DELAY UNCERTAINTY DUE TO ON-CHIP SIMULTANEOUS SWITCHING NOISE IN HIGH PERFORMANCE CMOS INTEGRATED CIRCUITS

Kevin T. Tang and Eby G. Friedman
Department of Electrical and Computer Engineering
University of Rochester
Rochester, New York 14627-0231

Abstract – On-chip parasitic inductance inherent to the power supply rails has become significant in high speed digital circuits. Therefore, current surges result in voltage fluctuations within the power distribution networks, creating delay uncertainty. On-chip simultaneous switching noise should therefore be considered when estimating the propagation delay of a CMOS logic gate in high speed synchronous CMOS integrated circuits. Analytical expressions characterizing the on-chip simultaneous switching noise voltage and the output voltage waveform of a CMOS logic gate driving both a capacitive and a resistive-capacitive load are presented in this paper. The waveform of the output voltage signal based on the analytical expressions is quite close to SPICE. The estimated propagation delay is within 5% as compared to SPICE while the average improvement in accuracy can reach 10% as compared to a delay estimated without considering on-chip simultaneous switching noise. The analytical expressions presented here provide an accurate timing model for non-negligible on-chip simultaneous switching noise in high speed synchronous CMOS integrated circuits.

I. INTRODUCTION

The trend in next generation integrated circuit (IC) technology is towards higher speeds and densities. The total capacitive load associated with the internal circuitry is therefore increasing in both current and next generation VLSI circuits [1], [2], [3]. As the operating frequency increases, the average on-chip current required to charge (and discharge) these capacitances also increases, while the time during which the current is switched decreases [4]. Therefore, a large change of the on-chip current occurs within a short period of time.

The primary sources of the on-chip current surges are those logic gates that switch close in time to the clock edges in a synchronous CMOS integrated circuit. Because of the non-negligible parasitic inductance inherent to the on-chip power distribution network, large current surges result in voltage fluctuations in the power distribution network [5]. These voltage fluctuations are called on-chip simultaneous switching noise.

For example, at gigahertz operating frequencies and high integration densities, power dissipation densities are expected to approach $20\,W/cm^2$ [4], [6], a power density limit for an air-cooled packaged device. Such a power density is equivalent to 16.67 amperes of current for a 1.2 volt power supply in a 0.1 μm CMOS technology. Assuming that the current is uniformly distributed along a 1 cm wide and 1 μm thick Al-Cu interconnect plane, the average current density is approximately $1.67\,mA/\mu m^2$. For a standard mesh structured power distribution network, the current density is even greater than $1.67\,mA/\mu m^2$. For a 1 mm long power buss line with a parasitic inductance of 2 nH/cm [7], and an edge rate of the current signal on the order of an overly conservative nanosecond, the amplitude of the on-chip simultaneous switching noise voltage is approximately 0.35 volts. This peak noise is not insignificant in very deep submicrometer (VDSM) CMOS integrated circuits [8].

On-chip simultaneous switching noise affects the signal delay, creating delay uncertainty since the power supply level temporally changes the local drive current [8], [9]. Furthermore, logic malfunctions may be created and excess power may be dissipated due to faulty switching if the power supply fluctuations are sufficiently large [10], [11]. A modified tapered buffer design technique is presented

0-7803-6488-0/00/$10.00 © 2000 IEEE

by Vemuru in [9] which considers delay uncertainty caused by off-chip simultaneous switching noise. In addition to the many constraints encountered in the design of on-chip clock and power distribution networks [12], [13], on-chip simultaneous switching noise should be considered when estimating the propagation delay of a logic gate in a synchronous CMOS integrated circuit.

A lumped RLC model is applied in this paper to characterize the power supply rails in a synchronous CMOS integrated circuit. The submicrometer MOS transistors are modeled by the nth power law I-V model [14]. Analytical expressions are developed to characterize the on-chip simultaneous noise and the output voltage waveform of a CMOS logic gate. The output voltage waveform based on the analytical expressions is quite close to SPICE. For a capacitive load, the maximum error of the propagation delay model based on the analytical expressions is within 5%, as compared to close to 16% of the estimate if on-chip simultaneous switching noise is not considered. The average improvement in accuracy is about 8%. For a resistive-capacitive load, the estimated propagation delay based on the analytical expressions is within 5%, while the error of the model which does not consider on-chip simultaneous switching noise can reach 18% with an average improvement in accuracy of 10%.

Analytical expressions modeling the noise voltage of m simultaneous switching logic gates are presented in Section II. Analytical expressions characterizing the output voltage for both a capacitive and a resistive-capacitive load are derived in Section III, while the output voltage and propagation delay based on these analytical expressions are also compared to SPICE. Some concluding remarks are addressed in Section IV.

II. MODELING ON-CHIP SIMULTANEOUS SWITCHING NOISE

The power supply in high complexity CMOS integrated circuits should provide sufficient current to support the average and peak power demand within all parts of an integrated circuit. An analytical expression characterizing the on-chip simultaneous switching noise voltage at the ground rail is developed in this section for a high-to-low output transition. $L_{V_{ss}}$, $R_{V_{ss}}$, and $C_{V_{ss}}$ are the parasitic inductance, resistance, and capacitance of the ground rail, respectively. In order to derive an analytical expression characterizing the on-chip simultaneous switching noise on the ground rail, the short-circuit current is neglected based on an assumption of a fast ramp input signal [15],

$$V_{in}(t) = \frac{t}{\tau_r} V_{dd} \quad \text{for} \quad 0 \le t \le \tau_r. \tag{1}$$

Once the input voltage exceeds the threshold voltage of an NMOS transistor, the NMOS transistor turns ON and is assumed to operate in the saturation region during the input transition. The current through the NMOS transistor (I_N), the parasitic inductance (I_L), and the simultaneous switching noise voltage (V_s) are given, respectively, as

$$I_N = B_n(V_{in} - V_{TN} - V_s)^n, \tag{2}$$

$$I_L = mI_N - C_{V_{ss}} \frac{dV_s}{dt}, \tag{3}$$

$$V_s = R_{V_{ss}} I_L + L_{V_{ss}} \frac{dI_L}{dt}, \tag{4}$$

where m is the number of simultaneously switching logic gates and B_n is the transconductance of the NMOS transistor. The solution of the simultaneous switch-

ing noise voltage in this region is [16]

$$V_s(t) = c_0[1 - e^{-\tau_1(t-\tau_n)}] + c_1\xi + c_2\xi^2 + c_3\xi^3 + c_4\xi^4 + c_5\xi^5 \quad \text{for} \quad \tau_n \le t \le \tau_r, \tag{5}$$

where $\tau_1 = \frac{1}{\gamma\tau_r}$. These coefficients are

$$\begin{aligned}
c_0 &= A_0\gamma - A_1\gamma^2 + 2A_2\gamma^3 - 6A_3\gamma^4 + 24A_4\gamma^5 - 120A_5\gamma^6, \\
c_1 &= A_1\gamma - 2A_2\gamma^2 + 6A_3\gamma^3 - 24A_4\gamma^4 + 120A_5\gamma^5, \\
c_2 &= A_2\gamma - 3A_3\gamma^2 + 12A_4\gamma^3 - 60A_5\gamma^4, \\
c_3 &= A_3\gamma - 4A_4\gamma^2 + 20A_5\gamma^3, \\
c_4 &= A_4\gamma - 5A_5\gamma^2, \\
c_5 &= A_5\gamma,
\end{aligned} \tag{6}$$

where

$$\gamma = \frac{R_{V_{ss}}C_{V_{ss}} + L_{V_{ss}}f_{1,m}}{(R_{V_{ss}}f_{1,m} + 1)\tau_r}. \tag{7}$$

The A_i for $i = 0 \ldots 5$ are

$$A_i = \frac{mR_{V_{ss}}B_nV_{dd}^n\tau_r}{R_{V_{ss}}C_{V_{ss}} + L_{V_{ss}}f_{1,m}}a_i + \frac{mL_{V_{ss}}B_nV_{dd}^n}{R_{V_{ss}}C_{V_{ss}} + L_{V_{ss}}f_{1,m}}b_i, \tag{8}$$

where $f_{1,m} = mnB_n(0.5V_{dd} - V_{TN})^{(n-1)}$ and a_i (b_i) for $i = 0 \ldots 5$ is determined by a polynomial expansion of the drain-to-source current during the input transition [16]. The on-chip simultaneous switching noise voltage reaches a maximum when the input voltage completes the transition, *i.e.*, $t = \tau_r$.

The NMOS transistor is assumed to remain saturated when the input transition is completed. The on-chip simultaneous switching noise voltage in this region can be expressed as

$$V_s(t) = V_{s,1} + [V_s(\tau_r) - V_{s,1}]e^{-\tau_2(t-\tau_r)} \qquad \text{for} \quad \tau_r \le t \le \tau_{sat}, \tag{9}$$

where

$$V_{s,1} = \frac{mB_nR_{Vss}(V_{dd} - V_{TN})^n}{R_{V_{ss}}f_{2,m} + 1}, \tag{10}$$

$$\tau_2 = \frac{R_{V_{ss}}f_{2,m} + 1}{R_{V_{ss}}C_{V_{ss}} + L_{V_{ss}}f_{2,m}}, \tag{11}$$

$$f_{2,m} = mnB_n(V_{dd} - V_{TN} - V_s(\tau_r))^{(n-1)}. \tag{12}$$

After τ_{sat}, the NMOS transistor operates in the linear region. In order to derive a tractable expression, the drain-to-source current is approximated by γ_nV_{DS}, where γ_n is the effective output conductance of the NMOS transistor. The on-chip simultaneous switching noise voltage in region IV is

$$V_s(t) = K_1e^{-\frac{\beta_1 t}{2}} + K_2e^{-\frac{\beta_2 t}{2}} \qquad \text{for} \quad t \ge \tau_{sat}, \tag{13}$$

where

$$B_1 = \frac{mR_{V_{ss}}C_L + R_{V_{ss}}C_{V_{ss}} + \frac{C_L}{\gamma_n}}{R_{V_{ss}}C_{V_{ss}}\frac{C_L}{\gamma_n} + mL_{V_{ss}}C_L + L_{V_{ss}}C_{V_{ss}}},$$

$$B_2 = \frac{1}{R_{V_{ss}}C_{V_{ss}}\frac{C_L}{\gamma_n} + mL_{V_{ss}}C_L + L_{V_{ss}}C_{V_{ss}}},$$

$$\beta_1 = B_1 + \sqrt{B_1^2 - 4B_2},$$

$$\beta_2 = B_1 - \sqrt{B_1^2 - 4B_2}. \tag{14}$$

C_L is the load capacitance. K_1 and K_2 are integration constants and can be determined from $V_s(\tau_{sat})$ and $V_s'(\tau_{sat})$. However, the effective output conductance of a MOS transistor also depends upon the output voltage in the linear region, changing from γ_{nsat} to $2\gamma_{nsat}$ [14]. In order to accurately characterize the on-chip simultaneous switching noise voltage, γ_n is chosen between γ_{nsat} and $2\gamma_{nsat}$.

$B_1^2 - 4B_2$ can be expressed as

$$B_1^2 - 4B_2 = (mR_{V_{ss}}C_L + R_{V_{ss}}C_{V_{ss}} + \frac{C_L}{\gamma_n})^2 - 4(R_{V_{ss}}C_{V_{ss}}\frac{C_L}{\gamma_n} + mL_{V_{ss}}C_L + L_{V_{ss}}C_{V_{ss}}). \tag{15}$$

If $B_1^2 - 4B_2$ is less than zero,

$$L_{V_{ss}} > \frac{(mR_{V_{ss}}C_L + R_{V_{ss}}C_{V_{ss}} + \frac{C_L}{\gamma_n})^2 - 4R_{V_{ss}}C_{V_{ss}}\frac{C_L}{\gamma_n}}{4(mC_L + C_{V_{ss}})}, \tag{16}$$

resulting in a complex solution, which means the on-chip simultaneous switching noise oscillates sinusoidally until exponentially reaching a steady state voltage in the linear region. The critical value defined in (16) depends upon the resistance and capacitance of the power supply rail ($R_{V_{ss}}$ and $C_{V_{ss}}$), the load condition (C_L), the effective output conductance of the NMOS transistor (γ_n), and the number of switching gates (m). In practical integrated circuits, the load condition, the size of the MOS transistors, and the number of switching gates should be optimized in order to satisfy the condition described by (16). The on-chip simultaneous switching noise voltage can be approximated by forcing the imaginary part to zero,

$$V_s(t) = V_s\tau_{sat}e^{-\frac{\beta_3(t-\tau_{sat})}{2}}\frac{cos(\beta_4 t)}{cos(\beta_4\tau_{sat})} \qquad \text{for } t \geq \tau_{sat}, \tag{17}$$

where $\beta_3 = B_1$ and $\beta_4 = \sqrt{4B_2 - B_1^2}$. If $R_{V_{ss}}$ and $C_{V_{ss}}$ are assumed to be zero, the solution represented by (17) and the condition defined by (16) are similar to the results presented in [9], which characterizes the simultaneous switching noise voltage caused by the off-chip bonding wires. The parasitic inductance within the on-chip power distribution network is typically less than the critical value defined in (16) [7]. However, for off-chip bonding wires, the condition defined in (16) may occur, where the simultaneous switching noise oscillates sinusoidally before exponentially reaching a steady state voltage [9]. Therefore, the model of the power supply rails presented here is a unified approach which can characterize both on-chip and off-chip simultaneous switching noise.

III. OUTPUT VOLTAGE WAVEFORM AND PROPAGATION DELAY MODELS

Analytical expressions characterizing the output voltage and propagation delay of a CMOS logic gate, which include the effects of on-chip simultaneous switching noise, are presented for both a capacitive and a resistive-capacitive load in subsections A and B, respectively. The analytical results are also compared to both an analytical model which does not consider on-chip simultaneous switching noise and SPICE.

A. Capacitive load

Analytical expressions characterizing the output voltage of a CMOS logic gate driving a capacitive load based on an assumption of a fast ramp input signal are listed in Table I, where $f_{1,s} = \frac{f_{1,m}}{m}$ and $f_{2,s} = \frac{f_{2,m}}{m}$. K_1 and K_2 are determined from $V_o(\tau_{sat})$ and $V_o'(\tau_{sat})$. Note that both $\frac{f_{1,s}\tau_r}{C_L}V_{s,2}(t)$ in (18) and $\frac{f_{2,s}}{C_L}V_{s,3}(t)$ in (20) cause the output voltage to drop slowly during a high-to-low output transition. Therefore, the on-chip simultaneous switching noise affects the waveform shape of the output voltage and causes a change in the propagation delay of a CMOS logic gate driving a capacitive load.

TABLE I
ANALYTICAL EXPRESSIONS CHARACTERIZING THE OUTPUT VOLTAGE OF A CMOS INVERTER DRIVING A CAPACITIVE LOAD

Region	Analytical expressions
$[\tau_n, \tau_r]$	$V_o(t) = V_{dd} - \frac{B_n \tau_r V_{dd}^n}{(n+1)C_L}\xi^{(n+1)} + \frac{f_{1,s}\tau_r}{C_L}V_{s,2}(t)$ (18) $V_{s,2}(t) = (\xi + \frac{e^{-\tau_1 \xi}}{\tau_1})c_0 + \frac{c_1}{2}\xi^2 + \frac{c_2}{3}\xi^3 + \frac{c_3}{4}\xi^4 + \frac{c_4}{5}\xi^5 + \frac{c_5}{6}\xi$ (19)
$[\tau_r, \tau_{sat}]$	$V_o(t) = V_o(\tau_r) - \frac{B_n}{C_L}(V_{dd} - V_{TN})^n(t - \tau_r) + \frac{f_{2,s}}{C_L}V_{s,3}(t)$ (20) $V_{s,3}(t) = V_{s,1}(t - \tau_r) + \frac{V_s(\tau_r) - V_{s,1}}{\tau_2}(1 - e^{-\tau_2(t-\tau_r)})$ (21)
$t \geq \tau_{sat}$	$V_o(t) = K_1 e^{-\frac{\beta_1 t}{2}} + K_2 e^{-\frac{\beta_2 t}{2}}$ (22)

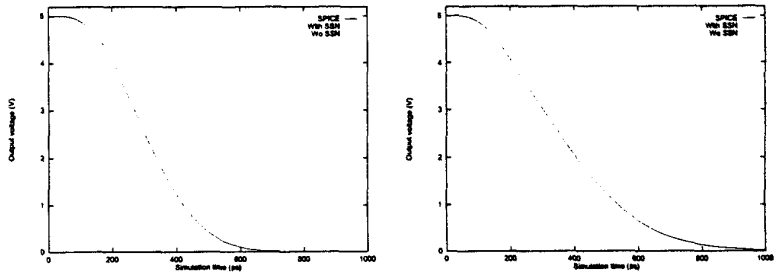
The time when the NMOS transistor leaves the saturation region τ_{sat} can be determined as

$$V_o(\tau_{sat}) - V_s(\tau_{sat}) = V_{sat}. \tag{23}$$

There is no explicit solution of τ_{sat}, but τ_{sat} can be determined by applying the Newton-Raphson technique. The technique typically requires two to four iterations.

If $B_1^2 - 4B_2$ is less than zero, a complex solution of the output voltage results. The output voltage can therefore be approximated as

$$V_s(t) = V_o(\tau_{sat})e^{-\frac{\beta_3(t-\tau_{sat})}{2}}\frac{cos(\beta_4 t)}{cos(\beta_4 \tau_{sat})} \qquad \text{for } t \geq \tau_{sat}, \tag{24}$$



(a) $C_L = 0.1\,\text{pF}$ and $\tau_r = 0.20\,\text{ns}$ (b) $C_L = 0.15\,\text{pF}$ and $\tau_r = 0.15\,\text{ns}$

Fig. 1. Comparison of the analytical output voltage with SPICE during a high-to-low output transition, $L_{V_{ss}} = 3.0\,\text{nH}$, $R_{V_{ss}} = 20.0\,\Omega$, $C_{V_{ss}} = 0.1\,\text{pF}$, $w_n = 1.8\,\mu\text{m}$, $w_p = 3.6\,\mu\text{m}$, and $m = 10$.

where β_3 and β_4 are defined in (17).

The output voltage based on these analytical expressions is compared to SPICE in Fig. 1 for a high-to-low output transition. Note that the analytical waveform which considers on-chip simultaneous switching noise voltage is quite close to SPICE. The difference in the linear region between the analytically derived waveform and SPICE is due to an assumption of a constant effective output conductance γ_{sat} of the MOS transistor in the analysis. However, the effective output conductance of the MOS transistor changes from γ_{nsat} to $2\gamma_{nsat}$ in the linear region [14], causing the analytical prediction which does not consider on-chip simultaneous switching noise to be more accurate than the analytical result based on the expressions listed in Table I for a portion of the linear region. The effects of the on-chip simultaneous switching noise on the propagation delay of a CMOS logic gate are also depicted in Fig. 1.

The propagation delay of a CMOS logic gates t_P is typically defined as the time from the 50% V_{dd} point of the input to the 50% V_{dd} point of the output. The high-to-low propagation delay of a CMOS logic gate can be determined by (20) or (22) using a Newton-Raphson iteration. Since β_1 is greater than β_2 in (13), the output voltage in region IV can be approximated as

$$V_o(t) = V_o(\tau_{sat})e^{-\frac{\beta_2}{2}(t-\tau_{sat})}. \tag{25}$$

The drain-to-source saturation voltage is typically greater than $0.5V_{dd}$; therefore, the high-to-low propagation delay can be expressed as

$$t_{P_{HL}} = \frac{2}{\beta_2} ln \frac{2V_o(\tau_{sat})}{V_{dd}} + \tau_{sat} - \frac{\tau_r}{2}. \tag{26}$$

Similarly, the low-to-high propagation delay of a CMOS logic gate can also be determined based on the time required to charge the load capacitor, which considers the simultaneous switching noise at the V_{dd} rail.

A comparison of the analytical propagation delay expressions with SPICE is listed in Table II for a high-to-low output transition. It is demonstrated in Table II that the delay uncertainty caused by on-chip simultaneous switching noise increases with increasing input slew rate, parasitic inductance, and resistance of the power supply

TABLE II
COMPARISON OF HIGH-TO-LOW PROPAGATION DELAY WITH SPICE OF A CMOS INVERTER DRIVING A CAPACITIVE LOAD INCLUDING THE EFFECTS OF ON-CHIP SIMULTANEOUS SWITCHING NOISE

Input Rise Time	Impedance of Ground Rail			Comparison of Propagation Delay				
				Simulation	Analytic Estimation			
τ_r	L	R	C	SPICE	Without SSN		With SSN	
(ps)	(nH)	(Ω)	(pF)	(ps)	(ps)	δ (%)	(ps)	δ (%)
200	1.0	5.0	0.1	182	172	5.5	183	1.1
		10.0	0.1	186	172	7.5	183	1.6
		15.0	0.1	190	172	9.5	186	2.1
		20.0	0.1	194	172	11.3	189	2.6
	2.0	5.0	0.1	186	172	7.5	183	1.6
		10.0	0.1	190	172	9.5	186	2.1
		15.0	0.1	194	172	11.3	189	2.6
		20.0	0.1	198	172	13.1	192	3.0
	3.0	5.0	0.1	191	172	10.0	187	2.1
		10.0	0.1	195	172	11.8	190	2.6
		15.0	0.1	199	172	13.6	192	3.5
		20.0	0.1	203	172	15.3	195	3.9
150	1.0	5.0	0.1	175	166	5.1	174	0.6
		10.0	0.1	180	166	7.7	177	1.7
		15.0	0.1	184	166	9.8	180	2.2
		20.0	0.1	189	166	12.2	183	3.2
	2.0	5.0	0.1	179	166	7.3	177	1.1
		10.0	0.1	184	166	9.7	180	2.2
		15.0	0.1	188	166	11.7	183	2.7
		20.0	0.1	193	166	14.0	185	4.1
	3.0	5.0	0.1	185	166	10.3	180	2.7
		10.0	0.1	188	166	11.7	182	3.2
		15.0	0.1	193	166	14.0	185	4.1
		20.0	0.1	197	166	15.7	188	4.6
Maximum error (%)					15.7		4.6	
Average error (%)					10.63		2.25	

rails. The maximum error of the propagation delay based on the analytical expressions is within 5%, as compared to nearly 16% of SPICE when not considering on-chip simultaneous switching noise. The average improvement in accuracy is about 8%.

B. **Resistive-capacitive load**

In this discussion, the interconnect is modeled as a resistive-capacitive impedance. R_L is the load resistance driven by a CMOS logic gate. Analytical expressions characterizing the output voltage of a CMOS logic gate driving a resistive-capacitive load impedance are listed in Table III. Similar to the capacitive load condition, τ_{sat} can be determined by applying the Newton-Raphson technique. K_1 and K_2 are also determined from $V_o(\tau_{sat})$ and $V_o'(\tau_{sat})$. Both $\frac{f_{1,\phi}\tau_r}{C_L}V_{s,2}(t)$ in (27) and $\frac{f_{2,\phi}}{C_L}V_{s,3}(t)$ in (28) cause the output voltage to drop slowly during a high-to-low output transition for a resistive-capacitive load impedance. Therefore, the on-chip simultaneous switching noise affects the waveform shape of the output voltage, increasing the propagation delay of a CMOS logic gate driving a resistive-capacitive load impedance.

TABLE III
ANALYTICAL EXPRESSIONS CHARACTERIZING THE OUTPUT VOLTAGE OF A CMOS INVERTER DRIVING A RESISTIVE-CAPACITIVE LOAD

Region	Analytical expressions
$[\tau_n, \tau_r]$	$V_o(t) = V_{dd} - \frac{B_n \tau_r V_{dd}^n}{(n+1)C_L}\xi^{(n+1)} - R_L B_n V_{dd}^n \xi^n + \frac{f_{1,s}\tau_r}{C_L} V_{s,2}(t)$ (27)
$[\tau_r, \tau_{sat}]$	$V_o(t) = V_o(\tau_r) - \frac{B_n}{C_L}(V_{dd} - V_{TN})^n (t - \tau_r) + \frac{f_{2,s}}{C_L} V_{s,3}(t)$ (28)
$t \geq \tau_{sat}$	$V_o(t) = K_1 e^{-\frac{\beta_5 t}{2}} + K_2 e^{-\frac{\beta_6 t}{2}}$ (29) $B_3 = \frac{mR_{V_{ss}}C_L + R_{V_{ss}}C_{V_{ss}} + \frac{C_L(1+R_L\gamma_n)}{\gamma_n}}{R_{V_{ss}}C_{V_{ss}}\frac{C_L(1+R_L\gamma_n)}{\gamma_n} + mL_{V_{ss}}C_L + L_{Vss}C_{V_{ss}}}$ $B_4 = \frac{1}{R_{V_{ss}}C_{V_{ss}}\frac{C_L(1+R_L\gamma_n)}{\gamma_n} + mL_{V_{ss}}C_L + L_{Vss}C_{V_{ss}}}$ $\beta_5 = B_3 + \sqrt{B_3^2 - 4B_4}$ $\beta_6 = B_3 - \sqrt{B_3^2 - 4B_4}$ (30)

For a resistive-capacitive load, $B_3^2 - 4B_4$ can be expressed as

$$B_3^2 - 4B_4 = (mR_{V_{ss}}C_L + R_{V_{ss}}C_{V_{ss}} + \frac{C_L(1+R_L\gamma_n)}{\gamma_n})^2 - 4(R_{V_{ss}}C_{V_{ss}}\frac{C_L(1+R_L\gamma_n)}{\gamma_n} + mL_{V_{ss}}C_L + L_{Vss}C_{V_{ss}}). \quad (31)$$

If $B_3^2 - 4B_4$ is less than zero, the critical value of the parasitic inductance defined in (16) becomes

$$L_{V_{ss}} > \frac{(mR_{V_{ss}}C_L + R_{V_{ss}}C_{V_{ss}} + \frac{C_L(1+R_L\gamma_n)}{\gamma_n})^2 - 4R_{V_{ss}}C_{V_{ss}}\frac{C_L(1+R_L\gamma_n)}{\gamma_n}}{4(mC_L + C_{V_{ss}})}. \quad (32)$$

The output voltage in region IV can be approximated as

$$V_s(t) = V_o(\tau_{sat})e^{-\frac{\beta_3(t-\tau_{sat})}{2}}\frac{cos(\beta_4 t)}{cos(\beta_4 \tau_{sat})} \qquad \text{for } t \geq \tau_{sat}, \quad (33)$$

where $\beta_3 = B_3$ and $\beta_4 = \sqrt{4B_4 - B_3^2}$.

Based on the same assumption as for a capacitive load, the output voltage in region IV can be approximated as

$$V_o(t) = V_o(\tau_{sat})e^{-\frac{\beta_6}{2}(t-\tau_{sat})}. \quad (34)$$

TABLE IV
COMPARISON OF HIGH-TO-LOW PROPAGATION DELAY WITH SPICE OF A CMOS INVERTER DRIVING A RESISTIVE-CAPACITIVE LOAD INCLUDING THE EFFECTS OF ON-CHIP SIMULTANEOUS SWITCHING NOISE

Input Rise Time	Impedance of Ground Rail			Comparison of Propagation Delay				
				Simulation	Analytic Estimation			
τ_r	L	R	C	SPICE	Without SSN		With SSN	
(ps)	(nH)	(Ω)	(pF)	(ps)	(ps)	δ (%)	(ps)	δ (%)
200	1.0	5.0	0.1	163	152	6.7	162	0.6
		10.0	0.1	167	152	9.0	165	1.2
		15.0	0.1	172	152	11.6	168	2.3
		20.0	0.1	176	152	13.6	171	2.8
	2.0	5.0	0.1	167	152	9.0	165	1.2
		10.0	0.1	171	152	11.1	168	1.8
		15.0	0.1	176	152	13.6	171	2.8
		20.0	0.1	180	152	15.6	174	3.3
	3.0	5.0	0.1	172	152	11.6	169	1.7
		10.0	0.1	176	152	13.6	172	2.3
		15.0	0.1	180	152	15.6	175	2.8
		20.0	0.1	184	152	17.4	177	3.8
150	1.0	5.0	0.1	157	146	7.0	156	0.6
		10.0	0.1	162	146	9.9	159	1.9
		15.0	0.1	166	146	12.0	162	2.4
		20.0	0.1	170	146	14.1	165	2.9
	2.0	5.0	0.1	161	146	9.3	159	1.2
		10.0	0.1	165	146	11.5	162	1.8
		15.0	0.1	170	146	14.1	165	2.9
		20.0	0.1	174	146	16.1	168	3.4
	3.0	5.0	0.1	166	146	12.0	162	2.4
		10.0	0.1	170	146	14.1	165	2.9
		15.0	0.1	174	146	16.1	168	3.4
		20.0	0.1	170	146	18.0	170	4.5
Maximum error (%)					18.0		4.5	
Average error (%)					12.61		2.38	

If the drain-to-source saturation voltage is greater than $0.5V_{dd}$, the high-to-low propagation delay can be expressed as

$$t_{P_{HL}} = \frac{2}{\beta_6} ln \frac{2V_o(\tau_{sat})}{V_{dd}} + \tau_{sat} - \frac{\tau_r}{2}. \tag{35}$$

A comparison of the propagation delay based on these analytical expressions with SPICE is listed in Tables IV for a high-to-low output transitions. The estimated propagation delay based on these analytical expressions is within 5%, while the error of the estimate which does not consider on-chip simultaneous switching noise can reach 18%. The average improvement in accuracy is about 10%.

If on-chip simultaneous switching noise cannot be neglected in VDSM synchronous CMOS integrated circuits, these analytical equations, (26) and (35), provide system level timing characteristics of a CMOS logic gate driving both a capacitive and a resistive-capacitive load. This timing information can be used to develop guidelines and methodologies for designing tapered buffers and inserting repeaters in order to improve interconnect-based circuit performance.

IV. CONCLUSIONS

It is necessary to consider on-chip simultaneous switching noise when determining the propagation delay of a CMOS logic gate in a high speed synchronous

CMOS integrated circuit. Analytical expressions characterizing on-chip simultaneous switching noise are presented in this paper. The effects of on-chip simultaneous switching noise on the waveform of the output voltage and the propagation delay of a CMOS logic gate are also discussed. The estimated propagation delay based on these analytical expressions is within 5% as compared to SPICE; the average improvement can reach 10% as compared to delay estimates which do not consider on-chip simultaneous switching noise. The analytical expressions presented in this paper provide an accurate timing model for repeater insertion, tapered buffer design, and related high performance design techniques for those high speed synchronous CMOS integrated circuits where on-chip simultaneous switching noise cannot be neglected.

V. ACKNOWLEDGEMENTS

This research was supported in part by National Science Foundation under Grant No. MIP-9610108, the Semiconductor Research Corporation under Contract No. 99-TJ-687, a grant from the New York State Science and Technology Foundation to the Center for Advanced Technology–Electronic Imaging Systems, and by grants from Xerox Corporation, IBM Corporation, Intel Corporation, Lucent Technologies Corporation, and Eastman Kodak Company.

REFERENCES

[1] D. W. Dobberpuhl *et al.*, "A 200-MHz 64-bit Dual-Issue CMOS Microprocessor," *IEEE Journal of Soiid-State Circuits*, Vol. SC–27, No. 11, pp. 1555–1565, November 1992.

[2] W. J. Bowhill *et al.*, "Circuit Implementation of a 300-MHz 64-bit Second-Generation CMOS Alpha CPU," *Digital Technical Journal*, Vol. 7, No. 1, pp. 100–118, 1995.

[3] P. E. Gronowski *et al.*, "High-Performance Microprocessor Design," *IEEE Journal of Solid-State Circuits*, Vol. SC-33, No. 5, pp. 676–686, May 1998.

[4] Semiconductor Industry Association, "The National Technology Roadmap for Semiconductors," 1997.

[5] G. Katopis, "ΔI Noise Specification for a High Performance Computer Machine," *Proceedings of the IEEE*, Vol. 73, No. 9, pp. 1405–1415, September 1985.

[6] Semiconductor Industry Association, "The National Technology Roadmap for Semiconductor," 1994.

[7] L.-R. Zheng and H. Tenhunen, "Effective Power and Ground Distribution Scheme for Deep Submicron High Speed VLSI Circuits," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 537–540, May 1999.

[8] Y.-M. Jiang and K.-T. Cheng, "Analysis of Performance Impact Caused by Power Supply Noise in Deep Submicron Devices," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 760–765, June 1999.

[9] S. R. Vemuru, "Effects of Simultaneous Switching Noise on the Tapered Buffer Design," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 5, No. 3, pp. 290–300, September 1997.

[10] J. J. Becerra and E. G. Friedman, *Analog Design Issues in Digital VLSI Circuits and Systems*. Norwell, Massachusetts: Kluwer Academic Publishers, 1997.

[11] E. G. Friedman, *High Performance Clock Distribution Networks*. Kluwer Academic Publishers, 1997.

[12] SRC, "SRC Physical Design Top Ten Problem," Technical Report, SRC Physical Design Task Force, Semiconductor Research Corporation, November 1998.

[13] SRC, "Design Sciences TAB: Physical Design Task Force Report," Technical Report, Semiconductor Research Corporation Physical Design Task Force, Semiconductor Research Corporation, April 1997.

[14] T. Sakurai and A. R. Newton, "A Simple MOSFET Model for Circuit Analysis," *IEEE Transactions on Electron Devices*, Vol. ED-38, No. 4, pp. 887–894, April 1991.

[15] N. Hedenstierna and K. O. Jeppson, "CMOS Circuits Speed and Buffer Optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. CAD-6, No. 2, pp. 270–280, March 1987.

[16] K. T. Tang and E. G. Friedman, "On-Chip Simultaneous Switching Noise in VDSM CMOS Circuits," *Proceedings of the International Conference on Modeling and Simulation of Microsystems*, pp. 313–316, March 2000.