$

$$\beta_{p2D} \equiv \beta_{p2}\left[\frac{\eta_{p2}}{2} B_0 - V'_{BINp2} - (\eta_{p2} - 1) V_{DD} - \frac{2}{3}\gamma_{sp2}(2\phi_{Fn} + V_{DD} - V_{F2})^{1/2}\right] \quad (15d)$$

$$V_{F2} \equiv v_{03}|_{t=t_e} = V_{Tpf} + \frac{1}{2}(V_{DD} - V_{Tpf}) \cdot \exp\left[-p_f(t_{df} - t_{df3})\right] \quad (15e)$$

$$\beta_{p2s} \equiv \beta_{p2}\left[\frac{\eta_{p2}}{2} C_0 - V'_{BINp2} - (\eta_{p2} - 1) V_{DD} - \frac{2}{3}\gamma_{sp2}(2\phi_{Fn} + V_{DD} - V_{F1})^{1/2}\right] \quad (15f)$$

$$V_{F1} \equiv v_{02}|_{t=t_e} = V_{Tpf} + \frac{1}{2}(V_{DD} - V_{Tpf}) \cdot \exp\left[-p_f(t_{df} - t_{df2})\right]. \quad (15g)$$

Note that the surface mobilities in the above expressions have been evaluated at the logic crossover point. Now, since the nonlinear capacitances and currents have been linearized, the whole equivalent circuit, as shown in Fig. 4, has become a linear circuit.

After detailed investigations on the characteristic waveforms of various CMOS logic gates, it has been found that generally during the fall- or rise-time periods, the input voltage has entered its tail region with small variation or nearly constant voltage level. Therefore, the input excitation has only negligible effect on the output response. Similar conclusion can be drawn for those waveforms not severely deviating from characteristic waveforms. In these cases, the output waveform can be well characterized by the poles and zeros of the linearized equivalent circuit, which are affected by the input voltage. Here, in order to obtain the closed-form expression for signal timing, the modified dominant-pole approximation is adopted. This approximation is in consistence with the above approximation of single-pole response. Under this approximation, the constant terms in drain currents may be discarded since they do not affect the poles and zeros of the linearized equivalent circuit.

According to our observations, simply considering the dominant pole is not accurate enough when there exists a zero sufficiently close to the poles of interest. This can be demonstrated by a simple example. Consider a stable circuit with two poles and one zero. Its transfer function $H(s)$ can be expressed as

$$H(s) = H_0 \frac{s + c}{(s + a)(s + b)}. \quad (16)$$

Using the conventional dominant pole calculation method [17], the dominant pole P^* can be determined despite of the zero c. The resultant expression of $1/p^*$ is

$$\frac{1}{p^*} = \frac{1}{a} + \frac{1}{b}. \quad (17)$$

However, if the zero c is sufficiently close to one of the two poles, say b, the transfer function can be simplified as

$$H(s) \simeq H_0 \frac{1}{(s + a)}. \quad (18)$$

Thus the dominant pole is a rather than that in (17).

Here, a new calculation method called the dominant-pole–dominant-zero (DPDZ) method is proposed in our modified dominant-pole approximation. In this method, the effective dominant pole P_e is calculated by

$$\frac{1}{p_e} = \frac{1}{a} + \frac{1}{b} - \frac{1}{c}. \quad (19)$$

Evidently, (19) reduces to $p_e = a$ when $b \simeq c$. Therefore, the DPDZ method is more adequate than the conventional method. If there exists more than one zero, the dominant zero must be used in (19).

Another modification proposed in the DPDZ method is for the case of closely spaced poles without any actual dominant pole. In this case, using the effective dominant pole of (19) to calculate the rise or fall time leads to a worst-case error greater than 20 percent [17] which occurs when two poles overlap (i.e., $1/a = 1/b - 1/c$). To completely reduce this error, the value of $1/p_e$ must be equal to $(2/a)/(1 + 1/5)$. If a factor of 1/2 is multiplied to all the terms in (19) concerning the larger poles, the resultant value of $1/p_e$ will be $3/2a$ which will have an error of 10 percent. In other cases with nonoverlapping poles, the error was found to be smaller than 10 percent with the correction factor of 1/2. Therefore, the effective dominant pole now can be expressed generally as

$$\frac{1}{p_e} = \frac{1}{a} + \frac{1}{2}\left(\sum_i \frac{1}{b_i} - \frac{1}{c}\right) \quad (20)$$

where b_i represents any pole larger than the pole a. This final form of p_e will be used in timing models of all CMOS combinational logic gates. Then the resultant error due to the modified dominant pole approximation will be within 10 percent.

Now, the fourth step in the model formulation is to calculate the effective dominate pole from the linearized equivalent circuit using the DPDZ method. In the case of 3-input NOR gates, applying the Kirchhoff's current law (KCL) to the circuit shown in Fig. 3 leads to the KCL equations.

Taking the Laplace transformation and discarding the constant terms, the resultant equations in matrix form can be written as

$$\begin{bmatrix} S(C_2 + C_3) + \beta_{n1D} + \beta_{p3D} & -\beta_{p3S} & 0 \\ -\beta_{p3D} & SC_4 + \beta_{p2D} + \beta_{p3S} & -\beta_{p2S} \\ 0 & -\beta_{p2D} & S(C_5 + C_6) + \beta_{p2S} \end{bmatrix} \begin{bmatrix} v_0 \\ v_{03} \\ v_{02} \end{bmatrix} = \begin{bmatrix} SC_2 \\ 0 \\ SC_6 \end{bmatrix} \begin{bmatrix} v_i \end{bmatrix}. \tag{21}$$

The transfer function V_0/V_i can be calculated from (21) and the inverse of the dominant pole can be calculated. In the expression of the dominant pole, the terms concerning both C_4 and $(C_5 + C_6)$ are the terms corresponding to the nondominant poles larger in magnitude than the pole generated by $(C_2 + C_3)$. Therefore, according to the DPDZ method, these terms must be multiplied by 1/2. The resultant expression is

$$\frac{1}{P_{fd}} = (C_2 + C_3)/\beta_{n1D} + \frac{1}{2}\left[C_4\left(\frac{1}{\beta_{p3S}} + \frac{1}{\beta_{n1D}}\frac{\beta_{p3D}}{\beta_{p3S}}\right) + (C_5 + C_6)\left(\frac{1}{\beta_{p2S}} + \frac{1}{\beta_{p3S}}\frac{\beta_{p2D}}{\beta_{p2S}} + \frac{1}{\beta_{n1D}}\frac{\beta_{p3D}}{\beta_{p3S}}\frac{\beta_{p2D}}{\beta_{p2S}}\right)\right]. \tag{22}$$

Similarly, the inverse of the dominant zero can be calculated. After multiplying the factor 1/2, the resultant expression is

$$\frac{1}{Z_{fd}} = \frac{C_2}{C_2 + C_6}\frac{1}{2}\left[C_4/\beta_{p3S} + (C_5 + C_6) \cdot \left(\frac{1}{\beta_{p3S}}\frac{\beta_{p2D}}{\beta_{p2S}} + \frac{1}{\beta_{p2S}}\right)\right] \tag{23}$$

Finally, the effective dominant pole can be written as

$$\frac{1}{P_f} = \frac{1}{P_{fd}} - \frac{1}{Z_{fd}}. \tag{24}$$

The fifth step is to calculate the fall or rise times from the corresponding effective dominant pole. In the case of *NOR* gate considered above, since the output voltage is of single-pole response with the fall pole p_f expressed in (24), its fall time T_F can be easily expressed as

$$T_F = (\ln 9)/P_f \tag{25}$$

Based upon the five model formulation steps described above, the rise time of the characteristic waveform in a CMOS 3-input NOR gate can be characterized. The resultant expression is

$$T_R = (\ln 9)/P_r = (\ln 9)\left(\frac{1}{P_{rd}} - \frac{1}{Z_{rd}}\right) \tag{26a}$$

$$\frac{1}{P_{rd}} = (C_2' + C_3')\left(\frac{1}{\beta_{p3D}'} + \frac{1}{\beta_{p2D}'}\frac{\beta_{p3S}'}{\beta_{p3D}'} + \frac{1}{\beta_{p1D}'}\frac{\beta_{p2S}'}{\beta_{p2D}'}\frac{\beta_{p3S}'}{\beta_{p3D}'}\right) \cdot \frac{1}{2}\left[C_4'\left(\frac{1}{\beta_{p2D}'} + \frac{1}{\beta_{p1D}'}\frac{\beta_{p2S}'}{\beta_{p2D}'}\right) + (C_5' + C_6')\frac{1}{\beta_{p1D}'}\right] \tag{26b}$$

$$\frac{1}{Z_{rd}} = \frac{1}{2}\{C_4'/(\beta_{p2D}' + \beta_{p3S}') + (C_5' + C_6')/(\beta_{p1D}' + \beta_{p2S}') - C_2'/\beta_{pl}V_{DD} - C_6'\,\beta_{p3S}'\,\beta_{p2S}'/[(\beta_{p2D}' + \beta_{p3S}') \cdot (\beta_{p1D}' + \beta_{p2S}')\,\beta_{pl}V_{DD}]\}. \tag{26c}$$

In (26), the capacitance Ci' can be similarly expressed as Ci in the case of fall time.

The conductance parameters in (26) can be expressed as

$$\beta_{p3D}' = \beta_{p3}\left[\frac{3\eta_{p3}}{4}V_{DD} - V_{BINp3}' - (\eta_{p3} - 1)\,V_{DD} - \frac{2}{3}\gamma_{Sp3}\left(2\phi_{Fn} + \frac{1}{2}V_{DD}\right)^{1/2}\right] \tag{27a}$$

$$\beta_{p2D}' = \beta_{p2}\left[\frac{\eta_{p2}}{2}X_{tr2} - V_{BINp2}' - (\eta_{p2} - 1)\,V_{DD} - \frac{2}{3}\gamma_{Sp2}\,(2\phi_{Fn} + V_{DD} - V_{R2})^{1/2}\right] \tag{27b}$$

$$\beta_{p1D}' = \beta_{p1}\left[\frac{\eta_{p1}}{2}X_{tr1} - V_{BINp1}' - (\eta_{p1} - 1)\,V_{DD} - \frac{2}{3}\gamma_{Sp1}\,(2\phi_{Fn} + V_{DD} - V_{R1})^{1/2} - V_{DD}\,2^{-T_R/T_F}\exp(-p_f t_{dr})\right] \tag{27c}$$

$$\beta_{p3S}' = \beta_{p3}\left[\frac{\eta_{p3}}{2}X_{tr2} - V_{BINp3}' - (\eta_{p3} - 1)\,V_{DD} - \frac{2}{3}\gamma_{Sp3}\,(2\phi_{Fn} + V_{DD} - V_{R2})^{1/2}\right] \tag{27d}$$

$$\beta_{p2S}' = \beta_{p2}\left[\frac{\eta_{p2}}{2}X_{tr1} - V_{BINp2}' - (\eta_{p2} - 1)\,V_{DD} - \frac{2}{3}\gamma_{Sp2}\,(2\phi_{Fn} + V_{DD} - V_{R1})^{1/2}\right] \tag{27e}$$

where

$$X_{tr2} \equiv 2V_{DD} - \tfrac{1}{2}(V_{DD} - V_{Tpf})\exp[-p_r(t_{dr} - t_{dr3})]$$

$$X_{tr1} \equiv 2V_{DD} - \tfrac{1}{2}(V_{DD} - V_{Tpf})\exp[-p_r(t_{dr} - t_{dr2})]$$

$$V_{R2} \equiv V_{DD} - \tfrac{1}{2}(V_{DD} - V_{Tpf})\exp[-p_r(t_{dr} - t_{dr3})]$$

$$V_{R1} \equiv V_{DD} - \tfrac{1}{2}(V_{DD} - V_{Tpf})\exp[-p_r(t_{dr} - t_{dr2})].$$

The mobility μ_{sp1} in the parameter β_{p1}' is evaluated at $V_{GS} = V_{DD}$ and $V_{DS} = V_{DD}/6$.

So far, the rise and fall times expressions have been explicitly formulated. The sixth step in the model formulation is to calculate the propagation delay time of the logic gate. Generally, the rise (fall) propagation delay T_{PLH} (T_{PHL}) is defined as the time interval from the point the input voltage lowers (raises) to $V_{DD}/2$ to the point the output voltage raises (lowers) to $V_{DD}/2$, as indicated in Fig. 2 for T_{PHL}. Based upon this definition, the delay times T_{PLH} and T_{PHL} can be expressed as

$$T_{PLH} = t_{dr} + \frac{\ln 2}{\ln 9} T_R - \frac{\ln 2}{\ln 9} T_F \tag{28}$$

$$T_{PHL} = t_{df} + \frac{\ln 2}{\ln 9} T_F - \frac{\ln 2}{\ln 9} T_R \tag{29}$$

where t_{dr} and t_{df} are the initial delay as indicated in Fig. 2, and T_F *and* T_R have been expressed in (25) and (26), respectively.

The pair delay of characteristic waveforms defined as the delay of two successive stages, can be expressed as

$$T_p \equiv T_{PLH} + T_{PHL} = t_{dr} + t_{df}. \tag{30}$$

III. Comparisons to Computer Simulation Results and Error Analyses

According to our observations, the initial delay at the output node denoted by t_{dr} or t_{df}, is mainly due to the voltage overshoot or undershoot caused by the capacitance feedthrough effect. For a better understanding, consider the equivalent circuit of Fig. 4. When the input voltage is raised from 0 V, this input voltage change is fed forward to the output node by the capacitance feedthrough path formed by the capacitance C_2 and the capacitance at the output node. At this time, the output voltage has an overshoot from its initial value V_{DD}. Therefore, the output voltage can not be lowered instantly when the input voltage is applied. This results in the initial delay time which is defined as the time the output voltage lowers back from its overshoot to V_{DD}. The above concept has been confirmed in modeling the initial delay of CMOS inverters under step inputs [18].

Based on the above mentioned concept, it is possible to develop models for all the initial delay times, namely, t_{df}, t_{df2}, t_{df3}, t_{dr}, t_{dr2}, t_{dr3}, etc. Then all the timing formulas derived in the previous section can be directly solved. We shall study this approach in the future.

Here, the initial delay times are determined semi-empirically. Thus the resultant timing formula can be more easily solved but still with reasonable accuracy.

Consider a logic gate driven by a slower rising input voltage which has a smaller rise pole P_r. In this case, the overshoot due to the capacitance feedthrough effect becomes larger. This leads to a larger t_{df}. Therefore the product $p_r t_{df}$ is expected to be nearly constant. So is the product $p_f t_{dr}$. These predictions have been verified for logic gates with a wide range of device channel dimensions, device parameter values, or load capacitances. In the case of CMOS 3-input NOR gates, the constant values of $P_r t_{df}$ and $p_f t_{dr}$ have been determined from the simulation results as $p_r t_{df} = p_f t_{dr} = 0.5$. These universal values have also been used for NOR gates with different fan-in number from 2 to 4.

The difference of delay times, such as $(t_{df} - t_{df2})$ and $p_f(t_{df} - t_{df3})$ have been found to be strongly related to the fall waveform of the output voltage v_0. If the fall pole p_f is smaller, i.e., the falling speed is slower, the voltage at the internal node v_{02} or v_{03} will be lowered more slowly. Thus the delay time difference $(t_{df} - t_{df2})$ or $(t_{df} - t_{df3})$ will be larger. Therefore, the products $p_f(t_{df} - t_{df2})$ or $(t_{df} - t_{df3})$ can be set to some universal constants. In the case of CMOS NOR gates, these constants are $p_f(t_{df} - t_{df2}) = -0.134$, $p_f(t_{df} - t_{df3}) = -0.113$, $p_r(t_{dr} - t_{dr2}) = 0.094$, and $p_r(t_{dr} - t_{dr3}) = 0.059$. These constants are universal for NOR gates with two to four inputs and different device channel dimensions, device parameters, or capacitive loads.

Now, the expression of T_F contains only a variable T_F/T_R whereas the expression of T_R contains a variable T_R/T_F. Therefore, a nonlinear equation of the variable T_F/T_R or T_R/T_F can be obtained by dividing one of these two expressions by the other. After a few number of iterations, T_F/T_R or T_R/T_F can be solved. Using the calculated T_F/T_R or T_R/T_F, T_F and T_R can be quickly calculated from their respective expressions.

Using the calculated values of T_F and T_R, the initial delay time of 3-input NOR gates can be calculated by using the following empirical laws:

$$t_{df} = 0.18\ T_R + 0.22\ T_F \tag{31a}$$

$$t_{dr} = 0.07\ T_R + 0.58\ T_F. \tag{31b}$$

Note that the above laws are universal for all types of CMOS NOR gates. Therefore, they can be determined either from computer simulations or actual experimental measurements.

Substituting (31) into (28)–(30), we can obtain the universal semi-empirical formulas for delay times. These formulas can be used in the calculations of delay times for various CMOS NOR gates with satisfactory accuracy, as will be seen later.

Before the comparisons between theoretical calculation and ESPICE simulations are made, the error sources of the proposed timing models are first investigated. Generally, there are three kinds of error sources. The first is the capacitance estimation error introduced in evaluating various voltage-dependent MOSFET channel capacitances and p-n junction capacitances by fixed values. The second is the drain current estimation error introduced from fixed mobilities, operating region determinations, and drain current linearization techniques. The last is the dominant-pole approximation error introduced from the calculation of dominant pole and the assumption of single-pole response.

In the case of large fixed capacitance load C_L, all the internal device capacitances become negligibly small as compared with C_L. Thus the error from the capacitance

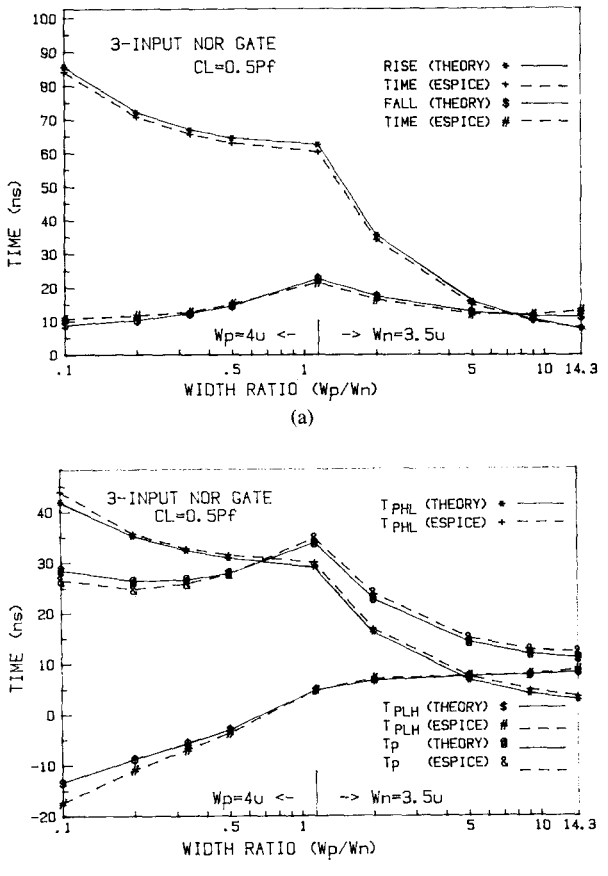

Fig. 5. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay and pair delay of characteristic waveforms for a CMOS 3-input NOR gate with C_L = 0.5 PF.

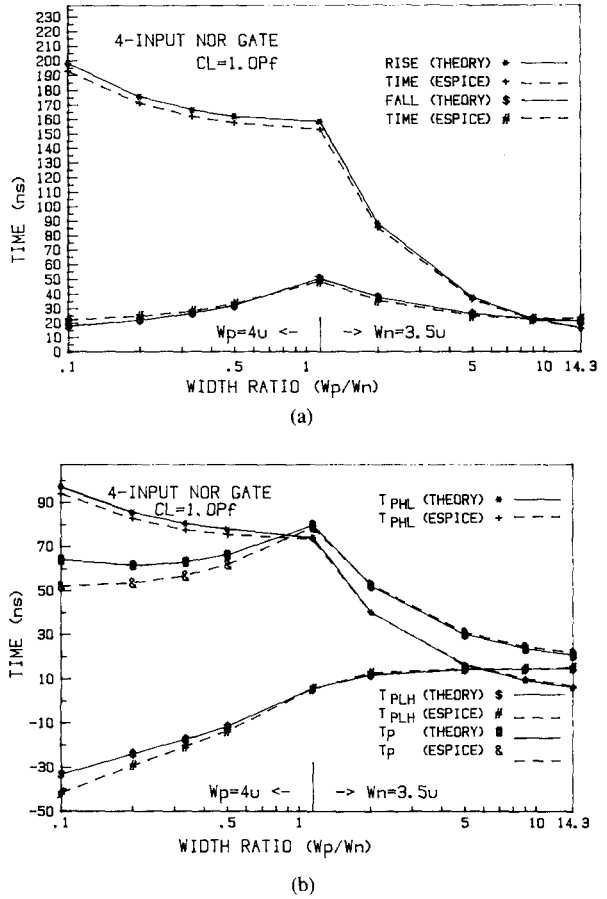

Fig. 6. Calculated and simulated (a) rise time, fall time; (b)rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS 4-input NOR gate with C_L = 1 PF.

estimation and the dominant-pole approximation is quite small. The only error source is the current estimation. For the 3-input NOR gate with a large C_L, the capacitances C_3 and C_3' become the most dominant capacitances. From (22) to (26) the expressions for T_F and T_R become

$$T_F \simeq (\ln 9)\, C_L/\beta_{n1D} \tag{32a}$$

$$T_R \simeq (\ln 9)\, C_L \left[\frac{1}{\beta'_{p3D}} + \frac{1}{\beta'_{p2D}}\frac{\beta'_{p3S}}{\beta'_{p3D}} + \frac{1}{\beta'_{p1D}}\frac{\beta'_{p2S}}{\beta'_{p2D}}\frac{\beta'_{p3S}}{\beta'_{p3D}}\right]. \tag{32b}$$

The error in T_F and T_R thus represents the error due to various approximations made on various conductance factors. Fig. 5(a) shows the comparisons between calculations and ESPICE simulations for the characteristic waveform timing of 3-input NOR gates with effective channel length L_{effn} = 1.9 μm, L_{effp} = 2.26 μm and C_L = 0.5 PF. The maximum error in T_F and T_R is limited to be below 15 percent for W_p/W_N = 4 μm/40 μm to W_P/W_N = 50 μm/3.5 μm, or equivalently for $T_R/T_F \simeq$ 10 to $T_R/T_F \simeq$ 0.5. Outside this range, the characteristic waveforms severely deviate from those shown in Fig. 2. Thus the timing error will be increased. For larger C_L or larger channel width, similar error characteristics have also been found in the same range of W_p/W_N or T_R/T_F. Generally, the values of T_F and T_R are found to be increased linearly with C_L, as predicted in (32).

For the rise delay time T_{PLH} and the fall delay time T_{PHL}, the maximum error exceeds 15 percent when the time is small, as may be seen from Fig. 5(b). However, this larger error will be diminished in the multi-stage delay time which is the sum of several rise and fall delay times, simply because the smaller delay times have only smaller contributions in the whole multi-stage delay time. This prediction has been confirmed by error characteristics of the pair delay time T_p shown in Fig. 5(b) for 3-input NOR gates with C_L = 0.5 PF. In that figure, maximum error of 10 percent is found between simulation results and calculation results from a simple universal equations.

The same error characteristics have also been observed in the cases of 2- and 4-input CMOS NOR gates. Fig. 6(a) and (b) show the comparisons for four-input NOR gates with C_L = 1 PF.

We had also derived the timing models for CMOS inverters and CMOS NAND gates with fan-in numbers up to 4. According to our observations on the comparisons between the calculations and computer simulations, similar error characteristics described above have also been found. Typical comparisons are shown in Fig. 7(a) and (b) for 4-input NAND gates with C_L = 1 PF.

From the above error analyses, it is concluded that the

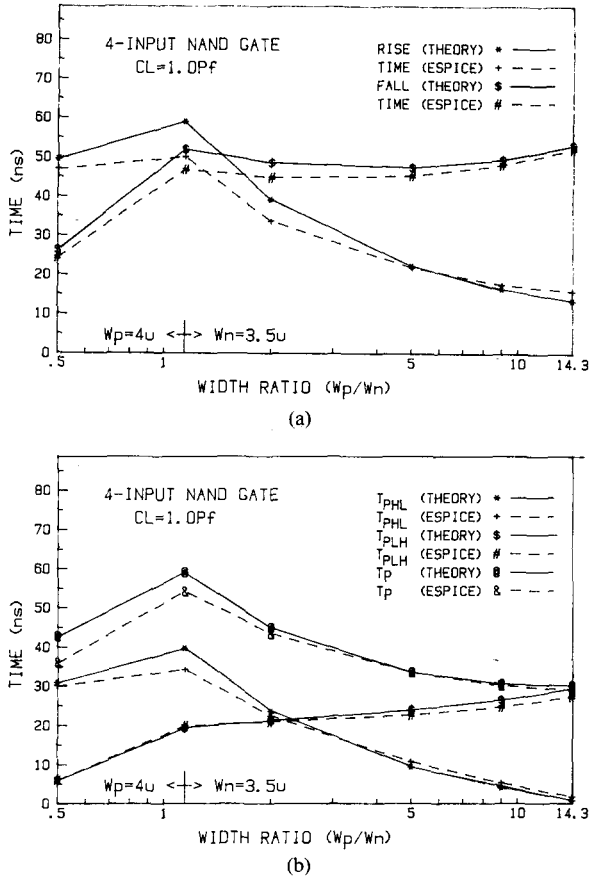

Fig. 7. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS four-input NAND gate with C_L = 1 PF.

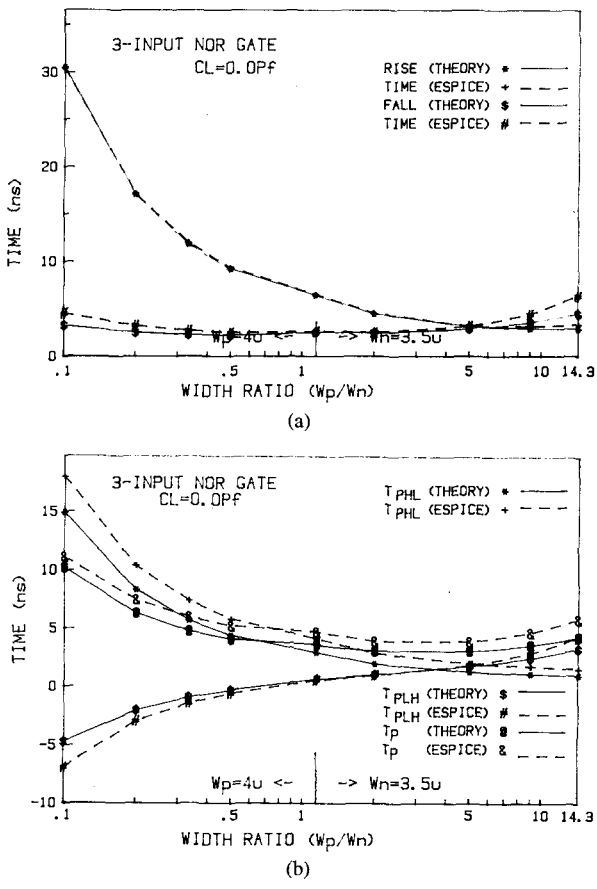

Fig. 8. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS three-input NOR gate with C_L = 0.

error introduced by the drain current approximations used in our timing model is below 15 percent. Therefore, under these approximations, the timing error for CMOS combinational gates with large C_L can be limited to be below 15 percent.

In the case of large fan-out number N such that the load capacitance is larger than internal capacitances, the dominant pole exists and the error sources are the current estimation and the capacitance estimation. It is found that the timing error in this case is limited to be under 25 percent. Therefore, the maximum error due to the capacitance estimation error is about 10 percent.

The worst-case error occurs when the capacitive load is not larger than internal gate capacitances such that the dominant pole does not exist. This is because that three error sources appear in these worst cases. Since the error due to the modified dominant-pole approximation is below 10 percent, the worse-case error in the proposed timing model can be confined to be below 35 percent.

One of these worst cases is that of the multi-input logic with no fixed capacitance load and with one fan-out gate. Fig. 8(a) and (b) show the calculated and the simulated signal timings of 3-input NOR gates whereas Fig. 9(a) and (b) show those of 2-input NOR gates. The maximum error is about 30 percent for the rise or fall times and is about 25 percent for the pair delay time. Similar error characteristics are also found in multi-input NAND gates with $N = 1$ and $C_L = 0$. Typical results are shown in Fig. 10(a) and (b) for CMOS 4-input NAND gates.

In the CMOS inverter case, the error is quite small because the dominant pole always exists as may be seen from its equivalent circuit. Typical inverter timing characteristics are shown in Fig. 11(a) and (b).

Under device parameter variations, the error characteristics of the timing model are found to be the same, being below 35 percent. As an example, the signal timing of CMOS 3-input NOR gates with increased V_{T0} and decreased mobility is shown in Fig. 12(a), (b), and Fig. 13(a) and (b), respectively.

In actual IC's, internal waveforms may not be characteristic waveforms. In these cases, the timing model can also be applied to characterize the signal timing with reasonable error if the waveforms do not deviate much from the characteristic waveforms. For example, under step input excitations, the worst-case error of the rise and fall times for the 3-input NOR gates with $C_L = 0$ is nearly 40 percent, as may be seen from Table II. Since the delay time in this case is very small, its error is larger. But for multi-stage delay, this large error will be diminished, as mentioned before. On the other hand, under slow exponential input excitation with time constant 5 ns, the maximum error of the rise and fall times is nearly 23 percent

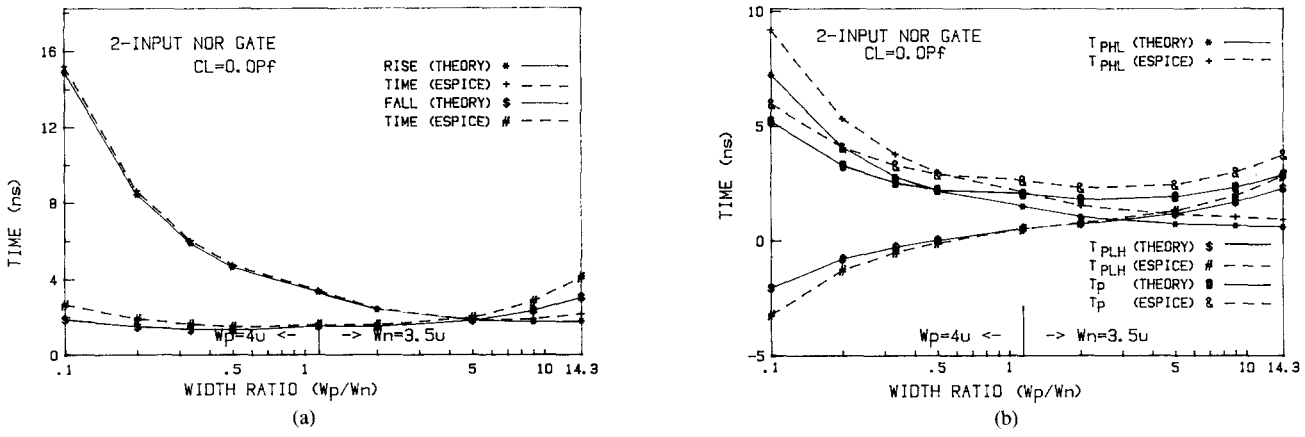

Fig. 9. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS 2-input NOR gate with $C_L = 0$.

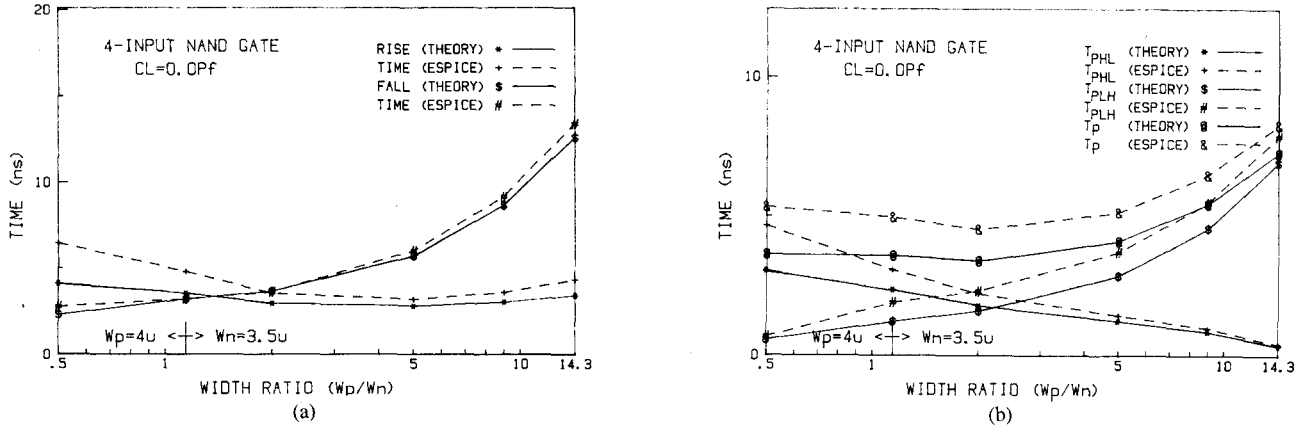

Fig. 10. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS 4-input NAND gate with $C_L = 0$.

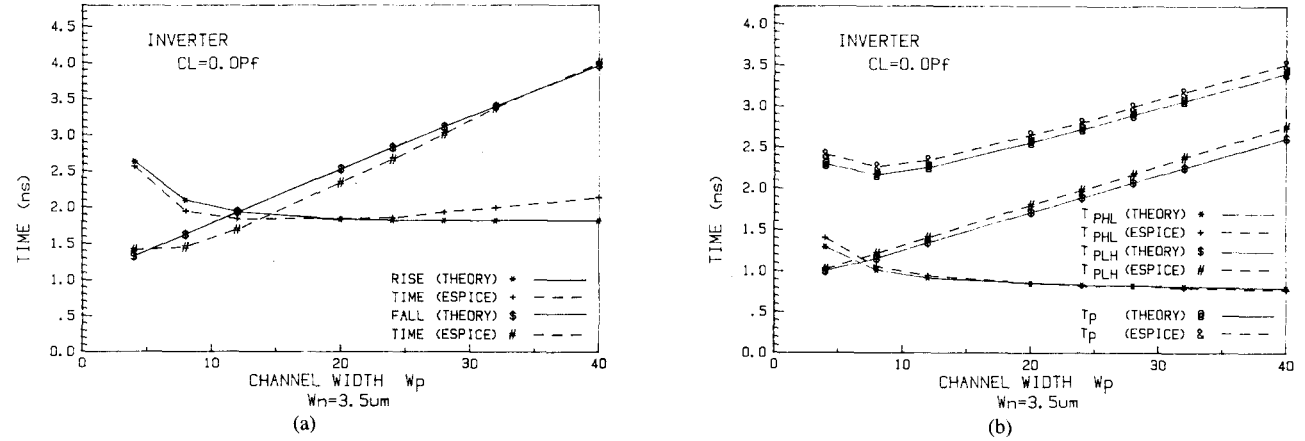

Fig. 11. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS inverter with $C_L = 0$.

and that of the delay time is larger, as may be seen from Table III. Note that the error characteristic described above is the worst-case one with $C_L = 0$. If C_L is added, better accuracy can be obtained. Thus the proposed timing model is versatile and is applicable in circuit timing analysis.

To study the actual application of our timing model in timing analysis, an experimental program called TISA [14] has been written to analyze the signal timing of CMOS combinational logic gates. Running TISA and ESPICE on a VAX-11/780 computer, the results are listed in Tables IV

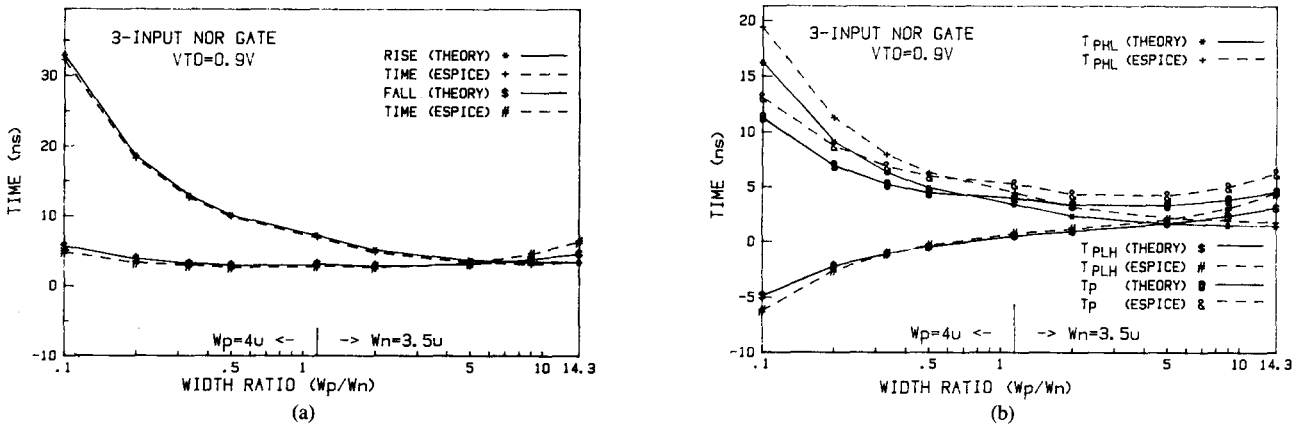

Fig. 12. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS 3-input NOR gate with $C_L = 0$ and $|V_{TO}| = 0.9$ V, a 0.205-V increase from that of Fig. 8.

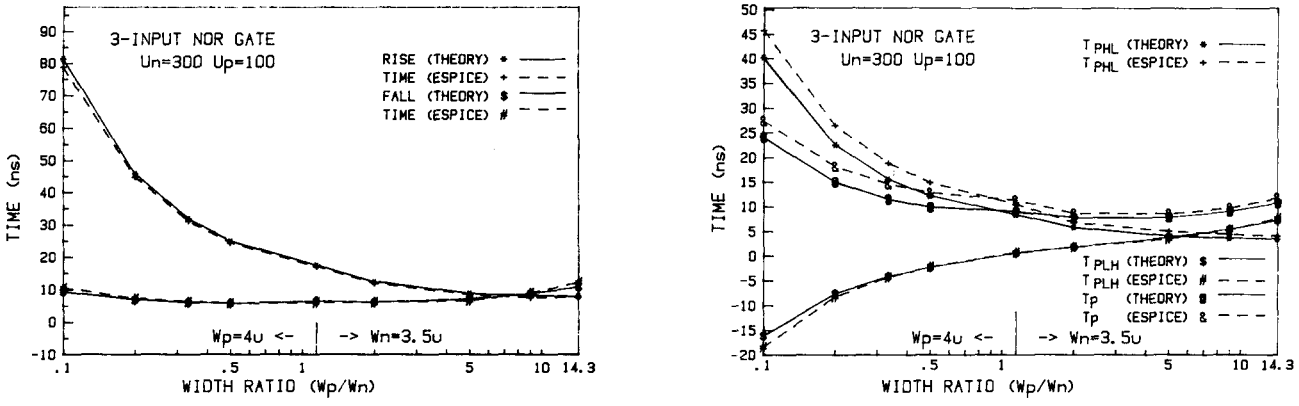

Fig. 13. Calculated and simulated (a) rise time, fall time; (b) rise delay, fall delay, and pair delay of characteristic waveforms for a CMOS three-input NOR gate with $C_L = 0$ and decreased mobility. The value of UO parameter for NMOS is decreased from 681 in Fig. 8 to 300 whereas that for PMOS is decreased from 262 in Fig. 8 to 100.

TABLE II
RISE TIME, FALL TIME, RISE DELAY AND FALL DELAY OF A CMOS 3-INPUT NOR GATE WITH $C_L = 0$ AND DRIVEN BY STEP INPUT

	W_p (μm)	4.0	4.0	4.0	4.0	4.0	7.0	17.5	31.5	50.0
	W_n (μm)	40.0	20.0	12.0	8.0	3.5	3.5	3.5	3.5	3.5
Rise Time (ns)	THEORY	29.59	16.76	11.65	9.122	6.314	4.311	2.839	2.422	2.232
	ESPICE	30.41	17.36	11.97	9.352	6.402	4.418	2.908	2.481	2.303
	ERROR(%)	2.69	3.46	2.67	2.46	1.37	4.42	2.37	2.37	3.08
Fall Time (ns)	THEORY	0.291	0.376	0.486	0.618	1.095	1.322	1.960	2.671	3.508
	ESPICE	0.376	0.399	0.457	0.515	0.877	1.249	2.521	4.008	5.932
	ERROR(%)	22.61	5.76	-6.35	-20.00	-24.86	-5.84	21.94	33.36	40.86
Rise Delay T_{PLH}(ns)	THEORY	14.80	8.378	5.827	4.561	3.157	2.155	1.419	1.211	1.116
	ESPICE	18.37	10.37	7.278	5.662	3.840	2.567	1.653	1.396	1.279
	ERROR(%)	19.43	19.21	19.94	19.44	17.79	16.05	14.16	13.25	12.74
Fall Delay T_{PHL}(ns)	THEORY	0.262	0.339	0.437	0.565	0.985	1.19	1.771	2.403	3.157
	ESPICE	0.169	0.185	0.206	0.235	0.362	0.52	1.232	2.278	3.610
	ERROR(%)	-55.03	-83.24	-112.14	-136.60	-172.10	-128.41	-43.75	-5.49	12.55

TABLE III
RISE TIME, FALL TIME, RISE DELAY AND FALL DELAY OF A 3-INPUT CMOS NOR GATE WITH $C_L = 0$ AND DRIVEN BY EXPONENTIAL INPUT EXCITATION WITH TIME CONSTANT 5 ns

	ω_p(μm)	4.0	4.0	4.0	4.0	4.0	7.0	17.5	31.5	50.0
	ω_n(μm)	40.0	20.0	12.0	8.0	3.5	3.5	3.5	3.5	3.5
Rise	THEORY	30.18	17.98	13.21	10.84	8.017	6.150	4.569	4.097	3.879
Time	ESPICE	30.70	17.63	12.48	10.04	7.523	5.936	4.703	4.237	3.957
(ns)	ERROR (%)	1.69	-1.99	-5.85	-7.97	-6.57	-3.61	2.85	3.30	1.97
Fall	THEORY	1.835	1.986	2.176	2.410	3.253	3.608	4.593	5.629	6.776
Time	ESPICE	1.939	1.951	2.116	2.460	3.433	3.807	5.753	7.072	8.753
(ns)	ERROR(%)	5.36	-1.79	-2.84	2.03	5.24	5.23	20.16	20.40	22.59
Rise	THEORY	13.99	7.891	5.507	4.321	2.987	1.977	1.187	0.951	0.841
Delay	ESPICE	21.31	13.67	10.33	8.434	5.882	4.161	2.443	1.829	1.532
T_{PLH}(ns)	ERROR(%)	34.35	42.28	46.69	48.77	49.22	52.49	51.41	48.00	45.10
Fall	THEORY	-1.093	-0.956	-0.787	-0.576	0.183	0.503	1.389	2.231	3.353
Delay	ESPICE	-1.140	-0.990	-0.791	-0.565	0.317	0.979	2.833	4.690	6.517
T_{PHL}(ns)	ERROR(%)	4.12	3.43	0.50	-1.95	42.71	48.62	50.97	50.51	48.55

TABLE IV
TIMING DATE OBTAINED FROM TISA AND ESPICE FOR A CHAIN OF 5-STAGE CMOS INVERTERS WITH DIFFERENT CAPACITIVE LOADS AND DRIVEN BY STEP INPUT

		ESPICE (ns)	TISA (ns)	ERROR (%)
0.2PF	T_F	3.42	2.55	25.4
	T_{PT}	2.16	1.61	25.5
0.4PF	T_R	5.94	5.89	0.8
	T_{PT}	5.72	4.50	21.3
0.8PF	T_F	11.05	10.04	9.1
	T_{PT}	12.40	11.20	9.6
1.6PF	T_R	21.44	21.60	-0.7
	T_{PT}	25.08	21.79	13.1
3.2PF	T_F	42.02	37.93	9.7
	T_{PT}	50.12	47.08	6.1
CPU Time (sec)		60.11	1.32	

ω_P = 12μm; ω_N=3.5 μm; step input excitation

T_{PT}: Total delay time

T_R : Rise time

T_F : Fall time

TABLE V
TIMING DATA OBTAINED FROM TISA AND ESPICE FOR A CHAIN OF 12-STAGE CMOS 3-INPUT NOR GATES WITH DIFFERENT CAPACITIVE LOADS AND DRIVEN BY STEP INPUT

		ESPICE (ns)	TISA (ns)	ERROR (%)
0.2 PF	T_F	4.74	3.14	33.7
	T_{PT}	4.26	2.82	33.8
0.4 PF	T_R	50.9	50.0	1.8
	T_{PT}	29.5	27.7	6.1
0.6 PF	T_F	28.7	22.2	22.6
	T_{PT}	40.0	34.9	12.8
0.8 PF	T_R	96.8	94.9	2.0
	T_{PT}	87.0	80.2	7.8
1.0 PF	T_F	49.9	38.3	23.2
	T_{PT}	102.8	90.9	11.6
2.0 PF	T_R	231.1	226.9	1.8
	T_{PT}	215.8	200.5	7.1
1.0 PF	T_F	78.8	60.6	23.1
	T_{PT}	217.4	198.3	8.8
0.8 PF	T_R	105.8	100.8	4.7
	T_{PT}	266.4	242.7	8.9
0.6 PF	T_F	41.5	31.4	24.3
	T_{PT}	272.0	245.8	9.6
0.4 PF	T_R	56.2	53.7	4.4
	T_{PT}	298.0	269.4	9.6
0.2 PF	T_F	19.3	14.6	24.4
	T_{PT}	298.5	269.2	9.8
0.8 PF	T_R	95.0	93.6	1.5
	T_{PT}	345.1	314.5	8.9
CPU Time (sec)		604.66	2.36	

W_P=4.0 μm; W_N=3.5 μm; step input excitation

T_{PT}: Total delay time; T_R: Rise time; T_F: Fall time

and V. In Table IV, the analyzed circuit is a chain of 5-stage identical inverters with different capacitive loads and under step-input excitation. In this case, internal waveforms are not characteristic waveforms. However, the error in the rise or fall time is below 25 percent whereas the error in the total delay time is reduced from 25.5 percent in the first stage to 6.1 percent in the last stage. The anal-

ysis time in CPU seconds of TISA is 45 times less that that of ESPICE.

In Table V, the signal timing of a chain of 12-stage identical 3-input NOR gates with different capacitive loads and under step-input excitation is listed. The device channel dimensions and capacitive load of each NOR gate in the chain are also listed in the table. Although the internal waveforms are not characteristic waveforms, the general error is below 34 percent whereas the error in the total delay time is reduced from 33.8 percent in the first stage to 9 percent in the last stage. The analysis time of TISA is about 250 times less than that of ESPICE.

From the above analyses, it is clear that the required CPU time and memory in the new timing analysis program is small because the signal timing is characterized by direct calculation rather than by point-by-point calculation in the time domain. Moreover, the error of the total delay time in an IC is reasonably small.

IV. Conclusion and Discussion

Based upon the s-domain analysis using the modified dominant-pole calculation method on the linearized equivalent circuit, the signal timing of a complex CMOS combinational logic gate has been explicitly expressed. Error analysis has shown that maximum error can be confined to 35 percent for inverters, multi-input NOR gates or multi-input NAND gates with effective channel length down to 1.9 μm and with different capacitive loads. Under considerable device parameter variations or waveform deviations, the timing model can also be applied with reasonable error characteristics. Applying the timing model in an experimental CAD program leads to faster timing analysis with CPU time being over an order of magnitude less than ESPICE for complex circuits.

From this work, it is felt that direct modeling a gate rather than modeling a device will be very helpful both in speed performance optimization and in circuit analysis of a digital IC. Using such models in a CAD program, the required CPU time and memory will be greatly decreased. Therefore, the approach of modeling a gate is worth developing especially in the VLSI era.

Further extension of the timing model will be done in three major areas. First, the signal timing of the CMOS gates with very short channel devices will be characterized by considering the velocity saturation effect. Second, the signal timing in non-worst-case timing case in which the input excitation does not enter the gate from the input node farthest to the output node will be modeled. Finally, the signal timing of NMOS gates and other CMOS gates will be characterized. Continuing efforts to improve the accuracy of the timing model will also be made.

Appendix

Using the expressions of $v_i(t)$ and $v_0(t)$ in (2) and (3), the voltage quadratic terms $v_i v_0(t)$ and $v_0^2(t)$ can be expressed as

$$\begin{aligned} v_i v_0(t) &= V_{DD} v_0(t) - V_{DD} \exp(-p_r t) \\ &\quad \cdot [u(t) - u(t - t_{dr})] - [V_{DD}^2 \exp \\ &\quad \cdot [-P_f(t - t_{df})]\, u(t - t_{df}) \\ &\quad \cdot \exp(-p_r t) \end{aligned} \tag{A1}$$

$$\begin{aligned} v_0^2(t) &= V_{DD}^2 u(t) - V_{DD}^2 u(t - t_{df}) + \{V_{DD}^2 \\ &\quad \cdot \exp[-p_f(t - t_{df})]\}\, u(t - t_{df}) \\ &\quad \cdot \exp[-p_f(t - t_{df})]. \end{aligned} \tag{A2}$$

In the above equations, the terms containing the product of two exponential functions must be linearized to eliminate one of the two exponential functions. The linearization technique adopted here is to evaluate one of the exponential functions at the output voltage crossover point, i.e., at $t = t_e$ where $v_0(t)$ is equal to $V_{DD}/2$. From (3), t_e may be solved as in (7) in the text.

Substituting (7) into the expression of $\exp(-p_f t)$ in the last term of (A1), $v_i v_0(t)$ when $t > t_{df}$ can be linearized as

$$\begin{aligned} v_i v_0(t) &= V_{DD} v_0(t) - [V_{DD} 2^{-p_r/p_f} \cdot \exp(-p_r t_{df})]\, v_0(t) \\ &= [V_{DD} - V_{DD} 2^{-p_r/p_f} \exp(-p_r t_{df})]\, v_0(t). \end{aligned} \tag{A3}$$

It is found that if the constant 2 in (A3) is replaced by 2.8, the error can be reduced. Thus we have

$$\begin{aligned} v_i v_0(t) &= F_o v_0(t) \\ &= [V_{DD} - V_{DD}(2.8)^{-T_F/T_R} \\ &\quad \cdot \exp(-p_r t_{df})]\, v_0(t). \end{aligned} \tag{A4}$$

Substituting (7) into the expression of $\exp[-p_f(t - t_{df})]$ in the last term of (A2), $v_0^2(t)$ after $t > t_{df}$ becomes

$$v_0^2(t) = \tfrac{1}{2} V_{DD} v_0(t) \tag{A5}$$

Using $v_{03}(t)$ in (4), the term $v_{03}^2(t)$ can be explicitly expressed as

$$\begin{aligned} v_{03}^2(t) &= V_{DD}^2 u(t) + [V_{Tpf}^2 - V_{DD}^2]\, u(t - t_{df3}) \\ &\quad + 2V_{Tpf}(V_{DD} - V_{TPf}) \exp[-p_f(t - t_{df3})] \\ &\quad \cdot u(t - t_{df3}) + \{(V_{DD} - V_{Tpf})^2 \exp[-p_f(t - t_{df3})] \\ &\quad \cdot u(t - t_{df3})\} \exp[-p_f(t - t_{df3})]. \end{aligned} \tag{A6}$$

Substituting (7) into the expression of $\exp[-p_f(t - t_{df3})]$ in the last term of (A6), $v_{03}^2(t)$ is linearized as

$$v_{03}^2(t) = B_o v_{03}(t) + V_{DD}^2 u(t) + [V_{Tpf}^2 - V_{DD}^2]\, u(t - t_{df3}) \tag{A7}$$

where the constant B_o can be written as

$$B_o = 2V_{Tpf} + \tfrac{1}{2}(V_{DD} - V_{Tpf}) \exp[-p_f(t_{df} - t_{df3})] \tag{A8}$$

After discarding the two constant terms in (A8), the expression of $v_{03}^2(t)$ is just as (10) in the text.

Similarly, $v_{02}^2(t)$ can be linearlized as

$$v_{02}^2(t) = \{2V_{Tpf} + \tfrac{1}{2}(V_{DD} - V_{Tpf}) \exp \cdot [-p_f(t_{df} - t_{df2})]\}\, v_{02}(t) \equiv C_o v_{02}(t). \quad \text{(A9)}$$

REFERENCES

[1] B. R. Chawla, H. K. Gummel, and P. Kozak, "MOTIS—An MOS timing simulator," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 901–910, Dec. 1975.

[2] A. R. Newton, "The simulation of large scale integrated circuits," *IEEE Trans. Circuits Syst.*, vol. CAS-26, pp. 741–749, Sept. 1979.

[3] E. Lelarasmee, A. E. Ruehli, and A. L. Sangiovanni-Vincentelli, "The waveform relaxation method for time-domain analysis of large scale integrated circuits," *IEEE Trans. Computer-Aided Design*, vol. CAD-1, pp. 131–145, July 1982.

[4] J. R. Burns, "Switching response of complementary-symmetry MOS transistor logic circuits," *RCA Rev.*, pp. 627–661, Dec. 1964.

[5] W. M. Penny and L. Lau, *MOS Integrated Circuit*. New York: Van Nostrand-Reinhold, 1972.

[6] E. Seewann, "Switching speed of MOS inverters," *IEEE J. Solid-State Circuits*, vol. SC-15, pp. 246–252, Apr. 1980.

[7] L. A. Glasser, "The analog behavior of digital integrated circuits," in *Proc. 18th Design Automation Conf.*, pp. 603–612, 1981.

[8] A. Kanuma, "CMOS circuit optimization," *Solid-State Electronics*, vol. 26, pp. 47–58, Jan. 1983.

[9] T. Tokuda, K. Okazaki, K. Sakashita, I. Ohkura, and T. Enomoto, "Delay-time modeling for ED MOS logic LSI," *IEEE Trans. Computer-Aided Design*, vol. CAD-2, pp. 129–134, July 1983.

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

[11] T. Sakurai, "Approximation of wiring delay in MOSFET LSI," *IEEE J. Solid-State Circuits*, vol. SC-18, pp. 418–426, Aug. 1983.

[12] R. Putatunda, "Auto-Delay: A program for automatic calculation of delay in LSI/VLSI chips," in *Proc. 19th Design Automation Conf.*, pp. 616–621, 1982.

[13] S. M. Kang, "A design of CMOS polycells for LSI circuits," *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 838–843, Aug. 1981.

[14] Chung-Yu Wu and Jen-Sheng Hwang, to be published.

[15] M. I. Elmasry, "Digital MOS integrated circuits: A tutorial," in *Digital MOS Integrated Circuits*, New York: IEEE Press, Mar. 1981, pp. 4–27.

[16] A. Vladimirescu and S. Liu, "The simulation of MOS integrated circuits using SPICE2," Univ. California, Berkeley, Electronic Res. Lab. Memo. UCB/ERL-M80/7, Oct. 1980.

[17] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*. New York: Wiley, 1977, p. 412.

[18] Chung-Yu Wu and Yeu-Haw Yang, to be published.

*

Chung-Yu-Wu (S'75-M'77) was born in Chiayi, Taiwan, Republic of China, on May 31, 1950. He received the B.S. degree from the Department of Electrophysics, and the M.S. and Ph.D. degrees from Institute of Electronics, National Chiao-Tung University, Hsinchu, Taiwan, Republic of China, in 1972, 1976, and 1980, respectively.

During 1975–1976, he studied ferroelectric films on silicon and their device applications. During 1976–1979, he has been engaged in the development of integrated differential negative resistance devices and their circuit applications, under the Grant support from the National Electronics Mass Plan (Semiconductor Devices and Integrated Circuit Technologies) of National Science Council. From 1980 to 1984 he was an Associate Professor in Institute of Electronics. Since 1984 he has been an Associate Professor in the Department of Electrical Engineering, Portland State University, Portland, Oregon. His research interests have been in device modeling, integrated-circuit technologies, semiconductor memories, analog integrated circuits, and timing analyses of digital integrated circuits.

*

Jen-Sheng Hwang was born in Kaushung, Taiwan, Republic of China. He received the B.S. and M.S. degree in electrical engineering from National Chiao Tung University, Hsin-Chu, Taiwan, in 1982 and 1984, respectively. Currently, he is working toward the Ph.D. degree at the Institute of Electronics, National Chiao Tung University. His research concerns MOS timing analysis and synthesis.

*

Chih Chang was born in Taiwan, R.O.C., on March 16, 1954. He received the B.S. and M.S. degrees in electrical engineering from National Cheng-Kung University, Taiwan, in 1976 and 1978, respectively.

After graduation, he joined Electronics Research and Service Organization, Industrial Technology Research Institute (ERSO/ITRI), Taiwan, R.O.C., where he was initially a product engineer for Digital Circuit Design Department. He worked on yield improvement of digital IC's and development of MOSFET modeling. He is currently the Manager of Standard Cell Design Section, responsible for the CMOS standard cell design for semicustoms.

*

Ching-Chu Chang received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1970 and the M.S. and Ph.D. degrees from Princeton University, Princeton, NJ, in 1973 and 1976, respectively, in the field of electrical engineering.

Since 1976, he has been with Electronics Research and Service Organization, Industrial Technology Research Institute (ERSO/ITRI). In 1976, he headed a group of engineers as trainees in the field of digital circuit testing at RCA. During 1977–1978, he was the technical leader of Product and Test Engineering Group, engaged in various tasks in supporting the production and testing of CMOS IC wafers. In 1978, he was assigned as the Manager of Bipolar Project Team dispatched by ITRI to RCA for the transfer of bipolar linear IC technology in the design, processing, and testing. His major work was to acquire the equipment information necessary to set up a bipolar product line. In 1979 he became the Manager of Product Development Group, supervising all the product development work in both MOS and bipolar fields. He is currently the Deputy Director of ERSO/ITRI, responsible for all the research and development of IC technology.

Mr. Chang is a member of Sigma Xi.