Analytical Delay Models for VLSI Interconnects Under Ramp Input

Andrew B. Kahng, Kei Masuko, and Sudhakar Muddu*

UCLA Computer Science Department, Los Angeles, CA 90095-1596 USA
Cadence Design Systems, Inc., San Jose CA 95134
Silicon Graphics, Inc., Mountain View, CA 94039

abk@cs.ucla.edu, masuko@cs.ucla.edu, sudhakar@cs.ucla.edu

Abstract

Elmore delay has been widely used as an analytical estimate of interconnect delays in the performance-driven synthesis and layout of VLSI routing topologies. However, for typical *RLC* interconnections with ramp input, Elmore delay can deviate by up to 100% or more from SPICE-computed delay since it is independent of rise time of the input ramp signal. We develop new analytical delay models based on the first and second moments of the interconnect transfer function when the input is a ramp signal with finite rise time. Delay estimates using our first moment based analytical models are within 4% of SPICE-computed delay, and models based on both first and second moments are within 2.3% of SPICE, across a wide range of interconnect parameter values. Evaluation of our analytical models is several orders of magnitude faster than simulation using SPICE. We also describe extensions of our approach for estimation of source-sink delays in arbitrary interconnect trees.

1 Introduction

Accurate calculation of propagation delay in VLSI interconnects is critical to the design of high speed systems, and transmission line effects now play an important role in determining interconnect delays and system performance. Existing techniques are based on either *simulation* or (closed-form) *analytical formulas*. Simulation methods such as SPICE give the most accurate insight into arbitrary interconnect structures, but are computationally expensive. Faster methods based on moment matching techniques are proposed in [12, 13, 14, 17], but are still too expensive to be used during layout optimization. Thus, Elmore delay [2], a first order approximation of delay under step input, is still the most widely used delay model in the performance-driven synthesis of clock distribution and Steiner global routing topologies. However, Elmore delay cannot be applied to estimate the delay for interconnect lines with ramp input source; this inaccuracy is harmful to current performance-driven routing methods which try to determine optimal interconnect segment lengths and widths (as well as driver sizes). Previous moment-based approaches [12, 14, 17] can compute a response for interconnects under ramp input within a simulation-based methodology, but no previous work has given *analytical* delay estimation models based on the first few moments.

Recently, [3] presented lower and upper bounds for the ramp input response; their delay model is the same as the Elmore model for ramp input (we refer to this model as analytical ramp input model (T_{AD}) in this paper). Delay estimates for the analytical ramp input model are off by as much as 50% from SPICE-computed delays for 50% threshold voltage, and the analytical ramp input model cannot be used to obtain threshold delay for various threshold voltages. The authors of [5] used Elmore delay as an upper bound on the 50% threshold delay for *RC* interconnection lines under arbitrary input waveforms. However, we find that Elmore delay is not at all close to SPICE-computed 50% threshold delay and, depending on the input slew time and driver resistance, can be either greater or less than SPICE-computed delay (see Section 7 below). This paper gives a new and accurate *analytical* delay estimate for distributed *RLC* interconnects under ramp input. To experimentally validate our analysis and delay formula, we model VLSI interconnect lines having various combinations of source, and load parameters, apply different input rise times, and obtain delay estimates from SPICE, Elmore delay and the proposed analytical delay model. Over our range of test cases, Elmore delay estimates can vary by as much as 100% from SPICE-computed delays. As the input rise time increases, Elmore delay deviates even further from SPICE-computed delays. In contrast, our single-pole delay estimates are within 4% of SPICE delays and our two-pole delay estimates are within 2.3% of SPICE delays.[1] Since our analytical models have the same time complexity of evaluation as the Elmore model, we believe that they are very useful for performance-driven routing methodologies.

The organization of our paper is as follows. In Section 2 we discuss delay models which have been previously proposed for interconnect lines under step input. Section 3 presents a new analytical delay definition for interconnect lines under ramp input. Section 4 discusses various threshold delay models for single-pole approximation of the interconnect transfer function; Section 5 gives various threshold delay models for two-pole approximation; and Section 6 extends our delay modeling approach to interconnection trees. Section 7 concludes with experimental results for various combinations of input rise times and interconnect parameters.

2 Previous Delay Models Under Step Input

The transfer function of an *RLC* interconnect line with source and load impedance (Figure 1) can be obtained using ABCD parameters [1] as

$$H(s) = \frac{1}{\left[\cosh(\theta h) + \frac{Z_S}{Z_0}\sinh(\theta h)\right] + \frac{1}{Z_T}\left[Z_0\sinh(\theta h) + Z_S\cosh(\theta h)\right]} = \frac{1}{1 + b_1 s + b_2 s^2 + \ldots + b_k s^k + \ldots} \quad (1)$$

where $\theta = \sqrt{(r+sl)sc}$ is the propagation constant and $Z_0 = \sqrt{\frac{R+sL}{sC}}$ is the characteristic impedance; $r = \frac{R}{h}, l = \frac{L}{h}, c = \frac{C}{h}$ are resistance, inductance, and capacitance per unit length and h is the length of the line. The variables b_k are called the coefficients of the transfer function and are directly related to the moments of the transfer function [8]. Expanding the transfer function into a Maclaurin series of s around $s = 0$ leads to an infinite series, and to compute the response the series is truncated to desired order. The method of Padé approximation has been widely used to compute the response from the transfer function [11, 12]. For the case of resistive source (R_S) and capacitive load (C_L) impedances, the coefficient of s in the transfer function can be obtained as [8] $b_1 = R_S C + R_S C_L + \frac{RC}{2} + RC_L$.

Efficient delay estimates for interconnect lines are typically derived by considering a single interconnect line with resistive source and capacitive load impedances; delay formulas for an interconnect tree come from recursive application of the formula for a single line. Elmore delay [2] is a first order delay estimate for interconnect lines under step input. It is equal to the first moment of the system impulse response, i.e., the coefficient of s or the first moment in the system transfer function $H(s)$. Applying this definition to $H(s)$ in Equation (1), we see that the Elmore delay is equal to the coefficient b_1.

*Principal author, to whom correspondence should be addressed: Sudhakar Muddu (muddu@mti.sgi.com). Andrew B. Kahng is currently on Sabbatical leave at Cadence Design Systems, ak@cadence.com.

[1] We use threshold delay to refer to delay measured from the point when the input signal is zero. To compute delay *relative to* the input signal, subtract the corresponding threshold delay of the input signal (e.g., for 50% threshold voltage, the delay for the input ramp is $\frac{T_R}{2}$).

ICCAD '96
1063-6757/96 $5.00 © 1996 IEEE

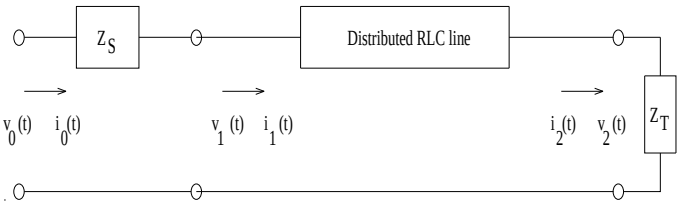

Figure 1: 2-port model of a distributed *RLC* line with source impedance Z_S and load impedance Z_T.

By considering only one pole in the transfer function, i.e, approximating the denominator polynomial to only the first moment, the single pole response can be obtained as in [4, 15]. The single pole of the transfer function is equal to the inverse of the Elmore delay T_{ED}. Hence, the delay at arbitrary thresholds of the single pole response can be directly related to Elmore delay (Elmore delay actually corresponds to the 63.2% threshold voltage of the single pole response). For example, delay at 50% threshold voltage is $0.69b_1$, and delay at 90% threshold voltage is $2.3b_1$. Although Elmore delay has been widely used for interconnect timing analysis, it cannot accurately estimate the delay for *RLC* interconnect lines, which are the appropriate representation for interconnects whose inductive impedance cannot be neglected [6].[2] More critically, Elmore delay cannot estimate delays when the input signal is a ramp.

3 Analytical Ramp Delay Definitions

In practice, the input at any gate or root of a tree is a ramp with finite rise (or fall) time, and there are no published analytical delay models for ramp input. We now propose various ramp delay definitions and also compute analytical delay expressions using the first one or two moments of the transfer function. We discuss delay models for rising ramp input only, since our analyses can be easily extended for falling ramp input [9].

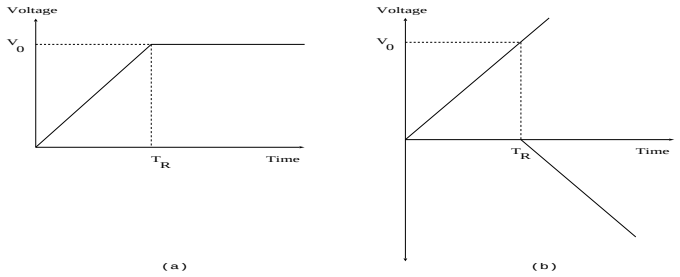

Figure 2: A ramp input function: (a) finite ramp with rise time T_R, and (b) finite ramp decomposed into two shifted infinite ramps.

Rising Ramp Input

The finite rising ramp input shown in Figure 2 can be expressed in the time domain as

$$v_{in}(t) = \frac{V_0}{T_R}[tU(t) - (t-T_R)U(t-T_R)] \quad \text{for all } t \geq 0$$

where $U(t)$ denotes the step function. The finite ramp input in the transform domain is $V_{in}(s) = \frac{V_0}{T_R} \cdot \frac{1}{s^2}[1-e^{-sT_R}]$. In the transform domain, the output response is

$$V_{out}(s) = V_{in}(s)H(s) = \frac{V_0}{T_R} \cdot \frac{1}{s^2}[1-e^{-sT_R}]H(s) \ .$$

[2] Recently, [8] have developed a more accurate analytical delay model considering inductive effects based on the first and second moments of the transfer function. Their model gives accurate estimates compared to SPICE-computed delays, but is valid only for step inputs.

We now give two distinct derivations of an analytical ramp delay estimate.

Elmore Definition.

Applying Elmore's original definition of delay for step input [2] yields an analytical delay T_{AD} for ramp input, i.e.,

$$T_{AD} = \frac{1}{V_0}\int_0^\infty t v'_{out}(t)dt \qquad (2)$$

where $v'_{out}(t)$ is the derivative of the output response under finite ramp input. Taking the Laplace transform of $v'_{out}(t)$,

$$V'_{out}(s) = \int_0^\infty v'_{out}(t)dt - s\int_0^\infty t v'_{out}(t)dt + \ldots$$

Equation (2) then implies that the analytical ramp input delay T_{AD} in the time domain is equal to the first moment of the derivative of the response. In the transform domain, T_{AD} is equal to the first moment (or coefficient of s) of the function $\frac{V'_{out}(s)}{V_0}$, which is equal to $s \cdot \frac{V_{out}(s)}{V_0}$. The derivative of the response in the transform domain is

$$V'_{out}(s) = V_0(1 - \frac{sT_R}{2} + \ldots)\frac{1 + a_1 s + a_2 s^2 + \ldots.}{1 + b_1 s + b_2 s^2 + \ldots.}$$

Therefore, the analytical ramp input delay is

$$T_{AD} = \frac{T_R}{2} + b_1 - a_1 = \frac{T_R}{2} + T_{ED} \qquad (3)$$

where T_{ED} is the Elmore delay for a step input (i.e., the first moment of the transfer function). Another definition of delay based on the formula given in [10] yields the same result of Equation (3) [9].

Group Delay Definition.

The concept of group delay was initially defined for step input by Vlach et al. [18]. We now give a group delay definition for computing ramp input delay similar to that in [18], and show that it converges to the same analytical expression of Equation (3).

Recall that *group delay* is defined as the negative of the rate of change of the phase characteristic ϕ of the output response $V_{out}(\omega)$ with respect to frequency, at zero frequency, i.e., $T_{GD} = \lim_{\omega \to 0} -\frac{\partial \phi}{\partial \omega}$. To compute the phase characteristic of the output response, we first compute the output response $V_{out}(s)$ in the transform domain and then substitute for the Laplace variable $s = j\omega$, i.e.,

$$\begin{aligned} V_{out}(\omega) &= \frac{V_0}{T_R} \cdot \frac{-1}{\omega^2}(1 - e^{-j\omega T_R}) \cdot H(\omega) \\ &= \frac{-V_0}{T_R \omega}\left[(\frac{\omega T_R^2}{2} - \frac{\omega^3 T_R^4}{3!} + \ldots) + j(T_R - \frac{\omega^2 T_R^3}{3!} + \ldots)\right] H(\omega) \\ &= \frac{-V_0}{T_R \omega}[M_1 + jM_2]H(\omega) \end{aligned}$$

where M_1 and M_2 are the real and imaginary parts of the input ramp function. Writing the transfer function in terms of numerator and denominator polynomials,

$$H(\omega) = \frac{(1 - a_2\omega^2 + \ldots) + j(a_1\omega - a_3\omega^3 + \ldots)}{(1 - b_2\omega^2 + \ldots) + j(b_1\omega - b_3\omega^3 + \ldots)} = \frac{N_1 + jN_2}{D_1 + jD_2}$$

Then, the phase characteristic of the output response is $\phi = \tan^{-1}\frac{M_2}{M_1} + \tan^{-1}\frac{N_2}{N_1} - \tan^{-1}\frac{D_2}{D_1}$. We obtain the group delay as

$$T_{GD} = \lim_{\omega \to 0} -\frac{\partial \phi}{\partial \omega} = \frac{T_R}{2} + b_1 - a_1$$

4 Single-Pole Analysis

If we approximate the system transfer function up to the first moment (or coefficient of s), $H(s) \approx \frac{1}{1+sb_1}$. Then, the output response under infinite ramp is[3]

$$U_{out}(s) = \frac{V_0}{T_R}\frac{1}{s^2}\frac{1}{1+sb_1} = \frac{V_0}{T_R}\left[\frac{1}{s^2} - \frac{b_1}{s} + \frac{b_1}{(s+1/b_1)}\right]$$

with corresponding time-domain response

$$u_{out}(t) = \frac{V_0}{T_R}\left[-b_1 + t + b_1 e^{\frac{-t}{b_1}}\right] \quad (4)$$

The time-domain response for a finite ramp is therefore

$$\begin{aligned} v_{out}(t) &= u_{out}(t) - u_{out}(t-T_R) \\ &= \frac{V_0}{T_R}\left[T_R + b_1 e^{\frac{-t}{b_1}} - b_1 e^{\frac{-(t-T_R)}{b_1}}\right] \end{aligned} \quad (5)$$

Note that as $t \to \infty$, $v_{out}(t)$ tends to a final value of V_0 as expected. [16] used a similar single-pole analysis to compute delay and transition times solving the above response equations by applying Newton-Raphson iteration.

4.1 Analytical Delay Model

It turns out that using the analytical ramp delay computed using the definition in Section 3 and the output response given in Equations (4) and (5) leads to the same result:

$$\begin{aligned} T_{AD} &= \frac{1}{V_0}\int_0^{T_R} t u'_{out}(t)dt + \frac{1}{V_0}\int_{T_R}^{\infty} t v'_{out}(t)dt \\ &= \frac{T_R}{2} + b_1 \end{aligned}$$

Threshold Voltage Corresponding To Analytical Ramp Delay.

Section 3 gave two different methods for computing an analytical ramp input delay from the output response. The threshold voltage corresponding to this analytical delay is not known, and must be computed by substituting T_{AD} for time in either the infinite or the finite ramp responses.
Computing the threshold voltage for the infinite ramp response in Equation (4) for $b_1 \ll \frac{T_R}{2}$, we get

$$u_{out}(t = T_{AD}) = \frac{V_0}{2}\left[1 + \frac{2b_1}{T_R}e^{-(1+\frac{1}{2b_1/T_R})}\right].$$

In the limit as $\frac{2b_1}{T_R} \to 0$ the threshold voltage reduces to $u_{out}(t = T_{AD}) = \frac{V_0}{2}$. Hence, for large rise-times or small first moment of the transfer function the analytical delay T_{AD} corresponds to 50% threshold voltage. When $b_1 \gg \frac{T_R}{2}$, using the finite ramp response in Equation (5) gives

$$v_{out}(t = T_{AD}) = V_0\left[1 + \frac{1}{2e}\frac{2b_1}{T_R}(e^{\frac{-1}{2b_1/T_R}} - e^{\frac{1}{2b_1/T_R}})\right].$$

In the limit as $\frac{2b_1}{T_R} \to \infty$ the threshold voltage reduces to $v_{out}(t = T_{AD}) = V_0(1 - 1/e) = 0.632V_0$. Hence, for small rise-times or large first moment of the transfer function the analytical delay T_{AD} corresponds to 63.2% threshold voltage. We see that for any choice of T_R and b_1 the threshold voltage corresponding to the analytical delay T_{AD} will be between 50% and 63.2%.

[3] In the transform and time domains, we respectively use $U(x,s)$ and $u(x,t)$ to indicate the response for infinite ramp input, and $V(x,s)$ and $v(x,t)$ to indicate the response for finite ramp input.

4.2 Threshold Delay Models

Condition for Computing Threshold Delay Using Finite or Infinite Ramp Response.

The ramp input delay at any threshold voltage can be computed using the infinite ramp response in Equation (4) if the ramp delay is less than the rise time T_R, or using the finite ramp response in Equation (5) if the ramp delay is greater than T_R. For example, the delay at threshold $Th1$ in Figure 3 is computed using the infinite ramp response, and the delay at threshold $Th2$ is computed using the finite ramp response. To determine when the infinite ramp response should be used, we write the threshold voltage corresponding to the rise-time T_R in terms of interconnect and rise time parameters:

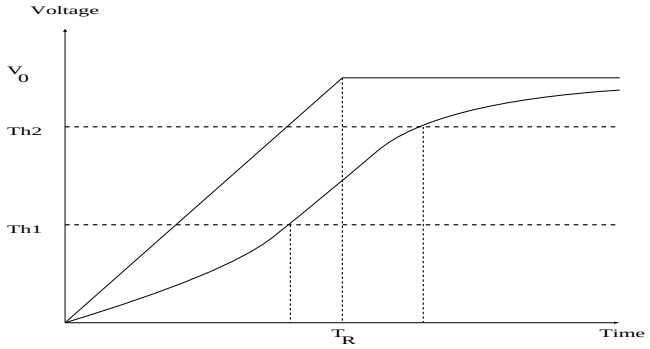

Figure 3: Ramp input delay at various threshold voltages.

$$v_{T_R} = \left[1 - \frac{b_1}{T_R}(1 - e^{\frac{-1}{b_1/T_R}})\right] \quad (6)$$

Here, v_{T_R} is the threshold voltage at which the delay through the interconnect is equal to T_R. Let v_{th} be the threshold voltage of interest for the finite ramp response, expressed as a fraction of the steady state voltage V_0. If $v_{th} \le v_{T_R}$, delay is calculated using Equation (4), and if $v_{th} > v_{T_R}$, delay is calculated using Equation (5).

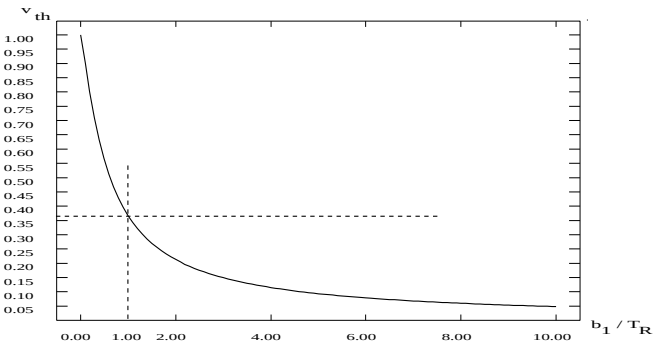

Figure 4: Variation of threshold voltage at delay equal to rise-time T_R with respect to the factor $\frac{b_1}{T_R}$.

Observe that Equation (6) can be rearranged to obtain a condition on $\frac{b_1}{T_R}$ for any given threshold voltage v_{th}: the condition for delay calculation using infinite ramp response is

$$\frac{b_1}{T_R}(1 - e^{\frac{-1}{b_1/T_R}}) \le (1 - v_{th})$$

and the condition for delay calculation using finite ramp response is

$$\frac{b_1}{T_R}(1 - e^{\frac{-1}{b_1/T_R}}) \ge (1 - v_{th})$$

Figure 4 shows the variation of v_{T_R} with respect to the factor $\frac{b_1}{T_R}$. At $b_1 = T_R$ the threshold voltage v_{T_R} is $0.368V_0$, i.e., 36.8%. Since most sub-micron interconnect networks have small rise-times and large propagation delays, the delays at threshold voltages of interest (50% or 90%) will likely be computed by considering the finite ramp response as developed in Equation (9) below.[4]

Threshold Delay Using Infinite Ramp Response.

Model 1. For the infinite ramp response of Equation (4), the threshold delay is

$$T_{RD1} + b_1 e^{\frac{-T_{RD1}}{b_1}} = u_{th}T_R + b_1$$

where u_{th} is the threshold voltage of interest for the infinite ramp response. We can solve such a recursive equation in less than 10 iterations of simple back-substitution (with T_{AD} as the starting value) for all the interconnect configurations we considered. To obtain a closed-form delay formula, we approximate T_{RD1} in the exponential term with some $f(T_{AD})$, which yields

$$T_{RD1} = u_{th}T_R + b_1(1 - e^{\frac{-f(T_{AD})}{b_1}}) \ . \qquad (7)$$

Here, $f(T_{AD})$ depends on the threshold voltage and T_{AD}.[5] The above delay estimate can be improved by expressing T_{RD1} as $u_{th}T_R + \tau_{RD1}$, since the threshold delay for the infinite ramp response T_{RD1} is greater than the threshold delay for the infinite ramp input $u_{th}T_R$. Making this change in delay variable in Equation (4), we get

$$-b_1 + \tau_{RD1} + b_1 e^{-\frac{u_{th}T_R + \tau_{RD1}}{b_1}} = 0$$

Expanding $e^{-\frac{\tau_{RD1}}{b_1}}$ as a Taylor series and considering only the first three terms yields

$$\tau_{RD1}^2 + 2b_1(e^{\frac{u_{th}T_R}{b_1}} - 1)\tau_{RD1} - 2b_1^2(e^{\frac{u_{th}T_R}{b_1}} - 1) = 0$$

Solving for τ_{RD1} in the above equation, the threshold delay can be expressed as

$$T_{RD1} = u_{th}T_R + b_1\left(1 - e^{\frac{u_{th}T_R}{b_1}} + \sqrt{e^{\frac{2u_{th}T_R}{b_1}} - 1}\right) \qquad (8)$$

Using this T_{RD1} value for $f(T_{AD})$ in the exponential term of the Equation (7), we obtain delay values that are very close to the values obtained by solving the equation through iteration.

Threshold Delay Using Finite Ramp Response.

Model 2. For the finite ramp response of Equation (5),

$$v_{th} = \frac{1}{T_R}\left[T_R + b_1 e^{\frac{-T_{RD2}}{b_1}} - b_1 e^{\frac{-(T_{RD2} - T_R)}{b_1}}\right] \ .$$

Collecting the threshold delay T_{RD2} terms, we obtain

$$T_{RD2} = b_1\left|\ln\left(\frac{b_1}{T_R}\cdot\frac{(e^{\frac{1}{b_1/T_R}} - 1)}{(1 - v_{th})}\right)\right| = b_1\left|\ln\left(\frac{F_1}{(1 - v_{th})}\right)\right| \qquad (9)$$

where the factor $F_1 = \frac{b_1}{T_R}(e^{\frac{1}{b_1/T_R}} - 1)$ can vary between ∞ and 0. With such a large variation in F_1, it is very difficult to fit the threshold delay T_{RD2} against the corresponding SPICE delay.

Model 3. Since the threshold delay computed from the finite ramp response is greater than T_R, an alternative formula for the threshold delay can be obtained by expressing T_{RD3} as $T_{RD3} = T_R + \tau_{RD3}$. Substituting into Equation (5) yields

$$v_{th} = \frac{1}{T_R}\left[T_R + b_1 e^{\frac{-T_R}{b_1}} e^{\frac{-\tau_{RD3}}{b_1}} - b_1 e^{\frac{-\tau_{RD3}}{b_1}}\right] \ .$$

Therefore, the delay is

$$T_{RD3} = T_R + b_1\left|\ln\left(\frac{F_2}{(1 - v_{th})}\right)\right| \qquad (10)$$

The factor $F_2 = \frac{b_1}{T_R}(1 - e^{\frac{-1}{b_1/T_R}})$ varies between 0 and 1.0 as shown in Figure 5. For $b_1 = T_R$ this factor is $F_2 = 0.632$. For $b_1 > T_R$ we can find a good approximation for F_2 by fitting against SPICE-computed delays, since the variation in F_2 values is very small. However, for the range of interconnect configurations studied both Model 2 and Model 3 gave essentially identical results and hence Section 7 reports results from Model 2 only.

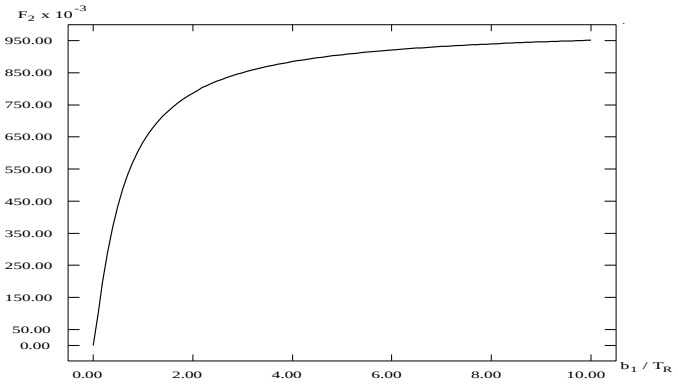

Figure 5: Variation of factor $F_2 = \frac{b_1}{T_R}(1 - e^{\frac{-1}{b_1/T_R}})$ with respect to $\frac{b_1}{T_R}$.

5 Two-Pole Analysis

The two-pole methodology for interconnect response computation under step input has been discussed in [4, 7, 20]. For interconnect trees (or lines) the transfer function has a special form in which the numerator polynomial is a constant, i.e., approximating to s^2 term yields $H(s) \approx \frac{1}{1 + sb_1 + s^2 b_2}$. For the case of resistive source (R_S) and capacitive load (C_L) impedances, the transfer function coefficients are given by [8]

$$\begin{aligned} b_1 &= R_S C + R_S C_L + \frac{RC}{2} + RC_L \\ b_2 &= \frac{R_S RC^2}{6} + \frac{R_S RCC_L}{2} + \frac{(RC)^2}{24} + \frac{R^2 CC_L}{6} + \frac{LC}{2} + LC_L \end{aligned} \qquad (11)$$

For this form of the transfer function, the output response under infinite ramp input is

$$\begin{aligned} U_{out}(s) &= \frac{V_0}{T_R}\frac{1}{s^2}\frac{1}{1 + sb_1 + s^2 b_2} \\ &= \frac{V_0}{T_R}\left[\frac{-b_1}{s} + \frac{1}{s^2} - \frac{1 + b_1 s_2}{s_1 - s_2}\frac{1}{s - s_1} + \frac{1 + b_1 s_1}{s_1 - s_2}\frac{1}{s - s_2}\right] \end{aligned}$$

[4] At 50% threshold, the condition for delay calculation using infinite ramp response is $\frac{b_1}{T_R} \leq 0.625$, with delay calculated using finite ramp response otherwise. Similarly, at 90% threshold, the condition for delay calculation using infinite ramp response is $\frac{b_1}{T_R} \leq 0.1$.

[5] In [9] the function $f(T_{AD})$ is approximated by $f(T_{AD}) = T_{AD}\ln(\frac{1}{1-u_{th}})$, which is threshold delay for the system with analytical delay as the time constant. The delay estimates using this approximation are reasonably close to SPICE-computed delays.

and the corresponding time-domain response is

$$u_{out}(t) = \frac{V_0}{T_R}\left[-b_1 + t + \frac{1+b_1s_2}{s_2-s_1}e^{s_1t} + \frac{1+b_1s_1}{s_1-s_2}e^{s_2t}\right]U(t) \qquad (12)$$

where $U(t)$ is the unit step function. The time-domain response for a finite ramp is

$$\begin{aligned} v_{out}(t) &= u_{out}(t) - u_{out}(t-T_R) \\ &= \frac{V_0}{T_R}\left[T_R + \frac{(1+b_1s_2)(e^{s_1t} - e^{s_1(t-T_R)})}{(s_2-s_1)} \right. \\ &\quad \left. + \frac{(1+b_1s_1)(e^{s_2t} - e^{s_2(t-T_R)})}{(s_1-s_2)}\right]U(t) \qquad (13) \end{aligned}$$

Note that the first and second moments of the transfer function can be obtained from the coefficients b_1 and b_2, i.e., $M_1 = b_1$ and $M_2 = b_1^2 - b_2$. We use the coefficient notation b_1, b_2 and the moment notation M_1, M_2 interchangeably according to the simplicity of the expression.

5.1 Threshold Delay Models

Depending on the sign of $b_1^2 - 4b_2$, the poles of the transfer function can be either real or complex. However, for most cases of interest the poles turn out to be real, and we now discuss delay models for the case of real poles. The condition for the poles to be real is $(b_1^2 - 4b_2) = (4M_2 - 3M_1^2) \geq 0$. Since the magnitude $|s_2|$ is greater than $|s_1|$, the second term in the time-domain response decreases rapidly compared to the first term. Hence, the two-pole infinite ramp response can be approximated as

$$u_{out}(t) \approx \frac{V_0}{T_R}\left[-b_1 + t + \frac{1+b_1s_2}{s_2-s_1}e^{s_1t}\right] \qquad (14)$$

and the finite ramp response as

$$v_{out}(t) \approx \frac{V_0}{T_R}\left[T_R + \frac{1+b_1s_2}{s_2-s_1}\left(e^{s_1t} - e^{s_1(t-T_R)}\right)\right]. \qquad (15)$$

Note that the residue $k_1 = \frac{1+b_1s_2}{s_2-s_1}$ is a positive quantity, and that the pole s_1 has to be negative in value for the response to converge.

Threshold Delay for Infinite Ramp Response.

Model 4. The delay T_{RD4} at threshold voltage u_{th} can be obtained as

$$T_{RD4} + \frac{1+b_1s_2}{s_2-s_1}e^{s_1T_{RD4}} = u_{th}T_R + b_1$$

Again, we can solve such a recursive equation in less than 10 iterations of simple back-substitution (with T_{AD} as the starting value) for all the interconnect configurations we considered. Another way to evaluate the above iterative equation is by substituting some $f(T_{AD})$ for T_{RD4} in the exponential term, which yields

$$T_{RD4} = u_{th}T_R + b_1 - \frac{1+b_1s_2}{s_2-s_1}e^{s_1f(T_{AD})} \qquad (16)$$

where $f(T_{AD})$ depends on the threshold voltage and T_{AD}. For example, for 50% threshold voltage $f(T_{AD}) = T_{AD}$ and for 90% threshold voltage $f(T_{AD}) = 2.3T_{AD}$. We found that the delay values using Equation (16) are close to the values obtained by solving the equation through iteration. Similar to the analysis of Model 1, a better approximation for the $f(T_{AD})$ term can be obtained by expressing T_{RD4} as $u_{th}T_R + \tau_{RD4}$, since T_{RD4} is greater than the threshold delay for the infinite ramp input ($u_{th}T_R$).

Threshold Delay for Finite Ramp Response.

Model 5. The delay T_{RD5} at threshold voltage v_{th} can be obtained from the response as

$$v_{th}T_R = T_R - \frac{1+b_1s_2}{s_2-s_1}(e^{-s_1T_R} - 1)e^{s_1T_{RD5}}.$$

Since the value of the pole s_1 is negative, the quantity $(e^{-s_1T_R} - 1)$ is positive and the residue $\frac{1+b_1s_2}{s_2-s_1}$ is also positive. Thus, the delay expression reduces to

$$T_{RD5} = \frac{1}{|s_1|}\left|\ln\left(\frac{F_3}{(1-v_{th})}\right)\right| \qquad (17)$$

where the factor $F_3 = \frac{(1+b_1s_2)(e^{|s_1|T_R}-1)}{(s_2-s_1)T_R}$ can vary widely.

Model 6. Since the threshold delay computed from the finite ramp response is greater than T_R, an alternative formula for the threshold delay can be obtained by assuming the form $T_{RD6} = T_R + \tau_{RD6}$. Substituting into Equation (15) yields

$$v_{th}T_R = T_R - \frac{1+b_1s_2}{s_2-s_1}(1 - e^{s_1T_R})e^{s_1\tau_{RD6}}.$$

Therefore, the delay is

$$T_{RD6} = T_R + \frac{1}{|s_1|}\left|\ln\left(\frac{F_4}{(1-v_{th})}\right)\right|. \qquad (18)$$

where the factor $F_4 = \frac{(1+b_1s_2)(1-e^{-|s_1|T_R})}{T_R(s_2-s_1)}$ varies over only a small range. For the range of interconnect configurations studied both Model 5 and Model 6 gave essentially identical results, and hence Section 7 reports results from Model 5 only. [9] gives a detailed discussion of two-pole models for the case of complex poles.

6 Interconnection Trees

Finally, we describe how to extend our analytical models to estimate delays in arbitrary interconnect trees. An *RLC* network is called an *RLC tree* if it does not contain a closed path of resistors and inductors, i.e., all resistors and inductors are floating with respect to ground, and all capacitors are connected to ground. Consider an *RLC* interconnect tree with root (or source) S and set of sinks (or leaves) $\{1, 2, \ldots, n\}$. The unique path from root S to the sink node i is denoted by $p(i)$ and is referred to as the *main path*. The edges/nodes not on the main path are referred to as the *off-path* edges/nodes. We model each edge on the main path of the tree using a lumped *RLC* segment, e.g., an **L**, **T**, or π model.[6]

We approximate the off-path subtree rooted at node i with its admittance. At any node i, the admittance Y_i is equal to (i) the capacitance of node i (C_i) if there is no subtree at node i, or (ii) to the sum of the capacitance of node i (C_i) and the subtree admittance $Y_{T(i)}$ otherwise. In other words,

$$\begin{aligned} Y_i &= sC_i && \text{if node } i \text{ has no off-path subtree} \\ &= sC_i + Y_{T(i)} && \text{if node } i \text{ has an off-path subtree} \end{aligned}$$

With this approximation, the main path reduces to an *RLY* equivalent circuit. Only two admittance moments need to be computed for an exact transfer function moment computation for the main path. The k^{th}

[6] Our model is not limited to traditional segment models, and accuracy of our results would likely improve if we use non-uniform segment models [7, 19] designed to perfectly match the low-order moments of the distributed *RLC* line.

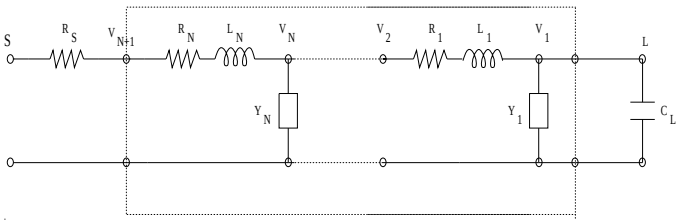

Figure 6: Representation of the main path in the tree, where each distributed line is modeled using *RLC* segments. Y_i indicates the off-path subtree admittance at node *i*.

coefficient b_k of the transfer function for the general *RLY* circuit of Figure 6 can be obtained using the recursive equation given in [7]. The first and second coefficients of the transfer function are

$$\begin{aligned} b_1^{SL} &= R_S \sum_{j=1}^{N} Y_{1,j} + R_N \sum_{j=1}^{N} Y_{1,j} + b_1^N \\ b_2^{SL} &= R_S \sum_{j=1}^{N} Y_{1,j} \cdot b_1^j + R_S \sum_{j=1}^{N} Y_{2,j} + R_N \sum_{j=1}^{N} Y_{1,j} \cdot b_1^j \\ &\quad + R_N \sum_{j=1}^{N} Y_{2,j} + L_N \sum_{j=1}^{N} Y_{1,j} + b_2^N \end{aligned} \tag{19}$$

The first and second moments are expressed in terms of coefficients as $M_1 = b_1$ and $M_2 = b_1^2 - b_2$. For any given source-sink pair the coefficients b_1 and b_2 can be computed in linear time by traversing the main path and using Equation (19) to obtain transfer function coefficients.

7 Experimental Results

We evaluate the above models by simulating various *RLC* interconnect lines with different source/load impedances and different input rise times. We consider typical interconnect parameters encountered in single-chip interconnects [8], with the length of the interconnect being 2000 μm. The source resistance is varied between 100 to 1000 Ω and the load capacitance is varied from 0.1 to 1.0 pf. We also consider 100 *ps* and 500 *ps* rise times for the input ramp.

For all our experiments, we compute exact 50% and 90% delays from the response at the load using the SPICE3e simulator. The step input delay is computed using the Elmore delay formula and then multiplying it with the appropriate constant for the given threshold voltage. For example, Elmore delay at 50% threshold voltage is $0.69b_1$ and at 90% threshold voltage is $2.3b_1$. Unlike [5], we find that Elmore delay is not at all close to SPICE-computed 50% threshold delays and, depending on the rise time of the signal and driver resistance, can be either greater or less than SPICE-computed delays (e.g., when the rise time is 500 *ps* the Elmore delay is for most cases less than the SPICE-computed delays). Also, increased rise time of the input signal causes the Elmore delay to deviate further from SPICE-computed delays (see Tables 1 and 3).

For comparison, we also present delay estimates using the analytical ramp delay model T_{AD}. When the rise time of the ramp input is increased from 100 *ps* to 500 *ps* the SPICE delays at 50% threshold are increased by approximately 200 *ps*, which suggests that delay at 50% threshold voltage is proportional to $\frac{T_R}{2}$. This effect of the rise time is well modeled in the analytical ramp delay model T_{AD}. To compute ramp input delays using the single-pole methodology we use either the Model 1 or Model 2, depending on the value of the first moment b_1 and the threshold voltage of interest. Similarly, to compute ramp input delays using the two-pole methodology we use either Model 4 or Model 5, again depending on the value of b_1 and the threshold voltage of interest. (If the delay is computed using the infinite ramp response then we mark those delays in the Table with (*).) Tables 1 and 2 give 50% and 90% delay estimates for ramp input with 100 *ps* rise time. Tables 3 and 4 give 50% and 90% delay estimates for ramp input with 500 *ps* rise time. Over our range of test cases, Elmore delay estimates can be as much as 100% away from the SPICE-computed delays. In contrast, our single-pole delay estimates are within 4% of SPICE delays and the two-pole delay estimates are within 2.3% of SPICE delays.

8 Conclusions

Fast delay estimation methods, as opposed to simulation techniques, are needed for incremental performance-driven layout synthesis. Estimation methods based on Elmore delay for a step input, although efficient, cannot accurately estimate the delay for *RLC* interconnect lines. We have obtained new analytical delay models under ramp input, based on the first and second moments of *RLC* interconnection lines. The resulting delay estimates are significantly more accurate than Elmore delay estimates. We also describe how to extend our delay models to estimate source-sink delays in arbitrary interconnect trees.

REFERENCES

[1] L. N. Dworsky, *Modern Transmission Line Theory and Applications*, Wiley, 1979.

[2] W. C. Elmore, "The Transient Response of Damped Linear Networks with Particular Regard to Wideband Amplifiers", *Journal of Applied Physics* 19, Jan. 1948, pp. 55-63.

[3] E. G. Friedman and J. H. Mulligan, Jr, "Ramp Input Response of RC Tree Networks", *IEEE ASIC Conference*, 1996.

[4] M. A. Horowitz, "Timing Models for MOS Circuits", *PhD Thesis*, Stanford University, Jan. 1984.

[5] R. Gupta et al., "The Elmore Delay as a Bound for RC Trees with Generalized Input Signals", *ACM/IEEE Design Automation Conference*, June 1995, pp. 364-369.

[6] C. C. Huang and L. L. Wu, "Signal Degradation Through Module Pins in VLSI Packaging", *IBM J. Res. and Dev.* 31(4), July 1987, pp. 489-498.

[7] A. B. Kahng and S. Muddu, "Two-pole Analysis of Interconnection Trees", *Proc. IEEE MCMC Conf.*, January 1995, pp. 105-110.

[8] A. B. Kahng and S. Muddu,' "Accurate Analytical Delay Models for VLSI Interconnects", *IEEE Int. Symposium on Circuits and Systems*, May 1996.

[9] A. B. Kahng and S. Muddu, "Analytical Delay Model for VLSI Interconnects Under Ramp Input", *UCLA CS Dept. TR-960015*, April 1996.

[10] T. Lin and C. A. Mead, "Signal Delay in General *RC* Networks", *IEEE Trans. on Computer-Aided Design*, Oct. 1984, pp. 331-349.

[11] S. P. McCormick and J. Allen, " Waveform Moment Methods for Improved Interconnection Analysis", *Proc. 27th ACM/IEEE Design Automation Conf.*, June 1990, pp. 406-412.

[12] L. T. Pillage and R. A. Rohrer, "Asymptotic Waveform Evaluation for Timing Analysis", *IEEE Trans. on CAD* 9, Apr. 1990, pp.352-366.

[13] V. Raghavan, J. E. Bracken and R. A. Rohrer, "AWESpice: A General Tool for the Accurate and Efficient Simulation of Interconnect Problems", *Proc. 29th ACM/IEEE Design Automation Conf.*, June 1992, pp. 87-92.

[14] C. L. Ratzlaff, N. Gopal, and L. T. Pillage, "RICE: Rapid Interconnect Circuit Evaluator", *Proc. 28th ACM/IEEE Design Automation Conf.*, June 1991, pp. 555-560.

[15] J. Rubinstein, P. Penfield and M. A. Horowitz, "Signal Delay in RC Tree Networks", *IEEE Trans. on CAD* 2(3), July 1983, pp. 202-211.

[16] N. Shirali, "Simple Expressions for Interconnect Delay and Input Transition Time", manuscript, 1995.

[17] M. Sriram and S. M. Kang, "Fast Approximation of The Transient Response of Lossy Transmission Line Trees", *Proc. ACM/IEEE Design Automation Conf.*, June 1993, pp. 691-696.

[18] J. Vlach et al., "Group Delay as an Estimate of Delay in Logic", *IEEE Trans. on Computer-Aided Design* 10, July 1991, pp. 949-953.

[19] Q. Yu and E. S. Kuh, "Exact Moment Matching Model of Transmission Lines and Application to Interconnect Delay Estimation", *IEEE Trans. VLSI Systems* 3, June 1995, pp. 311-322.

[20] D. Zhou, S. Su, F. Tsui, D. S. Gao and J. S. Cong,"A Simplified Synthesis of Transmission Lines with A Tree Structure", *Intl. Journal of Analog Integrated Circuits and Signal Processing* 5, Jan. 1994, pp. 19-30.

Interc. para.	Driver Res.	Load Cap.	SPICE	Elmore Delay	Analy. Delay	Single Pole	Two Pole
$\mathbf{r, l, c}$	$\mathbf{R_S}$	$\mathbf{C_T}$		$\mathbf{0.693b_1}$	$\mathbf{T_{AD}}$		
$/\mu m$	Ω	pf	ps	ps	ps	ps	ps
$0.0015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	83	25	87	83*	84*
”	500	0.01	178	126	232	178	178
”	1000	0.01	302	251	413	302	303
”	100	0.1	90	32	96	90*	92*
”	500	0.1	209	157	277	209	209
”	1000	0.1	364	314	503	365	365
”	100	1	150	96	189	149	151
”	500	1	522	471	730	522	522
”	1000	1	989	939	1406	990	990
$0.015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	87	29	92	87*	88*
”	500	0.01	181	129	237	182	182
”	1000	0.01	305	255	418	306	307
”	100	0.1	96	37	103	96*	97*
”	500	0.1	214	162	284	214	215
”	1000	0.1	369	319	510	370	371
”	100	1	172	118	220	171	173
”	500	1	543	493	761	544	545
”	1000	1	1010	961	1437	1012	1013

Table 1: The length of the interconnect line in these experiments is always $h = 2000\,\mu m$. The rise time of the input ramp is 100 ps. For single-pole delay estimates we use Model 1 or 2 and for two-pole estimates we use Model 4 or 5, depending on whether the delay point falls into the infinite ramp response range or the finite ramp response range. The delay estimates refer to 50% threshold voltage. (*) indicates that the delay is computed using the infinite ramp response models.

Interc. para.	Driver Res.	Load Cap.	SPICE	Elmore Delay	Analy. Delay	Single Pole	Two Pole
$\mathbf{r, l, c}$	$\mathbf{R_S}$	$\mathbf{C_T}$		$\mathbf{0.693b_1}$	$\mathbf{T_{AD}}$		
$/\mu m$	Ω	pf	ps	ps	ps	ps	ps
$0.0015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	141	85	87	145	143
”	500	0.01	468	418	232	470	469
”	1000	0.01	882	835	413	886	885
”	100	0.1	161	106	96	165	161
”	500	0.1	572	522	277	574	573
”	1000	0.1	1090	1042	503	1094	1093
”	100	1	366	319	189	372	366
”	500	1	1612	1563	730	1615	1614
”	1000	1	3167	3118	1406	3172	3170
$0.015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	150	96	92	156	151
”	500	0.01	476	430	237	482	479
”	1000	0.01	889	846	418	898	895
”	100	0.1	174	123	103	181	175
”	500	0.1	583	539	284	591	587
”	1000	0.1	1100	1059	510	1111	1107
”	100	1	429	392	220	445	435
”	500	1	1668	1636	761	1688	1682
”	1000	1	3218	3191	1437	3245	3238

Table 2: The length of the interconnect line in these experiments is always $h = 2000\,\mu m$. The rise time of the input ramp is 100 ps. The delay estimates refer to 90% threshold voltage.

Interc. para.	Driver Res.	Load Cap.	SPICE	Elmore Delay	Analy. Delay	Single Pole	Two Pole
$\mathbf{r, l, c}$	$\mathbf{R_S}$	$\mathbf{C_T}$		$\mathbf{0.693b_1}$	$\mathbf{T_{AD}}$		
$/\mu m$	Ω	pf	ps	ps	ps	ps	ps
$0.0015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	287	25	287	287*	287*
”	500	0.01	413	126	432	415*	415*
”	1000	0.01	529	251	613	530	530
”	100	0.1	296	32	296	296*	296*
”	500	0.1	445	157	477	445*	449*
”	1000	0.1	586	314	703	587	587
”	100	1	380	96	389	380*	381*
”	500	1	736	471	930	736	737
”	1000	1	1197	939	1606	1197	1198
$0.015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	291	29	292	292*	292*
”	500	0.01	416	129	437	419*	419*
”	1000	0.01	532	255	618	533	534
”	100	0.1	303	37	303	303*	303*
”	500	0.1	450	162	484	455*	455*
”	1000	0.1	591	319	710	591	592
”	100	1	405	118	420	406*	408*
”	500	1	757	493	961	758	759
”	1000	1	1217	961	1637	1219	1221

Table 3: The length of the interconnect line in these experiments is always $h = 2000\,\mu m$. The rise time of the input ramp is 500 ps. The delay estimates refer to 50% threshold voltage. (*) indicates that the delay is computed using infinite ramp response models.

Interc. para.	Driver Res.	Load Cap.	SPICE	Elmore Delay	Analy. Delay	Single Pole	Two Pole
$\mathbf{r, l, c}$	$\mathbf{R_S}$	$\mathbf{C_T}$		$\mathbf{0.693b_1}$	$\mathbf{T_{AD}}$		
$/\mu m$	Ω	pf	ps	ps	ps	ps	ps
$0.0015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	487	85	287	487*	487*
”	500	0.01	720	418	432	722	721
”	1000	0.01	1110	835	613	1113	1113
”	100	0.1	496	106	296	496*	496*
”	500	0.1	814	522	477	816	816
”	1000	0.1	1312	1042	703	1315	1315
”	100	1	633	319	389	638	634
”	500	1	1826	1563	930	1830	1828
”	1000	1	3374	3118	1606	3379	3378
$0.015\,\Omega$ $0.176\,ff$ $0.246\,ph$	100	0.01	491	96	292	492*	492*
”	500	0.01	727	430	437	732	730
”	1000	0.01	1116	846	618	1124	1122
”	100	0.1	503	123	303	504	504
”	500	0.1	825	539	484	832	833
”	1000	0.1	1322	1059	710	1332	1329
”	100	1	687	392	420	700	692
”	500	1	1882	1636	961	1902	1896
”	1000	1	3425	3191	1637	3452	3446

Table 4: The length of the interconnect line in these experiments is always $h = 2000\,\mu m$. The rise time of the input ramp is 500 ps. The delay estimates refer to 90% threshold voltage.