Repeater Design to Reduce Delay and Power in Resistive Interconnect

Victor Adler, *Student Member, IEEE,* and Eby G. Friedman, *Senior Member, IEEE*

Abstract—In large chips, the propagation delay of the data and clock signals can limit performance due to long resistive interconnect. The insertion of repeaters alleviates the quadratic increase in propagation delay with interconnect length while decreasing power dissipation by reducing short-circuit current. In order to develop a repeater design methodology, a timing model characterizing a complementary metal–oxide–semiconductor (CMOS) inverter driving a resistance–capacitance (RC) load is presented. The model is based on the Sakurai short-channel α-power law model of transistor operation.

The inverter model is applied to the problem of repeaters to produce design expressions for determining the optimum number of uniformly sized repeaters to be inserted along a resistive interconnect line for reduced delay. For a wide variety of typical RC loads, this analytical repeater model exhibits a maximum error of 16% as compared to a dynamic circuit simulator (SPICE). The advantage of uniformly sized repeaters versus tapered-buffer repeaters is also investigated using the repeater model presented in this paper. It is shown that uniform repeaters remain advantageous over tapered buffers and tapered-buffer repeaters even with relatively small resistive RC loads.

An expression for the short-circuit power dissipation of a repeater driving an RC load is presented. A comparison of the short-circuit power dissipation to the dynamic power dissipation in repeater chains and related power/delay tradeoffs are made.

Index Terms—Buffer insertion, delay optimization, RC interconnect, repeaters.

I. INTRODUCTION

AS THE SIZE of complementary metal–oxide–semiconductor (CMOS) integrated circuits continues to increase, interconnections have become increasingly significant. With a linear increase in length, interconnect delay increases quadratically due to a linear increase in both interconnect resistance and capacitance [1], [2]. Also, large interconnect loads not only affect performance, but degrade the waveform shape, causing excessive short-circuit power to be dissipated in the stage loading a CMOS logic gate.

Several methods have been introduced to reduce interconnect delay so that these impedances do not dominate the delay of a critical path. Bakoglu presents a method in which the delay of a repeater is characterized by the input capacitance

and output resistance based on the geometric size of the repeaters [1], [3]. Bakoglu equalizes the delay of the repeaters and the interconnect delay to optimize the number and size of the repeaters for a specific resistance–capacitance (RC) interconnect impedance.

In [4] and [5], Wu and Shiau describe a repeater implementation to reduce interconnect delay. Their method uses a linearized form of the Shichman–Hodges equations [6] at a specific operating point to determine the proper repeater insertion locations. Nekili and Savaria consider optimal methods for driving resistive interconnect in [7]. They introduce the concept of parallel regeneration in [8] in which precharge circuitry is added to the repeaters to decrease the evaluation time. Although this technique requires fewer repeaters, extra area is necessary, and a precharge signal is required to operate correctly.

Dhar and Franklin present a mathematical treatment for optimal repeater insertion in which elegant solutions are described to optimize repeaters with and without area constraints [9]. However, the repeater is modeled as a simple resistor and capacitor and no closed form solution is developed. Other repeater insertion methods are described in [10]–[12].

In this paper, CMOS inverting repeaters are presented as a simple yet effective way of reducing the total propagation delay and transition time characteristics of a system with highly resistive interconnect. A methodology is presented for determining the number and size of the repeaters to attain the minimum propagation delay based on an analytical expression derived from the α-power law model for short-channel devices [13], [14]. Using the α-power law model permits the development of a repeater design methodology that considers the short-channel transistor effect of velocity saturation which is not considered in any of the aforementioned repeater methodologies [1], [3]–[5], [7]–[12]. Furthermore, the proposed model is based on current versus voltage (I–V) equations rather than modeling a CMOS inverter as a discrete resistor and capacitor. Unlike previous work, the method presented in this paper does not separate the device model from the interconnect model.

Alternative methods to uniform repeaters driving RC loads are also considered in this paper. A tapered-buffer repeater structure provides high drive capability with low input capacitance; however, the additional buffer stages may add significant delay. It is shown here that for even relatively small resistances, uniform repeaters are more effective in driving RC loads than tapered buffers or tapered-buffer repeaters.

In addition to delay, power is considered. With the introduction of portable and massively parallel applications, power has
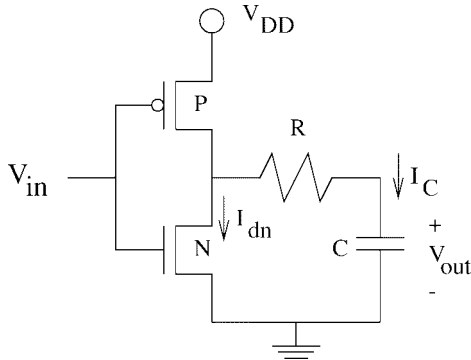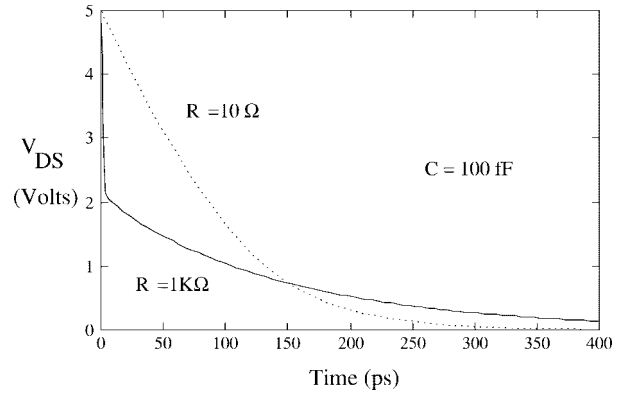
Fig. 1. A CMOS inverter driving an *RC* load.

Fig. 2. V_{ds} across the N-channel transistor of a CMOS inverter driving various load resistances R and a constant load capacitance ($C = 100$ fF).

become an increasingly important factor in the circuit design process [15]. For example, clock distribution networks can account for 40% of the total power dissipated on-chip [16]. A high performance clock distribution network can contain many thousands of repeaters due to the distributed *RC* nature of a clock tree. Thus, power consumption must be both accurately estimated and minimized when developing design techniques that improve the speed of the signal propagation through long resistive interconnects. The issue of minimizing power dissipation in repeater systems is therefore analyzed in this paper. Two components of the transient power dissipation are considered herein. A comparison of the power contribution of both the dynamic power and the short-circuit power in a CMOS inverter driving an *RC* line is made. An empirical analysis is presented for determining the optimal number of repeaters to attain the minimum power when considering both short-circuit and dynamic power dissipation.

The paper is organized as follows: in Section II, a timing model of a CMOS inverter driving a lumped *RC* load that forms the basis for the following repeater design methodology is presented. Equations characterizing the signal delay through a repeater chain are resented in Section III. A comparison of these analytic design expressions versus a dynamic circuit simulator (SPICE) are presented in Section IV. In Section V, the use of tapered-buffer repeaters versus uniformly sized repeaters is discussed. Power dissipation in repeater chains is examined in Section VI. Finally, some conclusions are presented in Section VII.

II. EXPRESSIONS FOR AN INVERTER DRIVING AN *RC* LOAD

The foundation for the repeater model is developed in this section. An analytical model describing the output voltage of a CMOS inverter driving an *RC* load (see Fig. 1) given a step input is presented. The model presented in this section utilizes a lumped *RC* load and a short-channel transistor model which is more accurate than the traditional Shichman–Hodges *I–V* equations [6].

The α-power law model accurately describes the effects of short-channel transistor behavior, such as velocity saturation, while providing a form of the *I–V* equation that is both accurate and tractable. The linear region form of the α-power model is used to characterize the *I–V* behavior of the ON transistor since a large portion of the circuit operation occurs within this

region under the assumption of a step or fast ramp input signal. When the input to the inverter is a unit step or fast ramp, V_{out} is initially larger than $V_{gs} - V_T$ for a shorter period of time than if the input to the inverter is a slow ramp. Therefore, the circuit operates in the linear region for a greater portion of the total transition time for a large *RC* load, particularly for large load resistances. When the load resistance is large, a significant *IR* voltage drop occurs across the load resistor once the capacitor begins to discharge, making V_{ds} nearly immediately less than $V_{gs} - V_T$, as shown in Fig. 2. The N-channel device operates in the linear region once the step input goes high when driving a large *RC* load. Note, however, if the input waveform increases more slowly or the load impedance is small, the inverter operates in the saturation region for a longer time before switching into the linear region [17].

Only the falling output (rising input) waveform is considered in this analysis. The following analysis, however, is equally applicable to a rising output (falling input) waveform. The lumped load is modeled as a resistor in series with a capacitor. The current through the output load capacitance is the same magnitude and opposite sign as the N-channel drain current (the P-channel current is ignored under the assumption of step or fast ramp input). The capacitive current is

$$i_C = C\frac{dV_{\text{out}}}{dt} = -i_d \tag{1}$$

where C is the output capacitance, V_{out} is the voltage across the capacitance C, i_C is the current discharged from the capacitor, and i_d is the drain current through the N-channel device.

The N-channel linear drain current is given by [13]

$$C\frac{dV_{\text{out}}}{dt} = i_d = \frac{I_{do}}{V_{do}}\left(\frac{V_{gs} - V_T}{V_{DD} - V_T}\right)^{\alpha}V_{ds}$$
$$\text{for } V_{gs} \geq V_T, V_{gs} - V_T \geq V_{ds}. \tag{2}$$

In the α-power law model, I_{do} represents the drive current of the MOS device and is proportional to W/L, V_{do} represents the drain-to-source voltage at which velocity saturation occurs with $V_{GS} = V_{DD}$ and is a process dependent constant, and α models the process dependent degree to which velocity saturation affects the drain-to-source current. α is within the range $1 \leq \alpha \leq 2$ where $\alpha = 1$ corresponds to a device
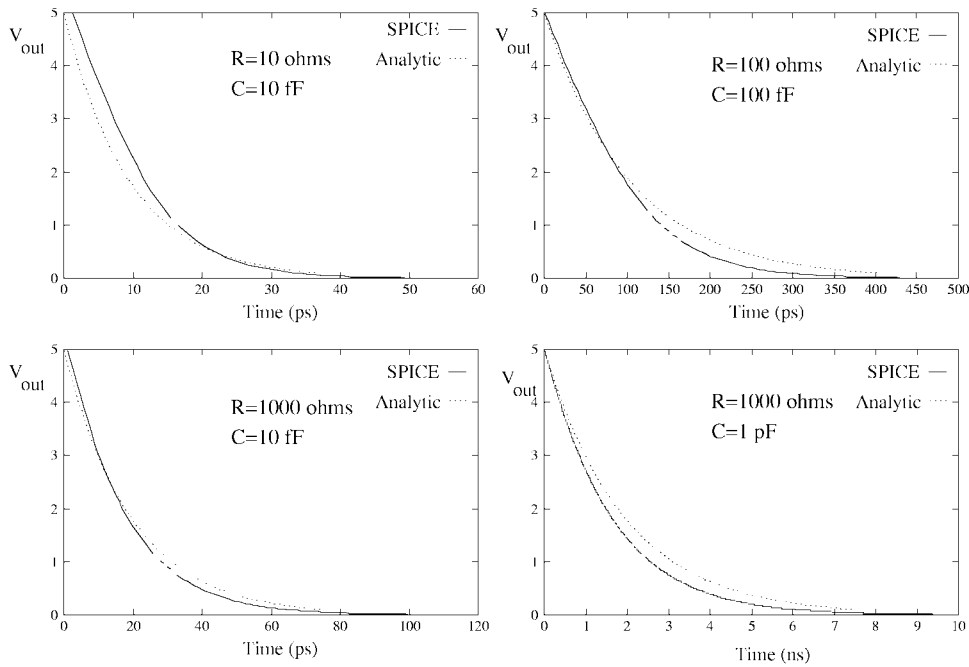
Fig. 3. Output response of a CMOS inverter driving a distributed RC load.

operating strongly under velocity saturation, while $\alpha = 2$ represents a device with negligible velocity saturation. V_{DD} is the supply voltage, and V_T is the MOS threshold voltage [where V_{TN} (V_{TP}) is the N-channel (P-channel) threshold voltage].

Assuming a unit step input is applied to the circuit shown in Fig. 1, V_{out} can be derived from (2). The linear equation, rewritten in Laplace form, is

$$SCV_{\text{out}} + S\mho_{do}RCV_{\text{out}} + \mho_{do}V_{\text{out}}$$
$$= CV_{\text{out}}(0) + \mho_{do}RCV_{\text{out}}(0) \qquad (3)$$

where $\mho_{do} = I_{do}/V_{do}$ is the saturation conductance.

Equation (3) yields

$$V_{\text{out}}(t) = V_{\text{out}}(0)e^{(-\mho_{do}/\mho_{do}RC+C)t}. \qquad (4)$$

The output voltage of a short-channel inverter driving an RC load is described by (4). This result is compared to SPICE [18] for various RC loads in Fig. 3 and exhibits an accuracy within 15%.

The information describing the waveform shape permits a more accurate delay estimation as compared to estimating the path delay based on the classical Elmore delay model [19]. Since the Elmore delay adds the products of a resistor (composed of the sum of the linearized model of an inverter and the interconnect resistance) and all of its downstream capacitors, the Elmore delay does not account for the interaction of an inverter with the RC interconnect nor does the Elmore delay consider the shape of the output signal waveform. Thus, by integrating a more accurate timing model of a CMOS inverter into methodology for inserting repeaters into an RC line, a more efficient circuit implementation can be achieved.

The expression for V_{out} can be rearranged to determine the time required for a CMOS inverter to reach an output voltage V_{out} given a step input signal:

$$t_{\text{out}} = \frac{\mho_{do}RC + C}{\mho_{do}} \ln\left(\frac{V_{DD}}{V_{\text{out}}}\right). \qquad (5)$$

Equation (5) can be used to express the 50% and 90% output delay with respect to a step input signal. These time delays are, respectively,

$$t_{50} = 0.693 \frac{(1 + \mho_{do}R_{\text{int}})C_{\text{int}}}{\mho do} \qquad (6)$$

and

$$t_{90} = 2.3 \frac{(1 + \mho_{do}R_{\text{int}})C_{\text{int}}}{\mho do}. \qquad (7)$$

These expressions are used in the following section to model the total delay required by a repeater chain to drive an RC load.

III. DELAY OF A REPEATER CHAIN DRIVING AN RC LOAD

Equations (5)–(7) presented in the previous section provide the basis for modeling the total delay of a repeater chain driving an RC load. Two other expressions are also presented in this section to complete the repeater delay model. The resulting delay model for an n-stage repeater is compared to SPICE and presented in this section.

Analytical expressions describing the behavior of a CMOS inverter driving a lumped RC load (as shown in Fig. 1) based on Sakurai's α-power law model are presented in the previous section. Equation (5) can be expanded to include the parasitic capacitances of the following inverting repeater, as

$$t_{\text{out}} = \frac{(1 + \mho_{do}R)(C_{\text{rep}} + C_{\text{int}})}{\mho_{do}} \ln\left(\frac{V_{DD}}{V_{\text{out}}}\right) \qquad (8)$$

where C_{rep} and C_{int} are the capacitances of the following inverter and the interstage load capacitance (see Fig. 4), respectively.
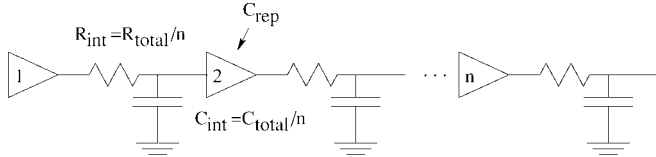
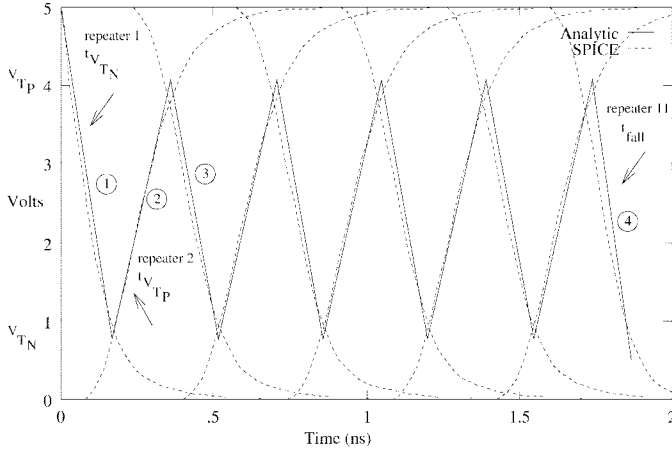Fig. 4. n equal sized CMOS inverting repeaters driving an RC load.

Fig. 5. The analytic and SPICE derived output waveforms of an 11-stage repeater chain driving an evenly distributed RC load of 1 kΩ and 1 pF.

The delay required to propagate a signal through a highly resistive interconnect can be reduced if the interconnect is broken up and distributed among a number of repeaters such as shown in Fig. 4. However, the delay of this signal path will increase if a nonoptimal number of repeaters is chosen. In order to choose the optimal number of repeaters for a given RC load, the delay from the input of the first repeater to the output of the last repeater must first be determined.

The analytical expression for the total time t_{total} from the input to the output of an n-stage repeater system is the sum of several expressions:

$$t_{\text{total}} = t_{\text{first stage}} + (n-2)t_{\text{int. stage}} + t_{\text{final stage}}. \quad (9)$$

The first term $t_{\text{first stage}}$ is the time required for the output of the first repeater to reach the turn-on voltage of the second repeater assuming the output voltage is initially at V_{DD}. The term $t_{\text{int. stage}}$ describes the time required for each repeater between the first and last stage to transition from $V_{DD} + V_{TP}$ to V_{TN} or vice versa. The time required for the output of the final repeater to reach 10%, 90%, or 50% of V_{DD} from a threshold voltage is described by the third component of (9), $t_{\text{final stage}}$ [20]. These three components of (9) are described in more detail below with reference to Fig. 5.

The first component of t_{total}, $t_{\text{first stage}}$, is the time required for the output signal of the first repeater to drop from V_{DD} to V_{TN}, the threshold voltage of the N-channel device (labeled 1 in Fig. 5) assuming a step input signal. V_{TN} is chosen as the end point because it is assumed during fast switching that the pull-up device of the following repeater turns on hard near the voltage at which the pull-down device turns off. In addition, it is assumed that the rising (falling) output of an inverting repeater reaches V_{TN} $(V_{DD} + V_{TP})$ by the time the falling input reaches V_{TN} $(V_{DD} + V_{TP})$. Thus, the signal waveforms

of the intermediate stages consistently operate between V_{TN} and $V_{DD} + V_{TP}$. The time for this switching to occur is

$$t_{V_{TN}} = \frac{(1 + \mho_{do_N} R_{\text{int}})(C_{\text{int}} + C_{\text{rep}})}{\mho do_N} \ln\left(\frac{V_{DD}}{V_{TN}}\right). \quad (10)$$

This equation also describes the time for the signal to transition from ground to $V_{DD} + V_{TP}$ when each N-channel transistor is replaced by a P-channel transistor. All of the following equations can be similarly expressed for a P-channel device. Note that V_{TP} is the P-channel threshold voltage and is negative for an enhancement mode device.

The delay of each successive stage, $(n-2)t_{\text{int. stage}}$, excluding the final stage, is modeled as the time required for the signal to transition from $V_{DD} + V_{TP}$ to V_{TN}. Equation (10) describes the time for the output signal to change from V_{DD} to V_{TN}. Therefore, the time for the signal to transition from V_{DD} to $V_{DD} + V_{TP}$ must be subtracted from (10). Equation (11) describes the time for the output signal to change from V_{DD} to $V_{DD} + V_{TP}$,

$$t_{t_P} = \frac{(1 + \mho_{do_N} R_{\text{int}})(C_{\text{int}} + C_{\text{rep}})}{\mho do_N} \ln\left(\frac{V_{DD}}{V_{DD} + V_{TP}}\right). \quad (11)$$

Therefore, an intermediate stage delay $t_{\text{int. stage}}$ is described by $(t_{V_{TP}} - t_{t_N})$ for a rising repeater output and $(t_{V_{TN}} - t_{t_P})$ for a falling repeater output (labeled 2 and 3, respectively, in Fig. 5). The two preceding expressions are alternately added to the total delay for each corresponding repeater stage up to the input of the final stage of the chain. The expression $(t_{V_{TP}} - t_{t_N})$ reduces to

$$t_N = \frac{(1 + \mho_{do_N} R_{\text{int}})(C_{\text{int}} + C_{\text{rep}})}{\mho do_N} \ln\left(\frac{V_{DD} + V_{TP}}{V_{TN}}\right). \quad (12)$$

t_P has a similar form of this expression.

The time t_{total} describes the output of the complete repeater system in terms of either: (1) the delay to reach 10% or 90% of V_{DD} from the input which is defined as the 90% output delay time t_{90} or (2) the delay at 50% V_{DD} which is defined as the 50% delay t_{50}. In order to determine the total delay to the 90% point, $t_{\text{final stage}}$ (labeled 4 in Fig. 5) is t_{90} [from (7)] minus t_{t_N} since (7) is from V_{DD} to 10% and the signal transition time to $V_{DD} + V_{TP}$ must be included. Similarly, to determine the total delay to the 50% point, $t_{\text{final stage}}$ is t_{50} [from (6)] minus t_{t_N}.

Having defined the delay of the components of the repeater system (labeled 1–4 in Fig. 5), the total time from the step input at the first repeater to the output of an even number of repeaters (for a 90% output change) is

$$t_{\text{total(even)}} = t_{V_{TN}} + \frac{(n-2)}{2}(t_N + t_P) + (t_{90} - t_{t_N}) \quad (13)$$

and for an odd number of repeaters, the time is

$$t_{\text{total(odd)}} = \frac{(n-1)}{2}(t_N + t_P) + t_{90}. \quad (14)$$

A plot of t_{total} versus the size and number of repeater stages n for an example CMOS technology and RC load is shown in
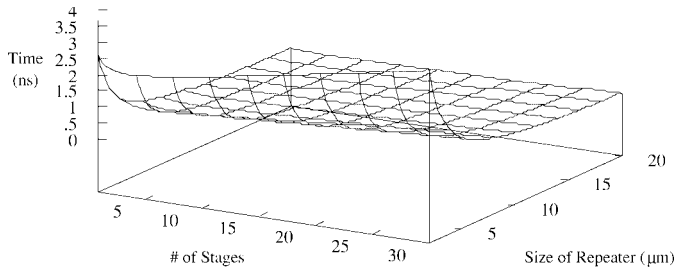
Fig. 6. The 90% output delay time for an interconnect line as a function of the number of repeaters and repeater width. ($R = 1$ kΩ, $C = 1$ pF, 0.8 μm CMOS technology).
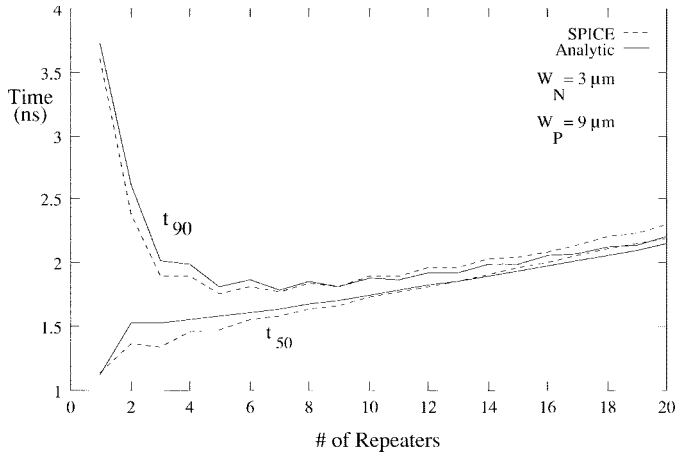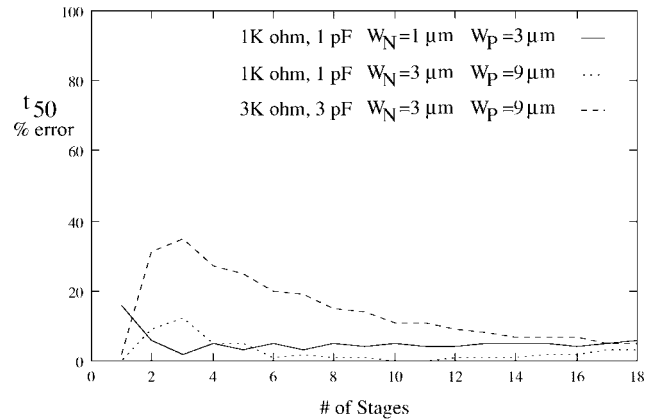

Fig. 7. The analytical and simulated 50% and 90% delay times for a 1 kΩ and 1 pF load evenly distributed across a number of uniformly sized repeaters.

(a)

(b)

Fig. 8. The percent error of the analytical value of the (a) 50% and (b) 90% output delays versus SPICE for various loads and repeater sizes.

Fig. 6. The optimal implementation of the number and size of the repeaters for this specific RC load is the minimum point on this graph. A similar graph can be determined for any RC load. Thus, (13) and (14) describe the total delay through an n-stage repeater system. These expressions are compared to SPICE in the following section.
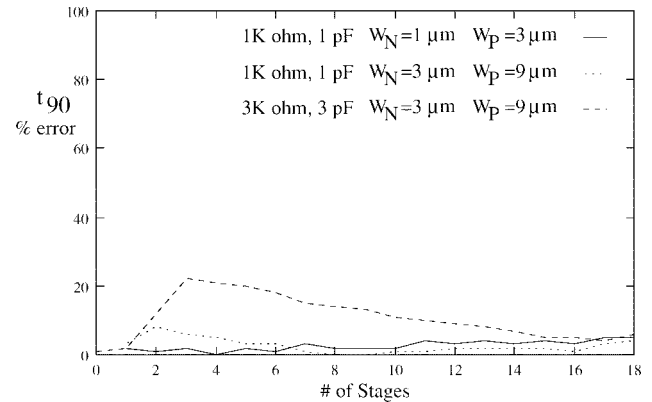
IV. ANALYTICAL DELAY MODEL VERSUS SPICE

The accuracy of the delay model for a repeater chain presented in the previous section is compared to SPICE in this section. Two different RC loads have been chosen to exemplify the effects of the interconnect resistance and capacitance on the repeater design methodology (the RC loads are 1 kΩ and 1 pF and 3 kΩ and 3 pF). These simulations are based on a 0.8 μm CMOS technology. The plots shown in Fig. 7 depict the 90% output delay t_{90} and the 50% output delay t_{50} of an RC load of 1 kΩ and 1 pF distributed evenly among one to 20 repeaters. The size of each repeater is uniform ($W_N = 3$ μm and $W_P = 9$ μm), although this analysis does not restrict the geometric widths to be uniform. The rise and fall time of each individual repeater is ratioed to maintain nearly equal transition times.

The 50% output delay of a chain of repeaters driving an RC load as a function of the number of repeater stages is shown in Fig. 7 for both the analytic expression and SPICE. The maximum error of the 50% and 90% output delays is 12% and 8%, respectively. Note that the greatest error occurs when

the repeater chain is two or three stages. The repeater model is most accurate when the loaded inverter operates predominately in the linear region. With only two or three repeaters, the inverters operate outside of the linear region for a longer period of time than with more than three repeaters. As shown in Fig. 7, there is close agreement between the analytical and simulated results for a repeater chain with more than four repeaters.

The error of the analytical delay as compared with the delay derived from SPICE for a given RC load, repeater size, and number of repeaters is shown in Table I and presented in graphical form in Fig. 8. In Table I, the number of stages into which the RC load is partitioned is shown in the first column. The propagation delay of the analytic expression and SPICE is shown in the second and third columns, respectively. The error of the analytic expression for the 50% output delay compared to SPICE is presented in the fourth column. The same information but for the 90% output delay time is listed in the fifth through seventh columns. These results are repeated for different loads and repeater sizes as denoted in the superior columns.

The deviation of the analytical result from SPICE for both t_{50} and t_{90} is shown as a function of the number of stages in Fig. 8. As shown in Fig. 8, for large RC loads (e.g., 3 kΩ and 3 pF), the model becomes less accurate since the repeaters operate for relatively less time within the linear region. At

TABLE I

PERCENT ERROR BETWEEN ANALYTICAL TOTAL DELAY MODEL (BOTH 50% AND 90% OUTPUT DELAY) VERSUS SPICE
FOR A GIVEN RC LOAD, REPEATER SIZE, AND NUMBER OF REPEATER STAGES (0.8 μm CMOS TECHNOLOGY)

| # of Stages | $R = 1$ KΩ, $C = 1$ pF $W_N = 1\mu$m, $W_P = 3\mu$m | | | | | | $R = 1$ KΩ, $C = 1$ pF $W_N = 3\mu$m, $W_N = 9\mu$m | | | | | | $R = 3$ KΩ, $C = 3$ pF $W_N = 3\mu$m, $W_N = 9\mu$m | | | | | |
| | t_{50} (ns) | | | t_{90} (ns) | | | t_{50} (ns) | | | t_{90} (ns) | | | t_{50} (ns) | | | t_{90} (ns) | | |
	Analytic	SPICE	Error	Analytic	SPICE	Error	Analytic	SPICE	Error	Analytic	SPICE	Error	Analytic	SPICE	Error	Analytic	SPICE	Error
1	1.98	2.37	16%	6.59	6.70	2%	1.12	1.13	0%	3.73	3.61	3%	7.53	7.39	2%	25.0	24.4	2%
2	3.11	3.37	6%	5.68	5.67	1%	1.49	1.37	9%	2.62	2.39	8%	8.03	6.11	31%	13.8	11.5	20%
3	3.37	3.45	2%	4.55	4.70	2%	1.51	1.34	12%	2.02	1.89	6%	6.95	5.14	35%	9.55	7.79	22%
4	3.53	3.73	5%	4.71	4.80	0%	1.53	1.46	5%	1.99	1.89	5%	6.46	5.08	27%	8.45	6.94	21%
5	3.62	3.74	3%	4.29	4.46	2%	1.56	1.47	5%	1.82	1.76	3%	6.03	4.81	25%	7.20	6.01	20%
6	3.70	3.92	5%	4.47	4.61	1%	1.58	1.56	1%	1.87	1.82	3%	5.80	4.84	20%	6.92	5.87	18%
7	3.77	3.92	3%	4.23	4.43	3%	1.62	1.58	2%	1.79	1.77	1%	5.59	4.71	19%	6.32	5.47	15%
8	3.83	4.05	5%	4.40	4.56	2%	1.65	1.64	0%	1.85	1.84	1%	5.47	4.77	15%	6.24	5.47	14%
9	3.89	4.06	4%	4.24	4.45	2%	1.69	1.67	4%	1.82	1.81	0%	5.37	4.69	14%	5.88	5.22	13%
10	3.94	4.16	5%	4.39	4.57	2%	1.72	1.73	0%	1.88	1.90	1%	5.30	4.75	11%	5.88	5.29	11%
11	4.00	4.18	4%	4.28	4.51	4%	1.76	1.77	0%	1.87	1.89	1%	5.25	4.73	11%	5.64	5.13	10%
12	4.04	4.26	4%	4.42	4.61	3%	1.80	1.82	1%	1.93	1.96	2%	5.21	4.78	9%	5.67	5.20	9%
13	4.10	4.31	5%	4.34	4.58	4%	1.84	1.86	1%	1.93	1.97	2%	5.18	4.78	8%	5.50	5.11	8%
14	4.14	4.37	5%	4.46	4.67	3%	1.88	1.91	1%	1.99	2.03	2%	5.17	4.82	7%	5.55	5.18	7%
15	4.19	4.44	5%	4.40	4.66	4%	1.92	1.96	2%	2.00	2.05	2%	5.16	4.82	7%	5.43	5.18	5%
16	4.24	4.46	4%	4.51	4.73	3%	1.96	2.00	2%	2.06	2.09	1%	5.16	4.82	7%	5.49	5.11	5%
17	4.29	4.53	5%	4.47	4.74	5%	2.00	2.06	3%	2.07	2.14	3%	5.16	4.90	5%	5.39	5.20	4%
18	4.33	4.63	6%	4.57	4.88	5%	2.04	2.11	3%	2.12	2.21	4%	6.17	4.89	5%	5.46	5.13	6%
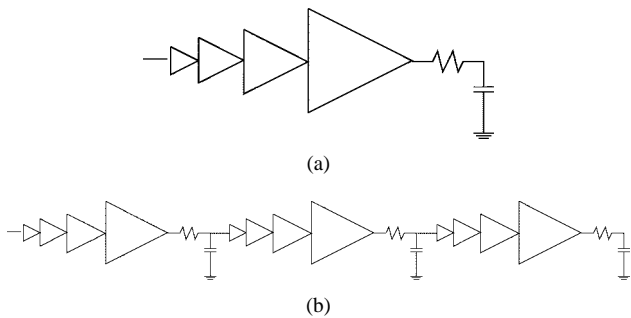
Fig. 9. (a) A single tapered buffer and (b) a three-stage tapered-buffer repeater system. The first stage is a minimum sized repeater. The tapering factor is e.

first glance, this behavior may seem to contradict the data indicated in Fig. 2; however, when each repeater is driving a large RC load, the input waveforms driving the intermediate repeater stages degrade, causing those repeaters with slow input waveforms to operate in the saturation region rather than in the linear region. However, as shown in Table I, with most repeater configurations the error is typically much less than 15%.

V. UNIFORM REPEATERS VERSUS TAPERED BUFFERS AND TAPERED-BUFFER REPEATERS

Depending on the magnitude of the RC load, the form of the repeater buffer structure to minimize the total delay may be expected to change. With larger RC loads or large capacitances, a tapered buffer or a tapered-buffer repeater system (as shown in Fig. 9(a) and (b) may decrease the total delay required to propagate a signal along a resistive line. Intuitively, an interconnect line that is highly capacitive and nonnegligibly resistive may exhibit characteristics similar to a purely capacitive line. Since a purely capacitive line is optimally driven by a tapered buffer [see Fig. 9(a)] [1], [21], a highly capacitive and moderately resistive line may possibly be more efficiently driven by a series of tapered buffers [22], [23]. The application of uniform repeaters versus tapered buffers and

tapered-buffer repeaters on an RC line is therefore discussed in this section.

An estimate of the total delay of a tapered-buffer repeater system is performed in a manner similar to that presented for a uniform repeater system. Some modifications, however, are made to accommodate the use of tapered buffers. C_{rep}, for example, is now the capacitance of a minimum-sized inverter since the first stage of each tapered-buffer repeater is a minimum-sized inverter. The drive current I_{DO} of the tapered-buffer repeater is related to the size of the final buffer in each tapered-buffer repeater stage.

The delay for a single tapered-buffer repeater is

$$
\begin{aligned}
t_{\mathrm{out}} &= t_{p,\,opt} + t_{\mathrm{rep}} \\
&= \ln\left(\frac{C_L}{C_i}\right)t_{p0} + \frac{(1 + \mho_{do}R)(C_{\min} + C_{\mathrm{int}})}{\mho_{do}} \\
&\quad \times \ln\left(\frac{V_{DD}}{V_{\mathrm{out}}}\right).
\end{aligned}
\tag{15}
$$

t_{out} for a tapered-buffer repeater is integrated into a similar expression as (9). The components of (15) are as follows: C_L is the gate capacitance of the final buffer in the repeater; C_i is the input gate capacitance of a minimum-size inverter; and t_{p0} is the propagation delay of a minimum-size inverter driving a capacitance $e \cdot C_i$ [24] since the tapering factor is assumed to be e. For each tapered buffer, the final inverter stage is of size W_{opt} and the number of stages in the repeater is $\ln(W_{\mathrm{opt}})$ (note that this value must be rounded to an integer).

A comparison of the efficacy of tapered buffers and tapered-buffer repeater systems versus uniformly sized repeaters for various loads is shown in Table II. Furthermore, the accuracy of the analytical models for both the uniform and tapered-buffer repeaters versus SPICE is also listed in the same table. The single tapered buffer has been optimized for the specified load capacitance. The results listed in columns 10 and 13 shown in Table II as compared to column five demonstrate the importance of interconnect resistance. Even small resistances have a large effect on the signal delay characteristics. RC loads in which the capacitance is the dominant component

TABLE II
90% OUTPUT TIME FOR OPTIMALLY SIZED UNIFORM REPEATERS, TAPERED-BUFFER REPEATERS,
AND TAPERED BUFFERS FOR VARIOUS LOADS AS COMPARED WITH SPICE

| Total RC Load | Uniform Repeaters | | | | | Tapered-Buffer Repeaters | | | | | | Single Tapered Buffer |
	# of repeaters	W_{opt} μm	Analytical t_{90} (ns)	SPICE t_{90} (ns)	Error %	# repeaters	# of Stages	W_{opt} μm	Analytical $t_{90}(ns)$	SPICE t_{90} (ns)	Error %	SPICE $t_{90}(ns)$
1 KΩ 1 pF	7	13	0.90	0.98	8	5	2	2	3.20	3.03	6	2.8
1 KΩ 5 pF	15	29	2.10	2.18	4	5	3	9	7.36	5.05	45	12.1
5 KΩ 2 pF	33	12	2.96	2.75	8	9	2	2	7.30	5.70	28	23.5
1 KΩ 20 pF	31	56	4.20	4.43	5	5	4	34	15.05	10.1	50	47
1 KΩ 100 pF	67	124	9.46	11.15	15	11	5	75	36.00	18.5	50	> 50

of the interconnect impedance are of primary interest when considering tapered-buffer repeaters. However, as shown in Table II, even when driving RC loads as large as 1 kΩ and 100 pF, the application of uniform repeaters remains more delay efficient than both tapered buffers and tapered-buffer repeaters.

VI. POWER DISSIPATION IN INVERTERS AND REPEATER CHAINS

Power consumption has become one of the premier issues in VLSI circuit design. There are two primary contributions to the total transient power dissipated by a CMOS inverter: dynamic power dissipation and short-circuit power dissipation [25]–[31]. Dynamic power dissipation is quantified by the familiar expression CV^2f, and in repeater chains is due to the input capacitance of each repeater. On the other hand, short-circuit power is often neglected since the dynamic power is assumed to be dominant. As described below and in [25]–[31], the magnitude of the short-circuit power is both input signal and load dependent. It is shown in this section that short-circuit power can be a significant portion of the total transient power dissipation.

An analytical expression for the short circuit power dissipated by a single inverter driving a capacitive load with a ramp input signal is presented in Section VI-A. An analysis of the accuracy of this expression is presented in Section VI-B. A comparison of short-circuit power to the total transient power is made in Section VI-C. Finally, the total power dissipation in a repeater chain is considered in Section VI-D.

A. Short-Circuit Power in a CMOS Inverter

During the temporal region when the input signal is transitioning between V_{TN} and $V_{DD}+V_{TP}$, a dc current path exists between V_{DD} and ground. The excess current dissipated during this region is called the short-circuit (or crossover) current [28]. Short-circuit current is due to a slow input transition, and for a balanced inverter, the peak current occurs near the middle of the input transition. The logic stage following a large RC load may dissipate significant amounts of short-circuit power due to the degraded waveform originating from the CMOS inverter driving an RC load (see Fig. 10). A pulse of short-circuit current is exemplified by the solid line in the lower graph of Fig. 11, i.e., the SPICE-derived data.

The total short-circuit current I_{SC} can be estimated by modeling I_{SC} as a triangle. Therefore, the integral of I_{SC}
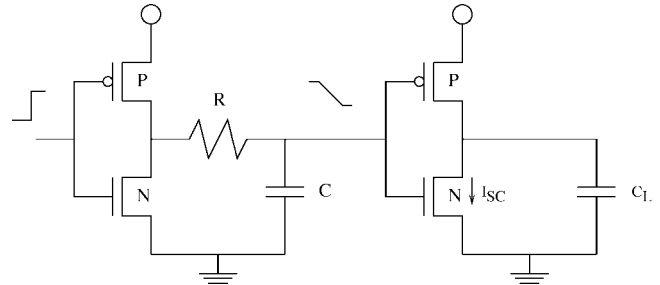

Fig. 10. Nonstep input signal driving CMOS inverter stage creates short-circuit power in the following inverter stage.
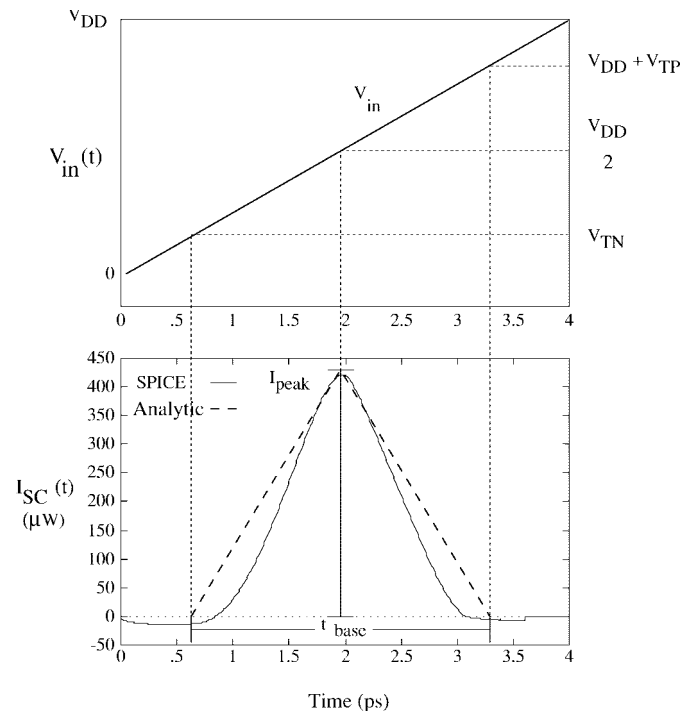

Fig. 11. Graphical estimation of short-circuit current (0.8 μm CMOS technology).

is the area of a triangle, $(1/2)$ base × height. In terms of the short-circuit current, the height can be modeled as I_{peak} and the base can be modeled as t_{base} (see Fig. 11). I_{peak} is the maximum saturation current of the load transistor and depends on both V_{GS} and V_{DS}, therefore, I_{peak} is both input waveform and load dependent. The value of t_{base} is the time during which the P-channel and the N-channel transistors are both turned on,

Load Resistance	Load Capacitance	Power (μW) $V_{DD} = 5.0$V $f = 10$MHz		% Error
		Analytic	SPICE	
10 Ω	.3 pF	1.4	.99	41%
10 Ω	.5 pF	3.9	3.22	21%
10 Ω	1 pF	12.4	11.1	12%
100 Ω	.3 pF	1.71	1.23	39%
100 Ω	.5 pF	4.68	3.83	22%
100 Ω	1 pF	13.8	12.7	9%
1000 Ω	.3 pF	5.85	5.2	12%
1000 Ω	.5 pF	13.0	12.2	7%
1000 Ω	1 pF	34.2	33.8	1%

permitting a dc current path to exist between V_{DD} and ground. This time occurs over the region, $V_{TN} \leq V_{\text{in}} \leq V_{DD} + V_{TP}$. Therefore, t_{base} is the difference between the time to reach the N-channel threshold voltage and the P-channel threshold voltage, $|(t_{V_{TP}} - t_{V_{TN}})|$. The area defined by this triangle is $(1/2)I_{\text{peak}} \times t_{\text{base}}$, which models the total short-circuit current I_{SC} sourced by a CMOS inverter due to a nonstep input signal [17], [29].

The total short-circuit current multiplied by f and V_{DD} is the short-circuit power. The short-circuit power dissipation P_{SC} of the following stage for one transition (either rising or falling edge) can therefore be approximated by

$$P_{SC} = \frac{1}{2} I_{\text{peak}} t_{\text{base}} V_{DD} f \qquad (16)$$

with the expression for t_{base} being

$$t_{\text{base}} = \left| \ln \left(\frac{V_{TN}}{V_{DD} + V_{TP}} \right) \right| \frac{C + \mho_{do} RC}{\mho_{do}}. \qquad (17)$$

By inserting this expression for t_{base} into (16), the short-circuit power dissipation P_{SC} of a CMOS inverter following a lumped RC load over both the rising and falling transitions is

$$P_{SC} = \left| \ln \left(\frac{V_{TN}}{V_{DD} + V_{TP}} \right) \right| \frac{C + \mho_{do} RC}{\mho_{do}} I_{\text{peak}} f V_{DD}. \qquad (18)$$

The results of this expression are compared to SPICE in the following subsection.

B. Accuracy of the Short-Circuit Power Dissipation Expression

The short-circuit power derived from (18) for a wide range of RC loads between the CMOS inverter stages shown in Fig. 10 is compared with SPICE in Table III. The RC load of the driving inverter is described in the first two columns of Table III. The short-circuit power predicted by (18) and derived from SPICE is shown in the third and fourth columns, respectively. The percent error between the analytical expression and SPICE is shown in the final column.

For smaller RC loads, hence, faster transition times, there is negligible short-circuit power since a direct path from the power supply to ground does not exist for any significant time. The short-circuit power becomes nonnegligible when larger interconnect loads between the two CMOS stages cause a transition time of significant magnitude, e.g., a transition time greater than 0.5 ns for a 0.8-μm CMOS inverter. At this

borderline value, the analytical P_{SC} differs from SPICE by a maximum error of 41%. As the RC load and transition time increase, the analytical model more closely predicts the short-circuit current derived from SPICE. For RC loads exceeding 0.1 ns, errors less than 15% are typical. Furthermore, the short-circuit power becomes a significant portion of the total power dissipation when the CMOS inverter is loaded by larger RC loads, creating long transition times. It is this region of highest accuracy that is of greatest interest when considering short-circuit power in resistively loaded CMOS inverters.

The error of the analytical expression for P_{SC} is bounded by the RC time constant characterizing the interconnect load impedance. For 0.8-μm CMOS technology, the percent error is less than 15% for an RC time constant greater than 0.1 ns. For an RC time constant less than 0.1 ns, the percent error increases to approximately 40%.

One source of error in estimating the short-circuit power derived from (18) is found by examining the transition time of the input waveform. Comparing the analytical solution to the transition time generally yields pessimistic results when compared to SPICE (see Table I). By inserting these pessimistic transition times into (18), the resulting short-circuit power is also pessimistic, as demonstrated by the data presented in Table III.

Another source of error not modeled in these repeater delay and power equations is caused by signal overshoot of fast transient waveforms. This overshoot may increase V_{DS} above V_{DD} or below ground and is caused by the diffusion capacitance of the inverter. This overshoot occurs early during the transition time and causes current to flow opposite to the expected direction, thereby reducing the total short-circuit current. This behavior, in turn, reduces the total short-circuit power, increasing the discrepancy between SPICE and (18), which does not consider transient overshoot. The phenomenon of signal overshoot can be observed in Fig. 11.

C. Short-Circuit Power as Compared to the Total Transient Power in a CMOS Inverter

For a given supply voltage and frequency, dynamic power dissipation depends only on the load capacitance and does not depend on the input waveform shape or load resistance. In contrast, the short-circuit power dissipation changes with both input waveform shape and output load resistance and capacitance. The ratio of the short-circuit power to the total transient power (the sum of the dynamic and short-circuit power) of a CMOS inverter with respect to the load resistance R for three values of load capacitance C is shown in Fig. 12. Note that with increasing load resistance, the short-circuit power dissipation cannot be neglected, since, as shown in Fig. 12, it can comprise more than 20% of the total transient power dissipation.

D. Power Dissipation in Repeater Chains

As the input transition slows, more short-circuit power is dissipated within the repeater stage. The input signal transition time is dependent upon the number of repeaters in the chain. If additional repeaters are inserted into a line to drive a long resistive interconnect, each repeater drives a smaller RC load with
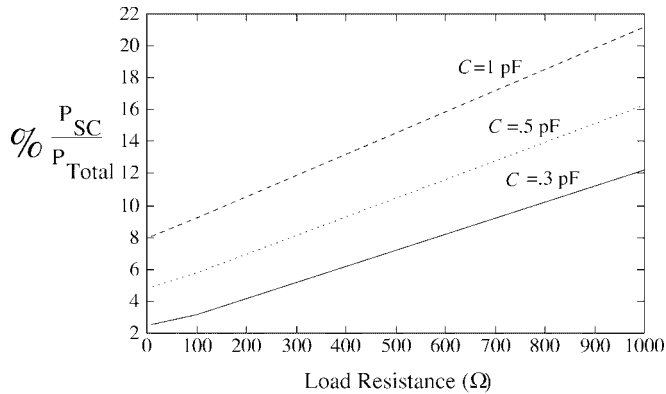
Fig. 12.   Ratio of short-circuit power to total transient power versus load resistance for varying load capacitance.
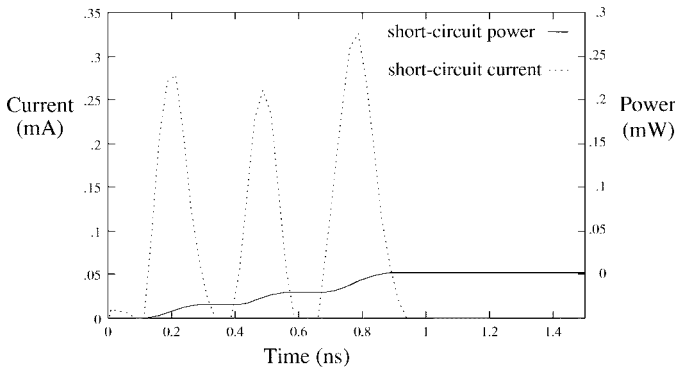

Fig. 13.   Short-circuit current and power dissipated in a four-stage repeater system with $W_N = 5$ μm and $W_P = 15$ μm, $f = 10$ MHz.
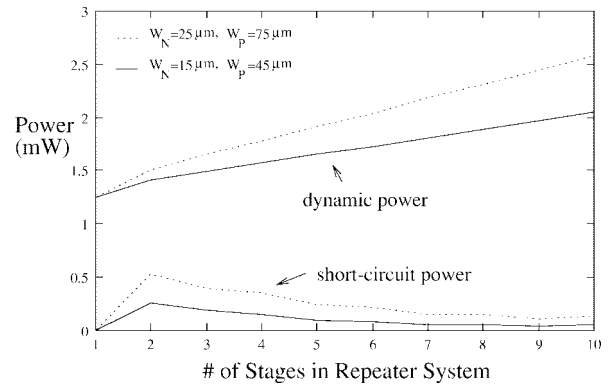

Fig. 14.   The short-circuit and dynamic power dissipation versus the number of stages in a repeater system. Note the small increase in short-circuit power from nine to ten stages for the larger sized repeater due to the increase in peak current with negligible improvement in transition time.

a waveform exhibiting a faster transition time, permitting the input transition of the following repeater to be faster. However, these additional repeaters may increase the short-circuit power of the total repeater system. The peak short-circuit current, which is proportional to the device width, is the other primary factor that determines the magnitude of the short-circuit power [17], [24], [29]. An example of short-circuit current and power in a repeater chain is shown in Fig. 13.

Simulations demonstrate that when device sizes are small, the contribution of short-circuit power is small in comparison to the dynamic power, typically ranging from 1% to 5%. As the geometric width of the repeaters is increased, the contribution of the short-circuit and dynamic power also increases. However, as the geometric width and the number of repeaters increase, dynamic power increases linearly, whereas short-circuit power changes nonlinearly. A comparison of short-circuit power versus dynamic power of a repeater system driving an RC load of 1 kΩ and 1 pF is shown in Fig. 14. Both the short-circuit power and the dynamic power dissipated within the repeater chain versus the number of repeaters are shown. For the larger sized repeater, the peak short-circuit power is about 30% of the dynamic power at two stages; at five stages the short-circuit power is 12% of the dynamic power; and at nine stages, about 5%. A five stage repeater system provides the minimum transition time for this RC load. Thus, reducing the repeater size to 15 and 45 μm from 25 and 75 μm saves 40% in area (≈ 200 μm^2), reduces the short-circuit

power by 60%, and reduces the dynamic power by 12% in return for a 5% increase in t_{50} delay. Note that the maximum short-circuit power savings occurs when the input transition time of each repeater is approximately equal to the repeater output transition time [24], [30].

VII. CONCLUSION

A closed form timing model of a CMOS inverter driving a resistive-capacitive load based on the α-power law device model has been presented. This new analytical expression differs from previous work because the short-channel transistor effect of velocity saturation is considered. The timing model for a CMOS inverter has been expanded to determine the overall delay of a signal propagating through a uniform repeater chain driving a large distributed RC load. Analytical estimates of delay with these design equations are within 16% of SPICE for loads representative of long resistive interconnect.

The performance characteristics of uniform and tapered-buffer repeaters are compared for a variety of RC loads. The resistance in RC lines is found to have a larger than expected effect on the delay of a signal propagating along a long line. Uniform repeaters outperform tapered buffers and tapered-buffer repeaters when driving even relatively low resistive RC loads. It is thus more advantageous to use a number of small uniform repeaters rather than a few (or one) tapered-buffer repeaters.

Power dissipation in CMOS inverters and repeaters driving RC loads has also been investigated. A closed form analytic expression for short-circuit power in a CMOS inverter driving an RC load is presented. In the region of interest, this expression exhibits a maximum error of 15% as compared to SPICE. It is also shown that short-circuit power can represent up to 20% of the total dynamic power dissipation. An empirical comparison of power in repeater chains is presented. The application of the repeater expressions developed in this paper to a specific repeater implementation demonstrate that a 4% decrease in input to output delay can be traded off for a 40% savings in area and a 15% savings in power.

In summary, inserting repeaters into an RC line can greatly improve the signal delay characteristics. A design method-

ology for accurately inserting repeaters into an *RC* line is presented in this paper.

REFERENCES

[1] H. B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI.* Reading, MA: Addison-Wesley, 1990.
[2] S. Bothra, B. Rogers, M. Kellam, and C. M. Osburn, "Analysis of the effects of scaling on interconnect delay in ULSI circuits," *IEEE Trans. Electron Devices,* vol. 40, pp. 591–597, Mar. 1993.
[3] H. B. Bakoglu and J. D. Meindl, "Optimal interconnection circuits for VLSI," *IEEE Trans. Electron Devices,* vol. ED-32, pp. 903–909, May 1985.
[4] C. Y. Wu and M. Shiau, "Accurate speed improvement techniques for RC line and tree interconnections in CMOS VLSI," in *Proc. IEEE Int. Symp. Circuits and Systems,* May 1990, pp. 2.1648–2.1651.
[5] ———, "Delay models and speed improvement techniques for RC tree interconnections among small-geometry CMOS inverters," *IEEE J. Solid-State Circuits,* vol. 25, pp. 1247–1256, Oct. 1990.
[6] H. Shichman and D. A. Hodges, "Modeling and simulation of insulated-gate field-effect transistor switching circuits," *IEEE J. Solid-State Circuits,* vol. SC-3, pp. 285–289, Sept. 1968.
[7] M. Nekili and Y. Savaria, "Optimal methods of driving interconnections in VLSI circuits," in *Proc. IEEE Int. Symp. Circuits and Systems,* May 1992, pp. 21–23.
[8] ———, "Parallel regeneration of interconnections in VLSI & ULSI circuits," *Proc. IEEE Int. Symp. Circuits and Systems,* May 1993, pp. 2023–2026.
[9] S. Dhar and M. A. Franklin, "Optimum buffer circuits for driving long uniform lines," *IEEE J. Solid-State Circuits,* vol. 26, pp. 32–40, Jan. 1991.
[10] C. Tretz and C. Zukowski, "CMOS transistor sizing for minimization of energy-delay product," in *Proc. IEEE Great Lakes Symp. VLSI,* Mar. 1996, pp. 168–173.
[11] C. Zukowski and C. Tretz, "Transistor sizing in CMOS logic chains to minimize energy-delay product," in *Proc. Workshop on Academic Electronics in New York State,* June 1996, pp. 221–226.
[12] C. J. Alpert, "Wire segmenting for improved buffer insertion," in *Proc. 34th IEEE/ACM Design Automation Conf.,* June 1997.
[13] T. Sakurai and A. R. Newton, "Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas," *IEEE J. Solid-State Circuits,* vol. 25, pp. 584–594, Apr. 1990.
[14] ———, "A simple short-channel MOSFET model and its application to delay analysis of inverters and series-connected MOSFET's," in *Proc. IEEE Int. Symp. Circuits and Systems,* May 1990, pp. 105–108, May 1990.
[15] A. P. Chandrakasan, S. Sheng, and R. W. Broderson, "Low-power CMOS digital design," *IEEE J. Solid-State Circuits,* vol. 27, pp. 473–483, Apr. 1992.
[16] D. Dobberpuhl *et al.,* "A 200 MHz 64-b dual issue CMOS microprocessor," *IEEE J. Solid-State Circuits,* vol. 27, pp. 1555–1567, Nov. 1992.
[17] V. Adler and E. G. Friedman, "Delay and power expressions for a CMOS inverter driving a resistive-capacitive load," *Analog Integrated Circuits for Signal Processing,* vol. 14, no. 1/2, pp. 29–40, Sept. 1997.
[18] A. Vladimirescu and S. Liu, "The simulation of MOS integrated circuits using SPICE2," Univ. of Calif., Berkeley, ERL Memo M80/7, Oct. 1980.
[19] W. C. Elmore, "The transient response of damped linear networks with particular regard to wideband amplifiers," *J. Appl. Phys.,* vol. 19, no. 1, pp. 55–63, Jan. 1948.
[20] V. Adler and E. G. Friedman, "Repeater design to reduce delay and power in resistive interconnect," in *Proc. IEEE Int. Symp. Circuits and Systems,* June 1997, pp. 2148–2151.
[21] R. C. Jaeger, "Comments on 'An optimized output stage for MOS integrated circuits'," *IEEE J. Solid-State Circuits,* vol. SC-10, pp. 185–186, June 1975.
[22] B. S. Cherkauer and E. G. Friedman, "A unified design methodology for CMOS tapered buffers," *IEEE Trans. VLSI Syst.*, vol. 3, pp. 99–111, Mar. 1995.
[23] ———, "Design of tapered buffers with local interconnect capacitance," *IEEE J. Solid State Circuits*, vol. 30, pp. 151–155, Feb. 1995.
[24] J. M. Rabaey, *Digital Integrated Circuits.* Englewood Cliffs, NJ: Prentice-Hall, 1996.
[25] L. Bisdounis, S. Nikolaidis, O. Koufopavlou, and C. E. Goutis, "Modeling the CMOS short-circuit power dissipation," in *Proc. IEEE Int. Symp. Circuits and Systems,* May 1966, pp. 4.469–4.472.
[26] A. M. Hill and S.-M. Kang, "Statistical estimation of short-circuit power in VLSI circuits," in *Proc. IEEE Int. Symp. Circuits and Systems,* May 1996, pp. 4.105–4.108.
[27] A. Hirata, H. Onodera, and K. Tamaru, "Estimation of short-circuit power dissipation for static CMOS gates," *IEICE Trans. Fundamentals of Electron., Commun., Comput. Sci.,* vol. E79-A, no. 3, pp. 304–311, Mar. 1996.
[28] ———, "Estimation of short-circuit power dissipation and its influence on propagation delay for static CMOS gates," in *Proc. IEEE Int. Symp. Circuits and Systems,* May 1996, pp. 4.751–4.754.
[29] V. Adler and E. G. Friedman, "Timing and power models for CMOS repeaters driving resistive interconnect," in *Proc. IEEE ASIC Conf.,* Sept. 1996, pp. 201–204.
[30] H. J. M. Veendrick, "Short-circuit dissipation of static CMOS circuitry and its impact on the design of buffer circuits," *IEEE J. Solid-State Circuits,* vol. SC-19, pp. 468–473, Aug. 1984.
[31] S. R. Vemuru and N. Scheinberg, "Short-circuit power dissipation estimation for CMOS logic gates," *IEEE Trans. Circuits Syst. I: Fundamental Theory and Applications,* vol. 41, no. 11, pp. 762–766, Nov. 1994.

Victor Adler (S'97) received the B.S. degree in electrical engineering and the B.A. degree in computer science from Duke University, Durham, NC, in 1992, the M.S. degree in electrical engineering from the University of Rochester, Rochester, NY, in 1993, and is currently working toward the Ph.D. degree in electrical engineering at the University of Rochester.

He was an IBM Watson Scholar and worked pre-professionally at IBM Microelectronics, Burlington, VT between 1988 and 1992 in the areas of final module test, packaging, circuit macro development, and standard cell design. He has been a Teaching and Research Assistant at the University of Rochester since 1993. In 1997, he worked on clock skew scheduling at Intel Corp. His research interests include design techniques for high performance CMOS and superconductive technologies.

Eby G. Friedman (S'78–M'79–SM'90) was born in Jersey City, NJ, in 1957. He received the B.S. degree from Lafayette College, Easton, PA, in 1979, and the M.S. and Ph.D. degrees from the University of California, Irvine, in 1981 and 1989, respectively, all in electrical engineering.

He was with Philips Gloeilampen Fabrieken, Eindhoven, The Netherlands, in 1978, where he worked on the design of bipolar differential amplifiers. From 1979 to 1991, he was with Hughes Aircraft Company, rising to the position of Manager of the Signal Processing Design and Test Department, responsible for the design and test of high performance digital and analog IC's. He has been with the Department of Electrical Engineering at the University of Rochester, Rochester, NY, since 1991, where he is an Associate Professor and Director of the High Performance VLSI/IC Design and Analysis Laboratory. His current research and teaching interests are in high performance microelectronic design and analysis with application to high speed portable processors and low power wireless communications. He has authored many papers and book chapters on the fields of high speed and low power CMOS design techniques and the theory and application of synchronous clock distribution networks, and has edited three books, *Clock Distribution Networks in VLSI Circuits and Systems* (IEEE Press, 1995), *High Performance Clock Distribution Networks* (Kluwer Academic Press, 1997), and *Analog Design Issues in Digital VLSI Circuits and Systems* (Kluwer Academic Press, 1997).

Dr. Friedman is a member of the editorial board of *Analog Integrated Circuits and Signal Processing,* a member of the CAS BoG, and a member of the technical program committee of a number of conferences. He was a member of the editorial board of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: ANALOG AND DIGITAL SIGNAL PROCESSING, Chair of the IEEE Rochester Section, Chair of the VLSI track for ISCAS '96 and '97, Technical Co-Chair of the 1997 IEEE International Workshop on Clock Distribution Networks, Editor of several special issues in a variety of journals, and a recipient of the Howard Hughes Masters and Doctoral Fellowships, an Outstanding IEEE Chapter Chairman Award, and a University of Rochester College of Engineering Teaching Excellence Award.