

Estimation of Power Distribution in VLSI Interconnects

Youngsoo Shin and Takayasu Sakurai

Center for Collaborative Research and Institute of Industrial Science,
University of Tokyo, Tokyo 106-8558, Japan

Abstract

The analysis and simulation of effects induced by interconnects become increasingly important as the scale of process technologies steadily shrinks. While most analyses focus on the timing aspects of interconnects, power consumption is also important. In this paper, the power distribution estimation of interconnects is studied using a reduced-order model. The relation between power consumption and the poles and residues of a transfer function is derived, and an appropriate driver model is developed, allowing power consumption to be computed efficiently. Application of the proposed method to RC networks is demonstrated using a prototype tool.

1 Introduction

As the scale of process technologies steadily shrinks and the size of designs increases, interconnects have increasing impact on the area, delay, and power consumption of circuits. Reduction in scale causes several effects: gate delays decrease due to the thinning gate oxide; interconnect resistances increase due to shrinking wire widths; the aspect ratios of interconnects have to be increased to compensate for increasing interconnect resistance; the lateral and fringing components of capacitance dominate the total capacitance of interconnects; and interconnect capacitance dominates total gate loading. These factors cause a continual increase in interconnect delays, although of course overall circuit performance continues to increase. In fact, interconnect delay is already a significant portion of the clock cycle time for large high-frequency chips [1].

As regards power, the situation is similar in that the portion of power associated with interconnects is increasing. This is an important fact because the conventional design, analysis, and synthesis of VLSI circuits are based on the assumption that gates are the main sources of on-chip power consumption. Furthermore, the power consumed by interconnects results in a phenomenon, called *self heating*, which reduces electromigration-induced mean time to failure (MTF) [2].

To verify the effects induced by interconnects, a combination of extraction and analysis is necessary. Extraction determines the capacitance and the resistance of interconnects, which can then be used to build a circuit model for the analysis of interconnect effects. For analysis (or estimation), extensive studies have been made of

the use of model order reduction over the last few years, following the introduction of Asymptotic Waveform Evaluation [3]. Model order reduction is based on approximating the Laplace-domain transfer function of a linear (or linearized) network by a relatively small number of dominant poles and zeros. Such reduced-order models can be used to predict the time-domain or frequency-domain response of the linear network.

Although there has been significant progress in the analysis and simulation of performance-related aspects of interconnects, less work has been devoted to the analysis of power consumption (or distribution) of interconnects. Furthermore, the analysis of power-related aspects of interconnects is limited to power distribution networks, and deals with quantities such as IR drop, ground bounce, and electromigration.

In this paper, we introduce, for the first time, a method based on a reduced-order model that allows the power distribution of interconnects to be estimated. We show that the power, which inherently involves improper integration, can be derived from the poles and residues of the transfer function, which requires only algebraic computation. When the interconnect is driven by MOSFETs and connected to the gates of MOSFETs, the load transistor can be satisfactorily approximated by a capacitor. And we show that the driver transistor can be modeled by a linear-region resistance with sufficient accuracy for power estimation.

2 Model Order Reduction

Practical circuits contain an extremely large number of poles, especially when circuit components are extracted from the geometry of the layout. Model order reduction is a technique that takes a circuit and reduces it to a smaller representation consisting of the dominant poles from the original circuit. There are two approaches to model order reduction: moment matching and matrix approximation [4]. In this section, we outline the method based on moment matching [3]. However, we stress the fact that *any kind of model order reduction method can be used* as part of the power distribution estimation which we present in the next section.

A lumped, linear, time-invariant circuit can be described by first-order differential equations

$$\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{bu},$$

$$y = \mathbf{c}^T \mathbf{x} + du, \quad (1)$$

where \mathbf{x} is an n -dimensional state vector, \mathbf{A} is an $n \times n$ matrix, u is the system's input, y is the output of interest, and d denotes the direct-coupling term. We wish to obtain the zero-state impulse response of a linear circuit described by (1), which in turn can be used to determine its response to any excitation. We apply the Laplace transform to (1) assuming zero initial conditions and ignoring the term du , which can be treated separately. Then, we obtain

$$\begin{aligned} s\mathbf{X} &= \mathbf{A}\mathbf{X} + \mathbf{b}U, \\ Y &= \mathbf{c}^T \mathbf{X}, \end{aligned} \quad (2)$$

where \mathbf{X} , U , and Y denote the Laplace transform of \mathbf{x} , u , and y , respectively. It follows from (2) that the transfer function, or the Laplace transform of the impulse response, defined as $H(s) = \frac{Y(s)}{U(s)}$, is given by

$$H(s) = \mathbf{c}^T (s\mathbf{I} - \mathbf{A})^{-1} \mathbf{b}, \quad (3)$$

where \mathbf{I} is an identity matrix. If $H(s)$ has a Taylor series expansion about $s = 0$ (i.e. Maclaurin series), then it can be described by

$$H(s) = \sum_{i=0}^{\infty} m_i s^i. \quad (4)$$

Substituting (4) into (3) and equating like powers of s , it can be shown that

$$m_i = -\mathbf{c}^T \mathbf{A}^{-i-1} \mathbf{b}, \quad i = 0, 1, \dots \quad (5)$$

Note that m_i can be computed recursively in (5) because $m_{i+1} = \mathbf{A}^{-1} m_i$, $i = 0, 1, \dots$, meaning that only a single LU -factorization is required.

In a reduced-order model, especially one obtained by moment matching, the transfer function is approximated by the reduced-order system of proper rational function of s having q -poles:

$$\hat{H}(s) = \frac{n_{q-1}s^{q-1} + n_{q-2}s^{q-2} + \dots + n_1s + n_0}{s^q + d_{q-1}s^{q-1} + \dots + d_1s + d_0}. \quad (6)$$

Because there are $2q$ unknowns in the reduced-order system, it is forced to correspond to the first $2q$ terms of (4) by using Padé approximation. In other words, $2q$ low-order moments are required to obtain the reduced-order system having q poles, yielding the following equality:

$$\frac{n_{q-1}s^{q-1} + n_{q-2}s^{q-2} + \dots + n_1s + n_0}{s^q + d_{q-1}s^{q-1} + \dots + d_1s + d_0} = m_0 + m_1s + \dots + m_{2q-1}s^{2q-1}. \quad (7)$$

Multiplying both sides of (7) by the denominator of the left-hand side yields a set of equations that can be solved

for $2q$ coefficients. After finding roots of the denominator of the reduced-order model, (6) can be expressed as a partial fraction expansion form given by

$$\hat{H}(s) = \sum_{i=1}^q \frac{r_i}{s - p_i}, \quad (8)$$

where r_i is a residue of $\hat{H}(s)$ at the pole p_i . It is then straightforward to obtain the approximated impulse response $\hat{h}(t)$ from (8).

3 Estimation of Power Distribution

In order to find the power consumption (or energy dissipation)¹ of a particular resistor element in a linear(ized) circuit, we first obtain the reduced-order model of current flowing through the resistor, denoted by $\hat{J}(s)$ (with the corresponding time-domain function $\hat{j}(t)$), using a model order reduction techniques such as the one described in the previous section. The approximate energy dissipated by R_i , denoted by \hat{E}_i , during time period $[\tau_1, \tau_2]$ is then given by

$$\hat{E}_i = R_i \int_{\tau_1}^{\tau_2} \hat{j}^2(t) dt. \quad (9)$$

If we are interested in the total energy dissipated by a specific resistor element during signal transition, we can choose to consider a semi-infinite interval of τ , without loss of generality. We make τ_1 the time origin and τ_2 infinite time. Then $\hat{j}(t)$ will reach a steady state, provided that $\hat{j}(t)$ corresponds to the reduced-order model of an individual transition. This leads us to the improper integral

$$\hat{E}_i = R_i \int_0^{\infty} \hat{j}^2(t) dt. \quad (10)$$

If $\hat{J}(s)$ is obtained in the form of a partial fraction expansion such as the one in (8) and the poles are distinct, then we can readily derive an algebraic equation involving poles and residues by substituting the combination of exponentials into $\hat{j}(t)$ in (10). However, this sort of direct computation from the improper integral cannot be applied if there are poles whose orders are larger than 1, or if $\hat{j}(t)$ is expressed as combination of functions other than exponentials [5]. Thus, in this paper, we resort to algebraic computation in the s -plane instead of improper integration in the time-domain.

First, we derive a general relation between improper integration in the time-domain and algebraic computation in the s -plane, which is expressed by the following theorem.

¹Power consumption and energy dissipation are used interchangeably. More precisely, power consumption in this paper means *average power consumption*, which is equal to energy dissipation divided by the time period of interest.

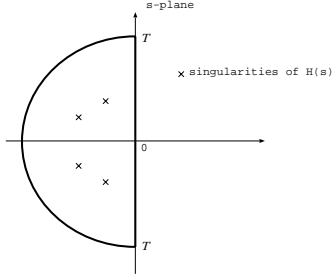


Fig. 1. The singularities of $H(s)$ and the contour of integration.

Theorem 1 *If the Laplace transform of a time-domain signal $h(t)$, denoted by $H(s)$, has q singularities in the left half of the s -plane, then*

$$\int_0^{\infty} h^2(t) dt = \sum_{i=1}^q \tilde{r}_i, \quad (11)$$

where \tilde{r}_i is a residue of $H(-s)H(s)$ at the singularity of $H(s)$.

Proof Let

$$I = \int_0^{\infty} h^2(t) dt. \quad (12)$$

From the definition of the Laplace transform, we have

$$\begin{aligned} I &= \left[\int_0^{\infty} h^2(t) e^{-st} dt \right]_{s=0} \\ &= [\mathcal{L}\{h(t)h(t)\}]_{s=0}. \end{aligned} \quad (13)$$

Since the Laplace transform of a product of two functions is equal to the convolution of the Laplace transforms of two functions, we find that

$$\begin{aligned} I &= \left[\frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\gamma-iT}^{\gamma+iT} H(s-\omega)H(\omega) d\omega \right]_{s=0} \\ &= \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\gamma-iT}^{\gamma+iT} H(-\omega)H(\omega) d\omega, \end{aligned} \quad (14)$$

where γ is chosen solely by the condition that it is to the right of the singularities of $H(s)$, meaning that γ can be chosen as any real number larger than or equal to 0. So we set $\gamma = 0$ and take the contour of integration as a semicircle of radius T with the line $\Re(s) = \gamma$ as diameter and to the left of it and the line segment $\Re(s) = \gamma, -T \leq \Im(s) \leq T$, as shown in Fig. 1. By taking T sufficiently large, we can guarantee that only the singularities of $H(s)$ fall inside the contour, because $H(-s)$ has singularities to the right of the s -plane. Then, by the Cauchy residue theorem, I reduces to the sum of residues of $H(-s)H(s)$ at the singularities of $H(s)$. This concludes the proof. \square

As an example, we consider

$$H(s) = \frac{s+3}{(s+1)^2} = \frac{2}{(s+1)^2} + \frac{1}{s+1}.$$

We form the product of $H(-s)$ and $H(s)$:

$$\begin{aligned} H(-s)H(s) &= \frac{(-s+3)(s+3)}{(s-1)^2(s+1)^2} = \\ &= \frac{2}{(s-1)^2} - \frac{5}{2(s-1)} + \frac{2}{(s+1)^2} + \frac{5}{2(s+1)}. \end{aligned} \quad (15)$$

It can be easily shown that

$$\int_0^{\infty} h^2(t) dt = \int_0^{\infty} (2te^{-t} + e^{-t})^2 dt = \frac{5}{2},$$

which is a residue of $H(-s)H(s)$ at the pole $s = -1$ (the coefficient of $\frac{1}{s+1}$ in (15)). Note that the only constraint required by Theorem 1 is that the transfer function has singularities to the left of the s -plane, which is a typical situation because we are concerned mostly with stable systems. Also, note that (12) corresponds to 0-th time moment of $h^2(t)$ if its Laplace transform has a Maclaurin series expansion. If we have a reduced-order model of $H(s)$, then \tilde{r}_i can be obtained by a matrix computation involving the moments of $\hat{H}(-s)\hat{H}(s)$ and the singularities of $\hat{H}(s)$ [3].

In the case when all the singularities are simple poles, we obtain the less complicated relation expressed by the following theorem.

Theorem 2 *If the Laplace transform of a time-domain signal $h(t)$, denoted by $H(s)$, has q simple poles in the left half of s -plane, then*

$$\int_0^{\infty} h^2(t) dt = \sum_{i=1}^q r_i H(-p_i), \quad (16)$$

where r_i is a residue of $H(s)$ at the pole p_i of $H(s)$.

Proof From Theorem 1, the residue of $H(-s)H(s)$ at the simple pole p_i (\tilde{r}_i) can be computed by

$$\tilde{r}_i = \lim_{s \rightarrow p_i} (s - p_i) H(-s) H(s). \quad (17)$$

Because

$$\lim_{s \rightarrow p_i} (s - p_i) H(s) = r_i, \quad (18)$$

we obtain the desired result from (11), (17), (18):

$$\int_0^{\infty} h^2(t) dt = \sum_{i=1}^q \tilde{r}_i = \sum_{i=1}^q r_i H(-p_i). \quad (19)$$

\square

Notice that the relations derived in Theorems 1 and 2 are exact, rather than approximate. Thus, when the reduced-order model $\hat{H}(s)$ is used in (11) or (16), the accuracy of energy dissipation is determined by the accuracies of the poles and residues of the reduced-order model. The relations can also be used to derive the exact energy dissipation if we have the Laplace transform of the exact time-domain function of current.

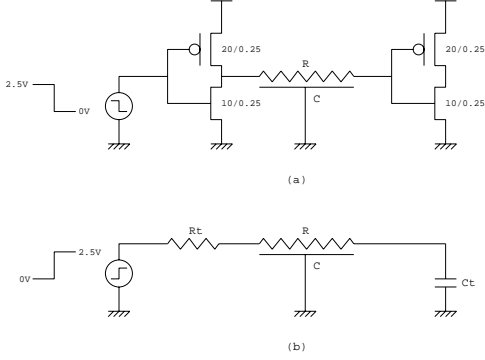


Fig. 2. A model for an interconnect driven by a MOSFET driver and loaded with a MOSFET receiver. The interconnect is $0.5 \mu\text{m}$ wide and 10 mm long in $0.25\text{-}\mu\text{m}$ technology, which results in $R=1.5 \text{ k}\Omega$ and $C=2.0 \text{ pF}$: (a) the circuit and (b) the model of an RC interconnect.

4 A MOSFET Model for Power Distribution Estimation

In order to estimate power distribution based on the method outlined in Sections 2 and 3, we need a simple linear model for non-linear devices such as MOSFETs to reduce the complexity of the estimation. In this section, we discuss the modeling of MOSFETs (when connected to the interconnect) for the purpose of power distribution estimation.

When the interconnect is driven by MOSFETs and connected to the gates of MOSFETs, as shown in Fig. 2(a), the drive transistor can be modeled by an equivalent resistance R_t and the load transistor by a capacitance C_t as shown in Fig. 2(b). It is well-known that the receiver MOSFET can be closely approximated by a capacitor. Approximating the drive transistor by a linear-region resistance R_t is sufficiently accurate for delay estimation [6]. However, the validity of such an approximation is not obvious in the case of power distribution estimation.

Fig. 3 corresponds to an analysis of the circuit in Fig. 2(a), with the interconnect approximated by 10 sections of π -ladder circuits² [6]. When the input of the driver MOSFETs goes from high to low, PMOS drives the interconnect. It is operated in a linear region most of the time, which is evident from the voltage response at the driving point, $V(1)$, shown in Fig. 3. Thus, the current flowing through each resistor, which depends on the voltages of nodes at the both sides of the resistor, is determined by PMOS in the linear region. This indicates that the approximation of R_t by a linear-region resistance is still valid for the estimation of power distribution.

To verify the validity of the resistance approximation when the input is a ramp rather than step, we compare two circuits: one shown in Fig. 3 and another with MOS-

²When $R_t = C_t = 0$, the exact energy distribution along a distributed RC interconnect can be derived, as shown in Appendix A.

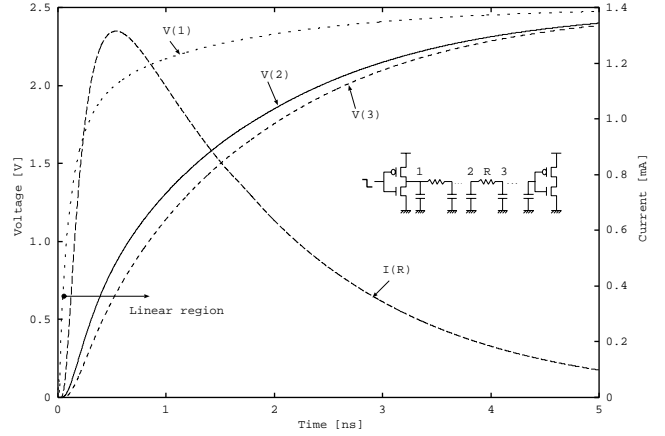


Fig. 3. The step response at the middle of an interconnect together with the driving-point response.

FET driver replaced by a linear-region resistance of PMOS (210Ω). We change the rising time (falling time in case of the first circuit) from 0 ps to 100 ps (10% of cycle time at a one GHz operating frequency), and compare the power consumption at each resistor obtained by SPICE. The average error over all resistor branches is up to 2.1%, while the maximum error is up to 7.9%. This indicates that the simple resistance model is a fairly good approximation for reasonably designed circuits.

5 Experimental Results

We implemented a prototype tool written in C++ and based on the results presented in Sections 2 and 3. The program reads in a circuit in a SPICE-like format and outputs the power distribution of the interconnect. Because the accuracy of power estimation depends on the accuracies of the poles and residues in the reduced-order model, the result presented in this section could be improved using more advanced techniques such as PVL [7].

For the first example, we consider the RC tree shown in Fig. 4, which has widely varying time constants (i.e. it is *stiff*). We compare the energy dissipation of each resistor branch obtained by SPICE with that obtained by our method (approximation by up to 3 poles), when a step voltage is applied. The results are shown in Table 1, where energy dissipation is of the order of pJ. For resistors from R3 to R8, approximation with a single pole is enough to yield an accurate result. To understand this, first note that the area under the current waveform when it is approximated by a single pole is equal to that under the exact waveform³. Because the exact waveform is bell-shaped (except for the

³If the reduced-order model of current consists of a single pole, it can be described by $\hat{J}(s) = \frac{r}{s-p}$. From (7), if we let $\hat{J}(s) = m_0 + m_1 s$ and solving for like powers of s , it can be shown that $p = m_0/m_1$ and $r = -m_0^2/m_1$. Thus, $\int_0^\infty \hat{j}(t) dt = \int_0^\infty r e^{pt} dt = -\frac{r}{p} = m_0$. From the definition of moment, this is equal to $\int_0^\infty j(t) dt$.

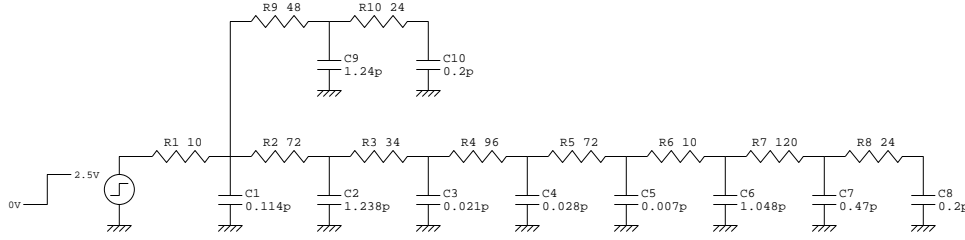


Fig. 4. An RC tree example.

Table 1. Comparison of the energy dissipation computed using SPICE and our method.

Resistor	SPICE	1-pole	2-poles	3-poles
R1	5.118	3.117	4.813	4.956
R2	8.418	7.183	8.394	8.380
R3	0.880	0.886	0.879	0.880
R4	2.424	2.428	2.420	2.424
R5	1.760	1.756	1.756	1.759
R6	0.242	0.242	0.242	0.242
R7	0.429	0.411	0.435	0.429
R8	0.008	0.007	0.008	0.008
R9	5.540	4.694	5.527	5.497
R10	0.051	0.044	0.050	0.051
Avg. error		9.4%	1.2%	0.5%
Max. error		39.1%	5.9%	3.2%

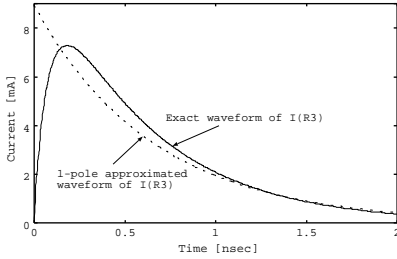


Fig. 5. Exact and approximated current waveforms at R3.

driving end) while the approximated one decays monotonically, the accuracy of energy approximation with a single pole depends on the peakness exhibited by the curve, because we are interested in *the area under the square of the waveform*.

As an example, in the case of waveforms for R3 (shown in Fig. 5), we would expect the squares of both waveforms to be a good match with the area under the curve. In the case of R9, on the other hand, the current waveform peaks near time 0 and has a long tail along the time axis, meaning that it is highly skewed leftwards. Because a single-pole approximation based on moments gives a waveform that follows the gross shape of the original waveform ($s = 0$, thus $t = \infty$), the approximated current waveform of R9 has a large error around its peak (although the area underneath is correct) and this error becomes more significant when we compute the square of the current waveform, as we

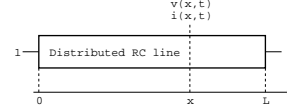


Fig. 7. A distributed RC interconnect.

must. However, approximations with more than one pole give satisfactory results for this example.

The second example consists of randomly generated RC tree networks. We vary the number of nodes from 100 to 500, randomly generate resistance and capacitance values in such a way that the resulting circuit has widely varying time constants (like the first example), and compare the energy distribution obtained by SPICE with that obtained by our method. As an example, Fig. 6 shows the result for circuits with 300 and 500 nodes. Although the approximation with a single pole depends on the stiffness of the circuit, the approximation with two poles gives accurate result for most cases, as can be seen in Fig. 6.

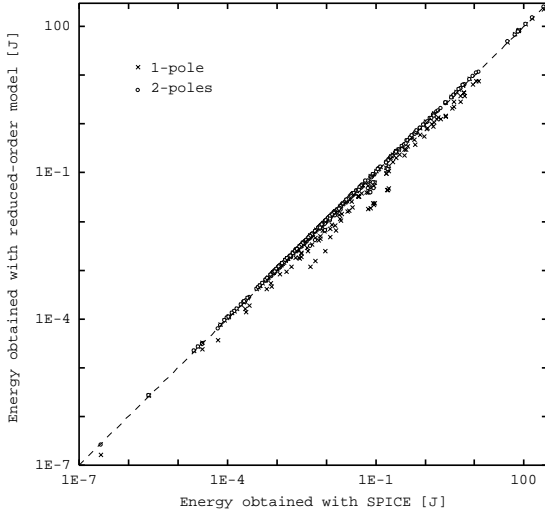
6 Conclusion

We describe a method for the power distribution estimation of an interconnect, based on a reduced-order model. We show that power consumption can be computed efficiently in the s -domain using an algebraic formulation, instead of improper integration in the time-domain. The proposed method of computing power consumption relies on the poles and residues of a transfer function (whether exact or approximate), and can thus be used in any kind of model order reduction technique. We also show that MOSFETs driving an interconnect can be approximated by a linear-region resistance for power distribution estimation.

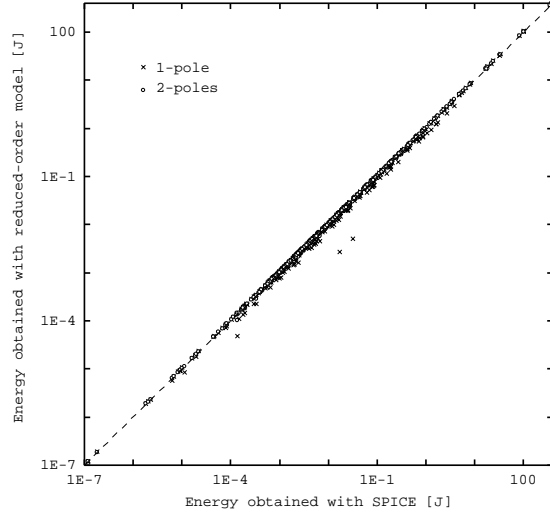
Appendix A

Although the focus of this paper is on the power distribution estimation of circuits consisting of lumped elements, we include the exact energy distribution of a *distributed* RC interconnect for completeness. We consider a distributed RC line as shown in Fig. 7. Suppose that point 1 is excited by a step input. Then, the Laplace transform of $v(x,t)$ is given by [6]

$$V(x,s) = \frac{\cosh\{(1-x')\sqrt{sCR}\}}{s \cosh\sqrt{sCR}}, \quad (20)$$



(a)



(b)

Fig. 6. Comparison of the energy distribution for randomly generated circuits: (a) circuit with 300 nodes and (b) circuit with 500 nodes.

where $x' = x/L$ and R and C are the total resistance and capacitance of the line, respectively. Because

$$i(x,t) = -\frac{1}{r} \frac{\partial v(x,t)}{\partial x}, \quad (21)$$

where r is the resistance of the line per unit length, from (20) and (21) we obtain:

$$I(x,s) = \sqrt{\frac{C}{R}} \frac{\sinh\{(1-x')\sqrt{sCR}\}}{\sqrt{s} \cosh\sqrt{sCR}}. \quad (22)$$

The poles of $I(x,s)$ are given by

$$\begin{aligned} p_k &= -\left(k - \frac{1}{2}\right)^2 \frac{\pi^2}{CR}, \quad k = 1, 2, \dots \\ &= -\frac{\sigma^2}{CR}, \quad \sigma = \frac{\pi}{2}, \frac{3\pi}{2}, \dots \end{aligned} \quad (23)$$

Because all singularities are simple poles in the left half of the s -plane, we can use the relation in Theorem 2. Now, from (22) and (23), we have

$$I(x, -p_k) = C \frac{\sinh\{(1-x')\sigma\}}{\sigma \cosh\sigma}, \quad (24)$$

and for the residue

$$r_k = \lim_{s \rightarrow p_k} (s - p_k) I(x,s) = \frac{2}{R} \cos x' \sigma. \quad (25)$$

Thus, the energy dissipation at an arbitrary position x is given by

$$\begin{aligned} E(x) &= r \sum_{k=1}^{\infty} r_k I(x, -p_k) \\ &= r \sum_{\substack{k=1 \\ \sigma=(k-1/2)\pi}}^{\infty} \frac{2C}{R} \frac{\sinh\{(1-x')\sigma\} \cos\sigma x'}{\sigma \cosh\sigma}, \end{aligned} \quad (26)$$

and it is graphically shown in Fig. 8.

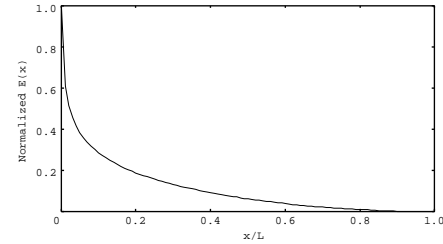


Fig. 8. The distribution of energy dissipation for a distributed RC interconnect.

References

- [1] M. T. Bohr, "Interconnect scaling – the real limiter to high performance ULSI," in *Proc. IEEE Int'l Electron Devices Meeting*, Dec. 1995, pp. 241–244.
- [2] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu, "On thermal effects in deep sub-micron VLSI interconnects," in *Proc. Design Automat. Conf.*, June 1999, pp. 885–891.
- [3] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. on Computer-Aided Design*, vol. 9, no. 4, pp. 352–366, Apr. 1990.
- [4] C. K. Cheng, J. Lillis, S. Lin, and N. Chang, *Interconnect Analysis and Synthesis*, John Wiley & Sons, Inc., 2000.
- [5] R. Kay and L. Pileggi, "PRIMO: Probability interpretation of moments for delay calculation," in *Proc. Design Automat. Conf.*, June 1998, pp. 463–468.
- [6] T. Sakurai, "Approximation of wiring delay in MOSFET LSI," *IEEE Journal of Solid-State Circuits*, vol. SC-18, no. 4, pp. 418–426, Aug. 1983.
- [7] P. Feldman and R. Freund, "Efficient linear circuit analysis by Padé approximation via the Lanczos process," *IEEE Trans. on Computer-Aided Design*, vol. 14, no. 5, pp. 639–649, May 1995.