

© 2011 Dmitri Vyacheslavovich Klokov

APPLICATION OF THE LATENCY INSERTION METHOD (LIM) TO THE ANALYSIS OF
LARGE CIRCUIT INTERCONNECT NETWORKS

BY

DMITRI VYACHESLAVOVICH KLOKOTOV

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2011

Urbana, Illinois

Doctoral Committee:

Professor José Schutt-Ainé, Chair
Professor Jennifer Bernhard
Professor Elyse Rosenbaum
Professor Martin Wong

ABSTRACT

The focus of this research is to explore the applications of the finite difference formulation based on the latency insertion method (LIM) to the analysis of circuit interconnects. Special attention is devoted to addressing the issues that arise in very large networks such as on-chip signal and power distribution networks. We demonstrate that the LIM has the power and flexibility to handle various types of analysis required at different stages of circuit design. The LIM is particularly suitable for simulations of very large scale linear networks and can significantly outperform conventional circuit solvers (such as SPICE).

To my family

ACKNOWLEDGMENTS

First and foremost, I would like to express my appreciation to my adviser, Dr. José Schutt-Ainé, for his attention, guidance, and insight during my thesis research.

I would like to thank the ECE faculty and staff for their help, support, and all the knowledge they shared with me over the years.

I am very grateful to my friends and colleagues Pavle Milosevic and Patrick Goh for being there for me.

Finally, special thanks to my wife Daria for everything, but most of all, for her patience.

TABLE OF CONTENTS

LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER 1 INTRODUCTION	1
1.1 Motivation	1
1.2 Background	1
1.3 Outline	3
CHAPTER 2 LATENCY INSERTION METHOD (LIM)	5
2.1 Basic Formulation	5
2.2 Stability of the Method	9
2.3 Simulation Examples	15
CHAPTER 3 APPLICATION OF THE LIM TO THE ANALYSIS OF POWER DISTRIBUTION NETWORKS	21
3.1 Problem Formulation	21
3.2 Overview of the Methodology	23
3.3 Experimental Results	32
3.4 Complete (DC and Transient) Analysis of PDN	35
CHAPTER 4 APPLICATION OF THE LIM TO CDM ESD MODELING	36
4.1 Problem Formulation	36
4.2 Overview of the Methodology	40
4.3 Experimental Results	43
4.4 Enhanced Model	45
CHAPTER 5 TEMPERATURE-AWARE ANALYSIS USING LIM	46
5.1 Background	46
5.2 Temperature-Dependent IR Drop Analysis	47
5.3 Analysis of Chip Substrate Temperature Profile	51
5.4 Analysis of the Transient Thermal Phenomena	53

CHAPTER 6 APPLICATION OF THE LIM TO THE ELECTRO-THERMAL CIRCUIT ANALYSIS AT THE EARLY DESIGN STAGES	55
6.1 Motivation	55
6.2 Problem Formulation	55
6.3 Pre-Layout Steady State IR Drop Analysis Using LIM	56
6.4 Electro-Thermal Analysis Using the LIM	59
6.5 Analysis of Thermal Transients Using the LIM	72
CHAPTER 7 CONCLUSIONS AND FUTURE DEVELOPMENT	76
7.1 Conclusions	76
7.2 Further Development	77
REFERENCES	80

LIST OF TABLES

3.1	Results for the basic circuit.	30
3.2	Run time results for large circuits.	32
4.1	Run time results.	44
6.1	Run time results.	71

LIST OF FIGURES

2.1	LIM structural elements	6
2.2	LIM branch equivalent circuit	6
2.3	LIM node equivalent circuit	6
2.4	LIM node and branch data structure	8
2.5	Circuit example	12
2.6	Spectral radius of the amplification matrix \mathbf{A}	13
2.7	Stable simulation	14
2.8	Unstable simulation	14
2.9	Spectral radius of the amplification matrix as the function of the time step (left) for the circuit example (right)	15
2.10	Simulation setup for the ideal transmission line	16
2.11	Waveforms at the near and far ends of the ideal transmission line	17
2.12	Equivalent circuit model of a transmission line with frequency-dependent conductor loss taken into account	17
2.13	Simulated waveforms at the near and far ends of the lossy transmission line .	18
2.14	Measured waveforms at the near and far ends of the lossy transmission line (the probe attenuation factor is 10)	18
2.15	Generalized model of a transmission line with frequency-dependent parameters	19
2.16	Simulated waveforms at the near and far ends of the lossy transmission line .	20
3.1	Simplified 3-D view of an on-chip power distribution network with three metal layers	22
3.2	Equivalent circuit model of a PDN network in Fig. 3.1	22
3.3	An average current through a power rail and a dynamic current waveform . .	23
3.4	Circuit model for a steady state power distribution network	24
3.5	A simple circuit example	25
3.6	The random walk “game” representation of the circuit in Fig. 3.5	26
3.7	Augmented section of the example circuit in Fig. 3.5 with fictitious latency elements	27
3.8	Spectral radius of the amplification matrix \mathbf{A} as a function of the time step .	29
3.9	Convergence of the LIM simulation as a function of the time step	30
3.10	Convergence of the LIM simulation	31
3.11	Convergence of the random walk method	31
3.12	Run time as a function of the number of nodes	33

3.13	IR drop profile	33
3.14	Voltage drop profile (left) and corresponding error profile (right)	34
4.1	Field-induced CDM ESD test setup. Chip in the “dead bug” position	39
4.2	Circuit model of CDM test setup	39
4.3	Simple circuit example	40
4.4	LIM-enabled simple circuit schematic	41
4.5	Small circuit ESD current waveform	42
4.6	Unit cell of a lumped resistive substrate model	42
4.7	CDM ESD current waveform	43
4.8	Transient voltage at one of the nodes inside the substrate	44
5.1	Power grid structure and a steady state equivalent circuit model	48
5.2	The flow of thermal-aware IR drop analysis	49
5.3	Multilayer structure of an IC	49
5.4	Analogy between thermal and electrical circuits	50
5.5	A lumped model of the interconnect thermal system	50
5.6	Analogy between the steady state thermal system and the resistive network	52
5.7	Temperature profile of a chip	52
5.8	Thermal-electrical analogous quantities	53
6.1	Model structure for IR drop analysis	57
6.2	Pre-layout IR drop analysis results	58
6.3	Test example setup	61
6.4	Thermal-electrical model structure	62
6.5	Simulation results for the benchmark example	62
6.6	Comparison of the results for the benchmark example	63
6.7	System of two chips with an interposer substrate (cross section)	64
6.8	3-D resistive grid model	65
6.9	1-D model for solder ball grid array	65
6.10	Model for convective boundary	66
6.11	Forced air cooling	66
6.12	Temperature distribution within the structure (cross section)	67
6.13	Temperature distribution within the structure. ANSYS Icepak	68
6.14	Temperature distribution within the structure (cross section) comparison	69
6.15	Top view of the structure	70
6.16	Controller off	70
6.17	Controller on	70
6.18	Two levels of discretization	71
6.19	A segment of an equivalent model for transient thermal simulations	73
6.20	Transient thermal behavior of the system in Fig. 6.7	74
6.21	Nonphysical transient response of the electro-thermal model	75
7.1	Prelayout power integrity analysis flow	77

CHAPTER 1

INTRODUCTION

1.1 Motivation

The efforts of the semiconductor industry to keep up with Moore's law have led to the dramatic increase in the number of devices implemented on a single die. As a result, length, density, and complexity of on-chip interconnects have increased accordingly, leading to multiple signal and power integrity, as well as reliability, issues. Addressing these issues at early design stages requires precise simulation and modeling. Very large size on-chip interconnect networks translate into very large circuit models with millions or even billions of nodes. The sheer size of such models often makes the use of traditional circuit simulation methods inefficient or simply prohibitive. Therefore, currently there is a big demand for powerful computer aided design (CAD) tools based on alternative algorithms capable of processing very large circuit networks consisting of very large numbers of nodes.

1.2 Background

A common approach to modeling interconnect-related effects is based on the representation of a system of interconnects with an equivalent discrete-distributed circuit model. Depending on the type of analysis (steady state or transient), such an equivalent circuit model can be either a purely resistive grid-like structure or an RLC (resistance, inductance, capacitance) or RLCG (resistance, inductance, capacitance, conductance) type of network with constant or time-varying sources. Once an equivalent circuit is obtained, a circuit solver is used to

find voltages and currents in the network. The current industry standard, general purpose simulator is SPICE. There exist multiple commercial CAD tools based on different “flavors” of SPICE (HSPICE, PSPICE, SPECTRE, etc.). SPICE employs the modified nodal analysis (MNA) technique, which solves a system of linear equations that represent the circuit. The MNA method is based on construction and inversion of a conductance matrix which consists of device representations (device “stamps”). When the circuit model size is in the millions of nodes, matrix storage requires large amounts of memory, and matrix inversion becomes prohibitively slow. Average workstations can simply run out of memory and the simulation then takes a very long time. Therefore, it is desirable to use some alternative method (one that would not involve building a large sparse matrix). The alternative method must be as accurate as SPICE, more memory efficient, faster, and, ideally, still must be as general as SPICE; that is, able to address a variety of issues arising in large systems of interconnects. Among other things, it must be able to perform both steady state (DC) and transient analysis of a network.

A number of methods have been proposed to address the size-based complexity of the problem. For example, the hierarchical analysis technique [1] manages the complexity by solving local grids separately, but this can compromise accuracy. Another approach [2] employs a grid-reduction scheme by solving several coarsened grids and then extrapolating the results to the original fine-grain grid. Therefore, the method in [2] solves the network approximately and thus suffers from errors. The statistical random walk method [3] guarantees a linear run time and the availability of memory, but becomes inefficient when voltages at all the nodes have to be computed or when there is a need for high accuracy. The node-based iterative scheme [4] in effect implements the classical successive over-relaxation (SOR) iterative method for solving linear systems. The node-based method is efficient for steady state analysis but not easily extendable to transient analysis. There also exists a number of techniques based on the conjugate gradient (CG) method [5], [6]. These methods demonstrate a lot of promise, but, as with the node-based method in [4], are only applied to DC problems.

This dissertation presents several novel applications of the latency insertion method (LIM). The LIM was initially proposed as a transient simulation technique to model fast transients in large networks efficiently. The method is based on the finite-difference formulation similar to the one used in the well-known finite-difference time-domain method (FDTD) [7]. While FDTD solves time-domain equations for electric and magnetic fields, LIM applies the same finite-difference technique to circuit equations for voltages and currents. LIM is meant to be an alternative to the traditional matrix-vector product-based circuit simulators such as SPICE, and is targeted primarily at simulation of very large networks. LIM iteratively solves circuit equations using a time-stepping scheme; therefore, it avoids construction and computation of a large matrix (as opposed to SPICE). The method demonstrates linear numerical complexity, enables analysis of very large linear networks, and greatly outperforms conventional MNA-based methods.

1.3 Outline

An introduction to the latency insertion method is given in Chapter 2. The algorithm formulation is provided, the stability of the method is explored, and several examples of typical simulations performed using the LIM are considered.

Application of the LIM to the analysis of on-chip power distribution networks (PDNs) is discussed in Chapter 3. It is shown that, while transient simulation of large-scale interconnect networks is a natural problem for the LIM, the method can also be successfully used for the steady state analysis of a PDN. Thus, it is demonstrated that the LIM can be considered as a general purpose simulator capable of performing both steady state and transient simulations.

Another problem that involves analysis of a very large scale network, simulation of the charged device model (CDM) electrostatic discharge (ESD) event, is considered in Chapter 4. The LIM is shown to be a suitable tool for CDM ESD analysis.

Application of the LIM to electro-thermal analysis of circuit interconnects is proposed in Chapter 5. Use of the LIM for temperature-dependent steady state IR drop calculation in an on-chip PDN is also discussed in that chapter.

In Chapter 6 special attention is paid to the analysis of both electrical and thermal phenomena at the very early stages of on-chip and off-chip power delivery system design. It is demonstrated that the LIM can be employed to obtain the steady state temperature profile of a whole system (IC and package) as well as its transient thermal behavior. Several actual commercial designs are analyzed.

Finally, in Chapter 7 some conclusions are reported and possible directions for further development of the LIM are discussed.

CHAPTER 2

LATENCY INSERTION METHOD (LIM)

2.1 Basic Formulation

The latency insertion method (LIM) [6] treats a circuit as a grid composed of nodes interconnected with branches. The approach is based on utilizing an equivalent model for circuit interconnects (distributed RLCG-based transmission line model), as shown in Fig. 2.1. The RLCG equivalent circuit representation of IC interconnects can be obtained by running a layout extraction procedure in a CAD tool used in the design process.

Each branch (a section of an interconnect) can be represented by a distributed series resistor-inductor model as a combination of a voltage source, an inductor, and a resistor in series (see Fig. 2.2). Current I_{ij} is assumed to be directed from node i at potential V_i to node j at potential V_j .

Each node is modeled as a combination of a current source H_i , a conductance G_i , and a capacitor to the ground C_i (as shown in Fig. 2.3).

Using Kirchhoff's voltage law (KVL) for the circuit in Fig. 2.2, we can write the following expression relating node voltages and the current flowing through the branch:

$$V_i - V_j = I_{ij}R_{ij} + L_{ij}\frac{\partial I_{ij}}{\partial t} - E_{ij} \quad (2.1)$$

We use finite-difference approximation to transform the partial derivative in (2.1) into

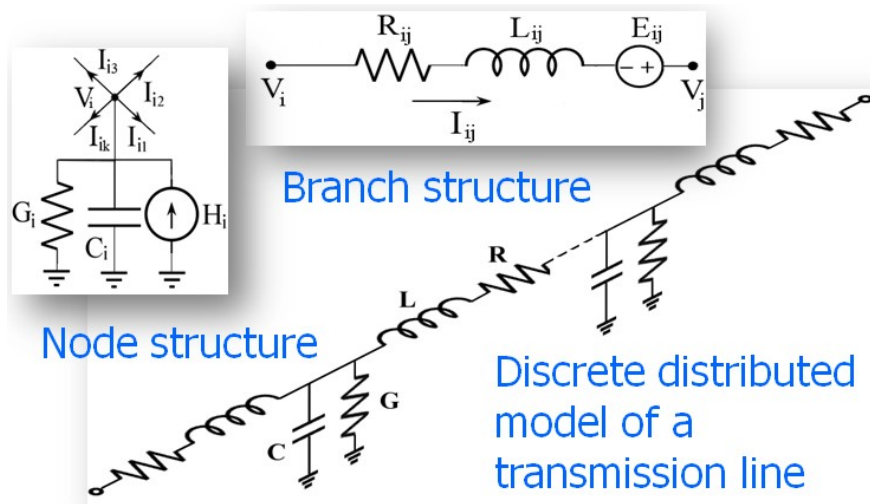


Figure 2.1: LIM structural elements

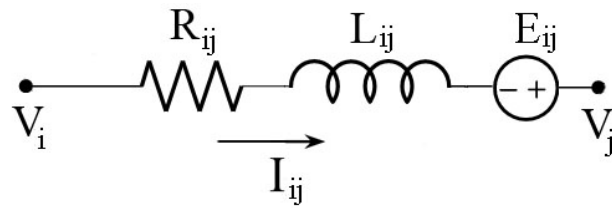


Figure 2.2: LIM branch equivalent circuit

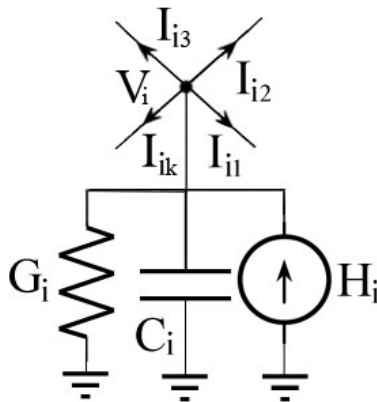


Figure 2.3: LIM node equivalent circuit

$$\frac{\partial I_{ij}}{\partial t} = \frac{I_{ij}^{n+1} - I_{ij}^n}{\Delta t} \quad (2.2)$$

Superscripts in (2.2) represent discrete time. For the $I_{ij}R_{ij}$ term in (2.1) we could choose to use I_{ij}^n or I_{ij}^{n+1} for the backward or forward difference schemes. Instead, we employ the central difference by using the average of the currents at the previous and the next current step. It can be demonstrated that the central difference scheme is second-order accurate, while the other two are only first-order accurate [8]. The discrete time form of (2.1) is then

$$V_i^{n+1/2} - V_j^{n+1/2} = L_{ij} \left(\frac{I_{ij}^{n+1} - I_{ij}^n}{\Delta t} \right) + R_{ij} \left(\frac{I_{ij}^{n+1} + I_{ij}^n}{2} \right) - E_{ij}^{n+1/2} \quad (2.3)$$

Note that currents are computed at whole time intervals $t = n\Delta t$, and voltages at half intervals $t = (n + 1/2)\Delta t$, where Δt is the time step. In this way, voltages and currents are staggered by half a time step. This is known as the leapfrog time stepping scheme [7]. It emphasizes the fact that both voltages and currents are updated at every time step Δt ; but voltages are computed first, and the new voltage values are then used in updating currents.

Solution of the discrete time equation (2.3) for the branch equivalent circuit in Fig. 2.2 yields the following expression for branch currents:

$$I_{ij}^{n+1} = 2\Delta t \frac{(V_i^{n+1/2} - V_j^{n+1/2})}{(2L_{ij} + R\Delta t)} + I_{ij}^n \frac{(2L_{ij} - R\Delta t)}{(2L_{ij} + R\Delta t)} \quad (2.4)$$

In a similar fashion using Kirchhoff's current law (KCL) for the LIM node equivalent circuit in Fig. 2.3, the discrete time equation can be written as

$$C_i \left(\frac{V_i^{n+1/2} - V_i^{n-1/2}}{\Delta t} \right) + G_i \left(\frac{V_i^{n+1/2} + V_i^{n-1/2}}{2} \right) - H_i^n = - \sum_{k=1}^{M_i} I_{ik} \quad (2.5)$$

where M_i is the total number of branches connected to node i . Solution of (2.5) yields the expression for node voltages:

$$V_i^{n+1/2} = V_i^{n-1/2} \frac{(2C_i - G_i \Delta t)}{(2C_i + G_i \Delta t)} - 2\Delta t \frac{\left(\sum_{k=1}^{M_i} I_{ik}^n - H_i^n \right)}{(2C_i + G_i \Delta t)} \quad (2.6)$$

Using (2.6) and (2.4), voltages and currents are updated at every time step throughout the entire network.

It is evident from (2.4) and (2.6) that the presence of latency generated by reactive elements L_{ij} and C_i is required for the algorithm to function. If reactive elements are not present in the circuit, small fictitious inductors must be inserted in each branch and capacitors must be added at every node to enable LIM formulation. It can be shown that if values of those fictitious elements are small enough, the accuracy of the transient simulation is not compromised [6]. The effects of these fictitious elements on the accuracy of the general LIM have been studied in [9], and closed-form expressions for computation of the fictitious elements values have been derived.

As opposed to the FDTD method, which has clearly distinct one-, two-, and three-dimensional models, there is no clear notion of dimension in the LIM. Unlike the FDTD problem, the LIM model can have more than three dimensions: in the LIM model, the dimensions weakly relate to the number of branches connected to a node, which can easily be more than three. In the computer code implementation of the method, the data is always stored in the form of two-dimensional arrays. There are two separate arrays, one containing node parameters and the other containing branch information. Every row in these arrays corresponds to a single branch or a node entry with columns holding parameters of that entry as in Fig. 2.4.

[Branch ID]	[Node1]	[Node2]	[Resistance]	[Inductance]	[Type]
[Node ID]	[Conductance]	[Capacitance]	[Current load]	[Type]	

Figure 2.4: LIM node and branch data structure

2.2 Stability of the Method

As with the traditional FDTD method [10], the LIM formulation is only conditionally stable. In other words, there is an upper bound on the time step that will result in a numerically stable solution to (2.4) and (2.6).

In the case when no more than two branches are connected at each node, the condition for the numerical stability of the method can be written as [11]

$$\Delta t < \min_{i=1}^{N_b} \left(\sqrt{L_i \min(C_i, C_{i+1})} \right) \quad (2.7)$$

where N_b is the number of branches in the network, L_i is the inductance of the branch i , and C_i and C_{i+1} are shunt capacitors connected on either side of the branch i . Condition (2.7) is analogous to the Courant criterion for wave propagation in a discrete grid [12], which is used to determine the limit on the time step in the FDTD method. In the FDTD method, the limit on the time step is based on the fact that numerical waves in a computational space cannot propagate faster than the speed of light. The same idea applies to the LIM; the speed of signal propagation is limited by the latency present in the circuit. However, in the case when there are more than two branches connected at a single node, the problem of determining the stable simulation time step is not as straightforward. Nevertheless, the exact limit on the time step of the LIM simulation can be found using the amplification matrix approach [13], [14]. The amplification matrix method is based on transforming the LIM equations into the matrix form, as we show below.

Equation (2.5) can then be written in a vector-matrix form as

$$\mathbf{C} \left(\frac{\mathbf{v}^{n+1/2} - \mathbf{v}^{n-1/2}}{\Delta t} \right) + \frac{1}{2} \mathbf{G} (\mathbf{v}^{n+1/2} + \mathbf{v}^{n-1/2}) - \mathbf{h}^n = -\mathbf{M}\mathbf{i}^n \quad (2.8)$$

where \mathbf{v} is the node voltage vector of dimension N_n (N_n being the total number of nodes in the circuit); \mathbf{i} is the branch current vector of dimension N_b (N_b is the total number of

branches); \mathbf{C} and \mathbf{G} are diagonal matrices, respectively, of dimensions N_n by N_n , with the values of the capacitors and conductances at each node on the main diagonal; \mathbf{h} is a vector of dimension N_n containing all the current sources at the nodes; and \mathbf{M} is the N_n by N_b incidence matrix defined as follows:

$\mathbf{M}_{qp} = \mathbf{1}$ if branch p is incident at node q and the current flows away from node q .

$\mathbf{M}_{qp} = -\mathbf{1}$ if branch p is incident at node q and the current flows into node q .

$\mathbf{M}_{qp} = \mathbf{0}$ if branch p is not incident at node q .

Solving (2.8) for $\mathbf{v}^{n+1/2}$ yields

$$\mathbf{v}^{n+1/2} = \left(\frac{\mathbf{C}}{\Delta t} + \frac{\mathbf{G}}{2} \right)^{-1} \left[\left(\frac{\mathbf{C}}{\Delta t} - \frac{\mathbf{G}}{2} \right) \mathbf{v}^{n-1/2} + \mathbf{h}^n - \mathbf{M}\mathbf{i}^n \right] \quad (2.9)$$

Similarly, we may rewrite equation (2.3) in vector-matrix form as

$$\mathbf{v}^{n+1/2} = \left(\frac{\mathbf{C}}{\Delta t} + \frac{\mathbf{G}}{2} \right)^{-1} \left[\left(\frac{\mathbf{C}}{\Delta t} - \frac{\mathbf{G}}{2} \right) \mathbf{v}^{n-1/2} + \mathbf{h}^n - \mathbf{M}\mathbf{i}^n \right] \quad (2.10)$$

where \mathbf{L} and \mathbf{R} are diagonal matrices respectively of dimensions N_b by N_b , with the values of the inductances and resistances at each branch on the main diagonal, and \mathbf{e} is a vector of dimension N_b containing all the voltage sources at the branches. Solving (2.10) for \mathbf{i}^{n+1} yields

$$\mathbf{i}^{n+1} = \left(\frac{\mathbf{L}}{\Delta t} + \frac{\mathbf{R}}{2} \right)^{-1} \left[\left(\frac{\mathbf{L}}{\Delta t} - \frac{\mathbf{R}}{2} \right) \mathbf{i}^n + \mathbf{e}^{n+1/2} + \mathbf{M}^T \mathbf{v}^{n+1/2} \right] \quad (2.11)$$

Equations (2.9) and (2.11) can then be used in place of (2.6) and (2.4) as the update equations to calculate the voltage and currents at each time step.

The advantage of the vector-matrix formulation lies in its ability to accurately predict if a time step will be stable. To see this, we return to (2.9) and (2.11) and expand them to get

$$\mathbf{v}^{n+1/2} = \mathbf{P}_+ \mathbf{P}_- \mathbf{v}^{n-1/2} - \mathbf{P}_+ \mathbf{M} \mathbf{i}^n + \mathbf{P}_+ \mathbf{h}^n \quad (2.12)$$

$$\mathbf{i}^{n+1} = \mathbf{Q}_+ \mathbf{Q}_- \mathbf{i}^n + \mathbf{Q}_+ \mathbf{M}^T \mathbf{v}^{n+1/2} + \mathbf{Q}_+ \mathbf{e}^{n+1/2} \quad (2.13)$$

where we have made the definitions

$$\mathbf{P}_+ = \left(\frac{\mathbf{C}}{\Delta t} + \frac{\mathbf{G}}{2} \right)^{-1} \quad \mathbf{P}_- = \left(\frac{\mathbf{C}}{\Delta t} - \frac{\mathbf{G}}{2} \right) \quad (2.14)$$

$$\mathbf{Q}_+ = \left(\frac{\mathbf{L}}{\Delta t} + \frac{\mathbf{R}}{2} \right)^{-1} \quad \mathbf{Q}_- = \left(\frac{\mathbf{L}}{\Delta t} - \frac{\mathbf{R}}{2} \right) \quad (2.15)$$

Substituting (2.12) into (2.13) and rearranging the terms, we obtain

$$\begin{aligned} \mathbf{i}^{n+1} = & \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{P}_- \mathbf{v}^{n-1/2} + (\mathbf{Q}_+ \mathbf{Q}_- - \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{M}) \mathbf{i}^n \\ & + \mathbf{Q}_+ \mathbf{e}^{n+1/2} + \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{h}^n \end{aligned} \quad (2.16)$$

Equations (2.12) and (2.16) can then be grouped together to obtain

$$\begin{aligned} \begin{bmatrix} \mathbf{v}^{n+1/2} \\ \mathbf{i}^{n+1} \end{bmatrix} = & \begin{bmatrix} \mathbf{P}_+ \mathbf{P}_- & -\mathbf{P}_+ \mathbf{M} \\ \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{P}_- & \mathbf{Q}_+ \mathbf{Q}_- - \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{v}^{n-1/2} \\ \mathbf{i}^n \end{bmatrix} \\ & + \begin{bmatrix} 0 & \mathbf{P}_+ \\ \mathbf{Q}_+ & \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \end{bmatrix} \begin{bmatrix} \mathbf{e}^{n+1/2} \\ \mathbf{h}^n \end{bmatrix}. \end{aligned} \quad (2.17)$$

Equation (2.17) defines a discrete linear time invariant system (DLTI) in the form of

$$x(t+1) = \mathbf{A}x(t) + \mathbf{B}u(t) \quad (2.18)$$

The DLTI given in (2.18) is asymptotically stable if and only if all the eigenvalues of \mathbf{A} have magnitude strictly smaller than one. Comparing (2.17) and (2.18), we define the matrix \mathbf{A} as

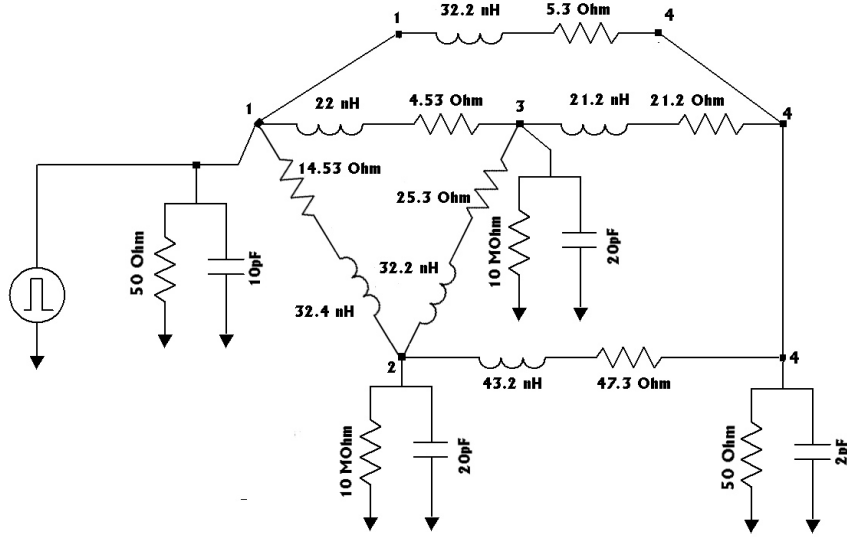


Figure 2.5: Circuit example

$$\mathbf{A} = \begin{bmatrix} \mathbf{P}_+ \mathbf{P}_- & -\mathbf{P}_+ \mathbf{M} \\ \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{P}_- & \mathbf{Q}_+ \mathbf{Q}_- - \mathbf{Q}_+ \mathbf{M}^T \mathbf{P}_+ \mathbf{M} \end{bmatrix} \quad (2.19)$$

and call it the amplification matrix, since in the absence of input, the voltages and the currents in the circuit will be amplified by the matrix \mathbf{A} at each time step. From the above discussion, we see that all the eigenvalues of the amplification matrix defined in (2.19) must have magnitude strictly smaller than one for the simulation to be stable. Thus, we can use the amplification matrix to predict the stability of a time step Δt .

We consider a simple example circuit in Fig. 2.5 to demonstrate the difference between a stable and an unstable simulation. In the example, node 1 is excited with a pulse source and the response is observed at node 4.

We calculate the amplification matrix \mathbf{A} for the circuit in Fig. 2.5 and sweep its eigenvalues to obtain the plot of the spectral radius of the matrix \mathbf{A} as a function of the time step (Fig 2.6).

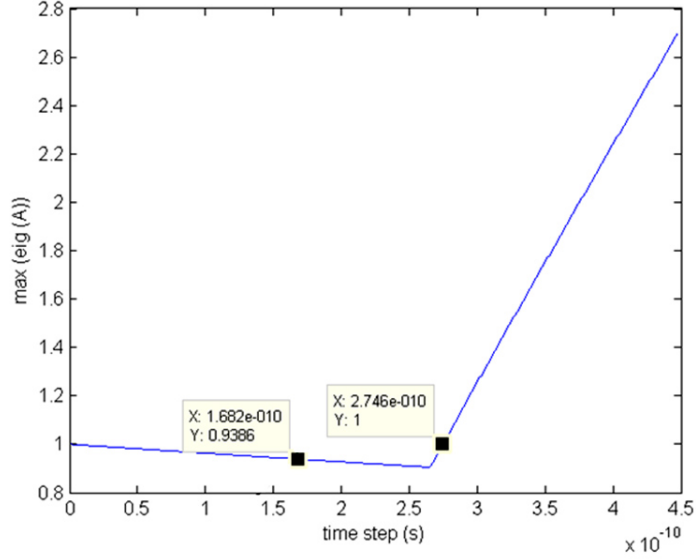


Figure 2.6: Spectral radius of the amplification matrix \mathbf{A}

In Fig. 2.6 we can clearly see the point at which the eigenvalues of \mathbf{A} become larger than unity and, hence, the time step becomes unstable. We pick two values of the time step (one slightly smaller than the limit observed in Fig. 2.6 and another one slightly larger than the limit) and run two simulations for our circuit example. The response in Fig. 2.7 stays stable and diminishes over time as the pulse passes. However, in Fig. 2.8 the response oscillates and the node voltages in the circuit eventually start to increase uncontrollably.

While using (2.19), we can find the exact limit on the time step of the LIM simulation; construction of the amplification matrix and finding its eigenvalues is not always practical. A different approach to finding a stable Δt can be employed. In [9] and [15] the upper bound on the time step of the transient LIM simulation was derived using the direct Lyapunov method (also known as the energy method). In the case when more than two branches are connected to a single node, the stability condition can be written as

$$\Delta t \leq \sqrt{2} \min_{i=1}^{N_n} \left(\sqrt{\frac{C_i}{N_b^i} \min_{p=1}^{N_b^i} (L_{i,p})} \right) \quad (2.20)$$

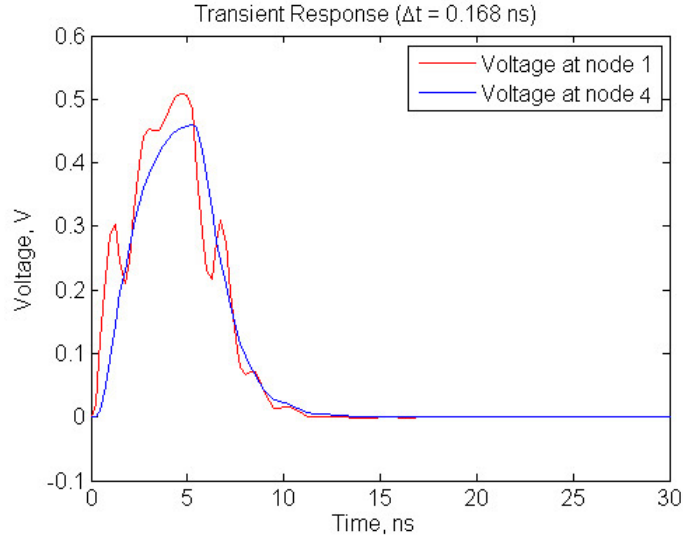


Figure 2.7: Stable simulation

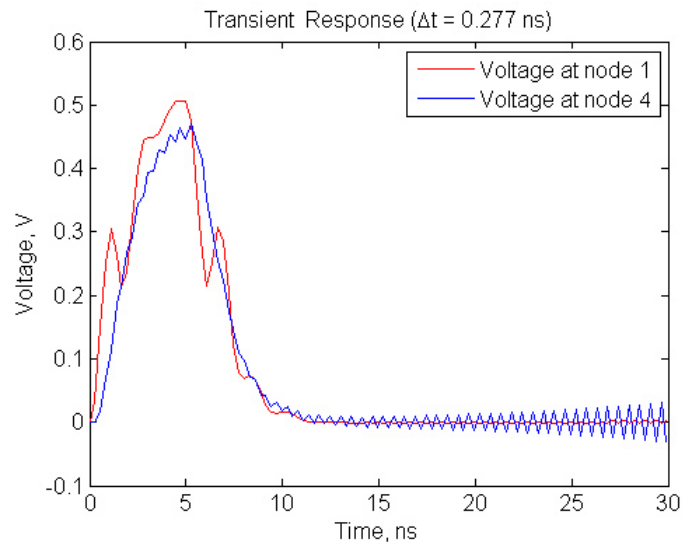


Figure 2.8: Unstable simulation

where N_n is the total number of nodes in the circuit; C_i is the shunt capacitance at node i ; N_{bi} denotes the number of branches connected to node i ; and $L_{i,p}$ denotes the value of p th inductor connected to node i . The result of (2.20) is the square root of the smallest $L_{i,p}C_i$ product among all nodes of the circuit divided by the number of connections at the node. In a one-dimensional case, (2.20) becomes (2.7). The condition in (2.20) has the sufficient

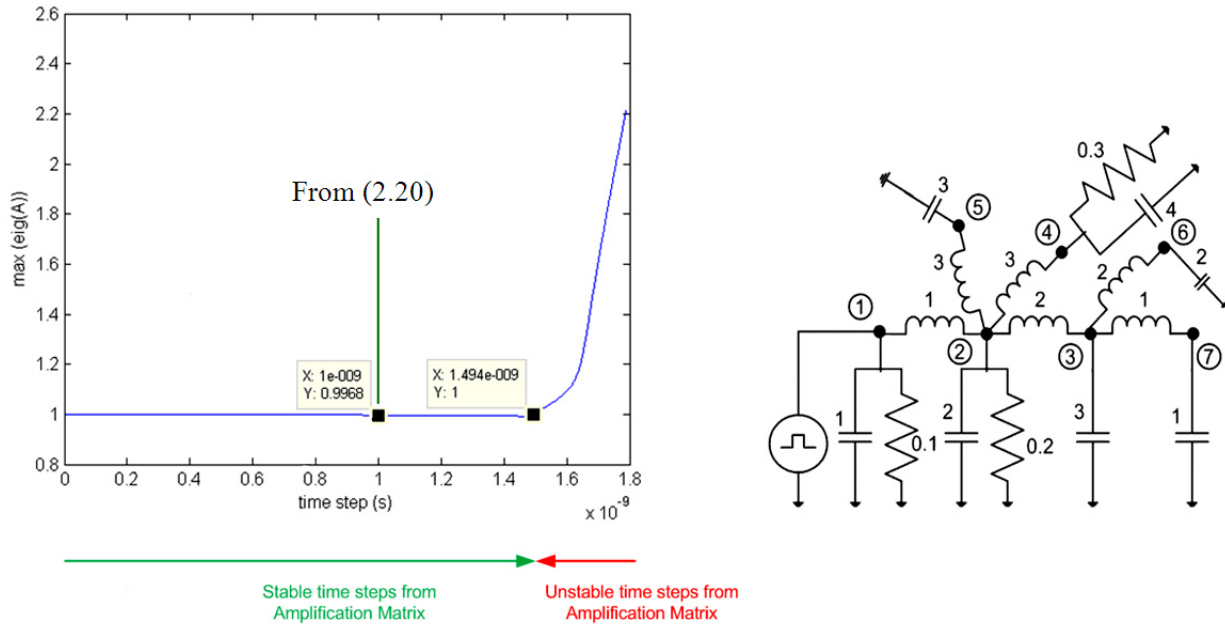


Figure 2.9: Spectral radius of the amplification matrix as the function of the time step (left) for the circuit example (right)

nature. Figure 2.9 shows the comparison between the time step value obtained from (2.20) and the value at which one of the eigenvalues of the amplification matrix in (2.19) becomes unity. A small circuit example is used for calculations (Fig. 2.9 (right)).

The condition in (2.20) is only proven for RLC and GLC circuits (and is the same for both cases). There is no strict proof of (2.20) for the general RLGC configuration. However, intuition and experience show that the same condition is true for the general case.

2.3 Simulation Examples

The LIM node and branch structures are derived from the discrete distributed model for a transmission line. Therefore, the most straightforward application of the LIM method is the transient simulation of cables and board-level interconnects that are well represented by such discrete distributed models.

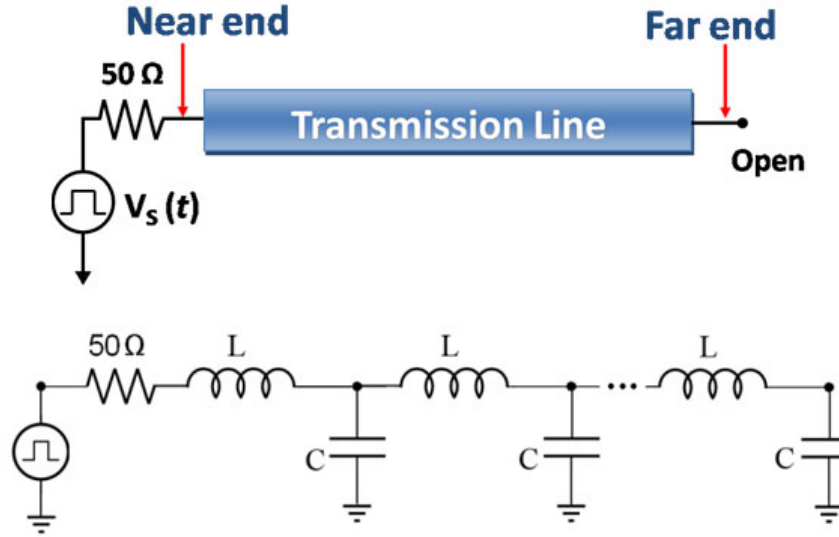


Figure 2.10: Simulation setup for the ideal transmission line

We can consider a simulation of signal propagation in an ideal (uniform, lossless) transmission line as the simplest example. Figure 2.10 shows the simulation setup and the corresponding circuit model.

The line is represented by a number of LC cells connected in series. Assuming that per unit length values of the electrical parameters of the line and its length are known, such an equivalent circuit can be easily constructed. The rule of thumb is to have at least 10 cells per wavelength at the highest frequency. In the simulation, a single pulse is sent down the line. The source pulse magnitude is 4 V, source impedance is 50 Ω . The line has the characteristic impedance of 100 Ω and is open-ended. The waveforms at the near and far ends of the line are observed and shown in Fig. 2.11.

As we can see from Fig. 2.11, in this ideal scenario the pulse travels without distortions and attenuation, and is being properly reflected from impedance discontinuities at the ends of the line.

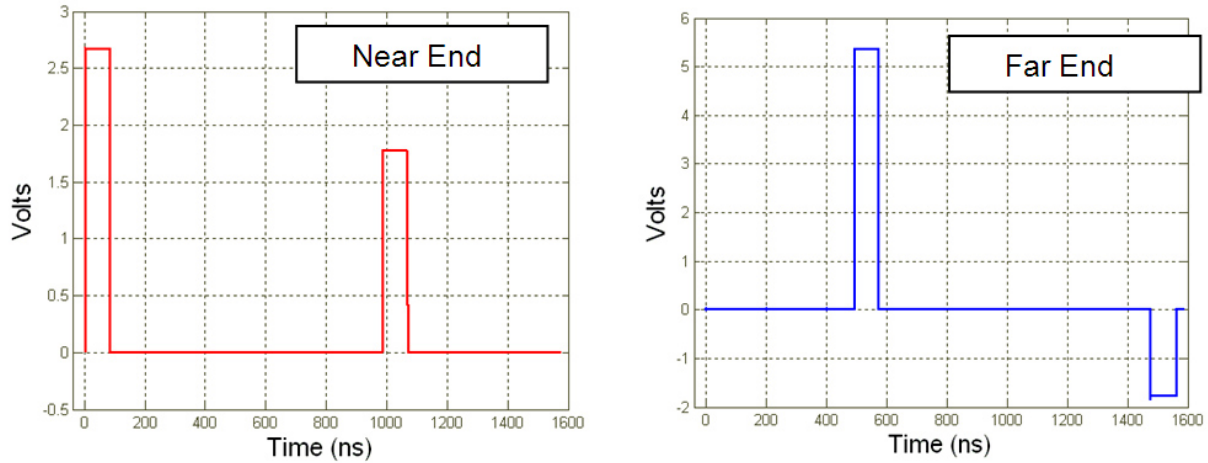


Figure 2.11: Waveforms at the near and far ends of the ideal transmission line

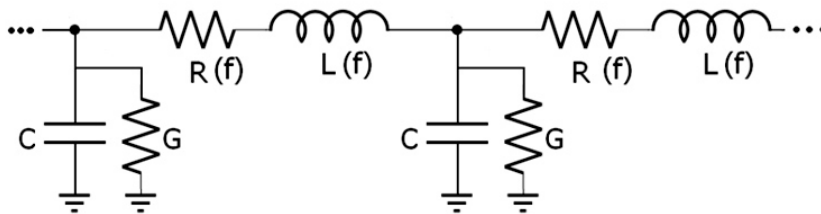


Figure 2.12: Equivalent circuit model of a transmission line with frequency-dependent conductor loss taken into account

Obviously, actual physical interconnects inevitably exhibit some DC resistance. Moreover, at higher frequencies the skin-effect losses, as well as the substrate loss, become major contributors and have to be accounted for in the simulation.

It can be shown [16] that the basic LIM formulation can be adjusted to handle a more realistic simulation of high-frequency signal propagation in a transmission line that exhibits frequency-dependent skin-effect loss. The equivalent model for such simulation is shown in Fig. 2.12.

Figure 2.13 demonstrates the effects of the frequency-dependent loss on the signal. Instead of sharp, undistorted waveforms in Fig. 2.11, in Fig. 2.13 we observe significant signal attenuation and degradation.

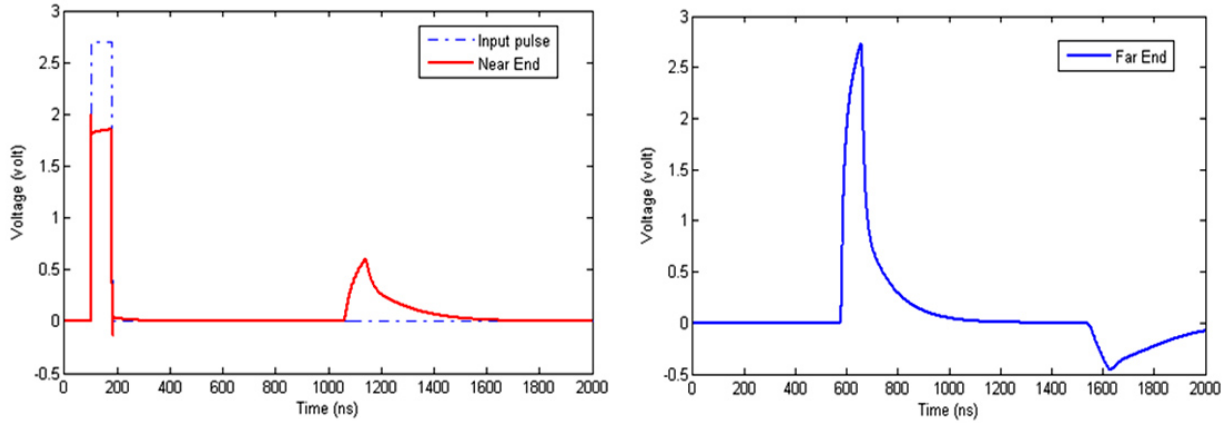


Figure 2.13: Simulated waveforms at the near and far ends of the lossy transmission line

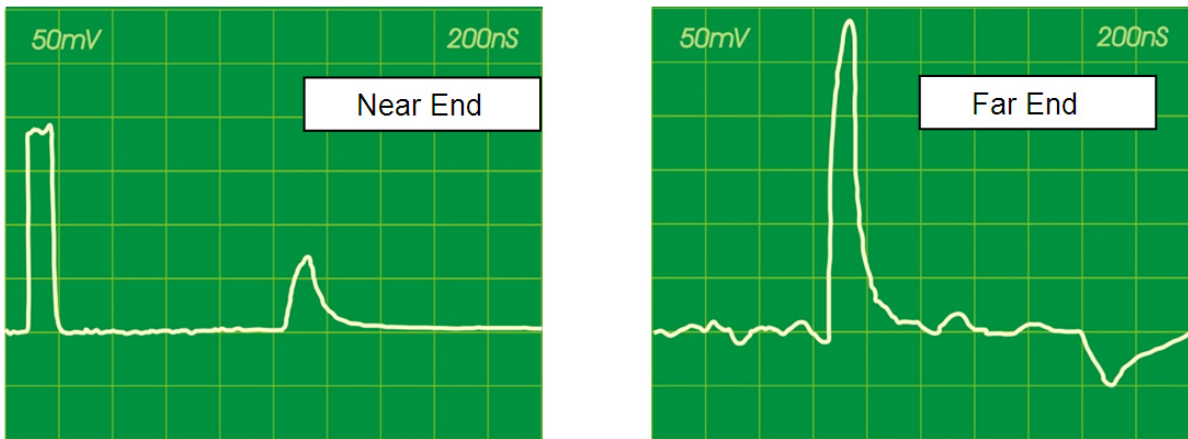


Figure 2.14: Measured waveforms at the near and far ends of the lossy transmission line (the probe attenuation factor is 10)

A time domain reflectometer (TDR) measurement was performed to evaluate the results in Fig. 2.13. The measured waveforms are shown in Fig. 2.14. A very good correlation between the experimental and simulated results can be observed. In [17] the LIM formulation is extended even further to account for the frequency-dependent dielectric loss, which is particularly prominent in PCB environments. The model for the transmission line that takes into account both skin-effect and dielectric loss is shown in Fig. 2.15. The model in Fig. 2.15 consists of generalized frequency-dependent series impedances and shunt admittances.

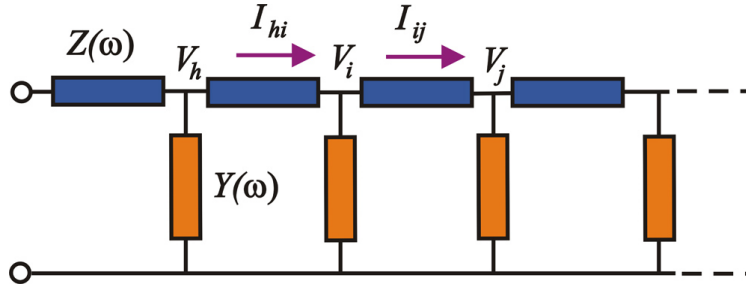


Figure 2.15: Generalized model of a transmission line with frequency-dependent parameters

Frequency-dependent impedances and admittances are represented with low-order rational fraction approximations in the frequency domain. Solution in the time domain is then found using recursive convolution. Coefficients for the rational fraction approximations are found by applying a vector fitting routine to the measured data.

Figure 2.16 shows the results of the simulations. First, a lossless case was modeled; frequency-dependent conductor loss was then added to the simulation, and frequency-dependent substrate loss was finally taken into account. The above examples demonstrate the use of the LIM for modeling signal propagation in stand-alone interconnects. Further in this dissertation we will show that the LIM can be applied to the analysis of various phenomena in very large systems of interconnects.

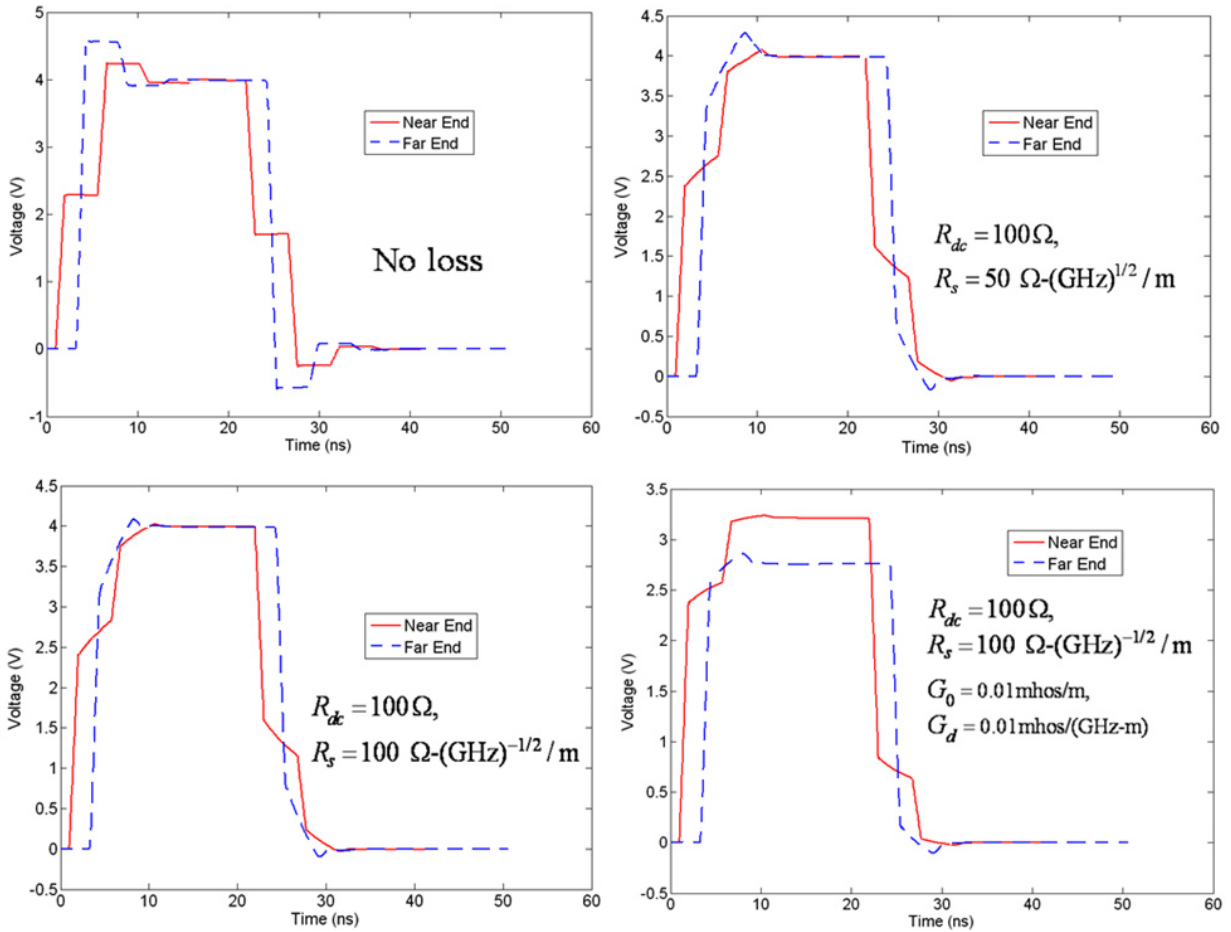


Figure 2.16: Simulated waveforms at the near and far ends of the lossy transmission line

CHAPTER 3

APPLICATION OF THE LIM TO THE ANALYSIS OF POWER DISTRIBUTION NETWORKS

3.1 Problem Formulation

Due to process scaling, the number of devices on a chip has increased dramatically. Consequently, modern on-chip power distribution networks (PDNs) are required to carry very large amounts of supply current. At the same time, power dissipation issues in high-performance high-frequency devices are forcing the use of lower supply voltages [18]. Also, as interconnects become narrower, wire resistances increase. As a result, large amounts of current are distributed through large-scale power grids with narrow interconnects at low voltages, which causes significant voltage drop (IR drop) on the grid, seriously compromising the performance of the chip. Therefore, efficient and accurate analysis of the IR drop in a power grid becomes one of the crucial steps in the design process.

A simplified view of an on-chip power distribution network [9] is shown in Fig. 3.1.

A typical power grid model consists of wire resistances, wire inductances, wire capacitances, decoupling capacitors, VDD pads, and current sources that represent currents drawn by logic gates or functional blocks [3]. The equivalent model of the PDN structure [9] in Fig. 3.1 is shown in Fig. 3.2.

The exact structure of a PDN model depends on the extraction method that was used to obtain the equivalent circuit and underlying assumptions. We are not concerned with the details of the extraction methodology, since one way or the other the resulting model is some form of an RLGC circuit.

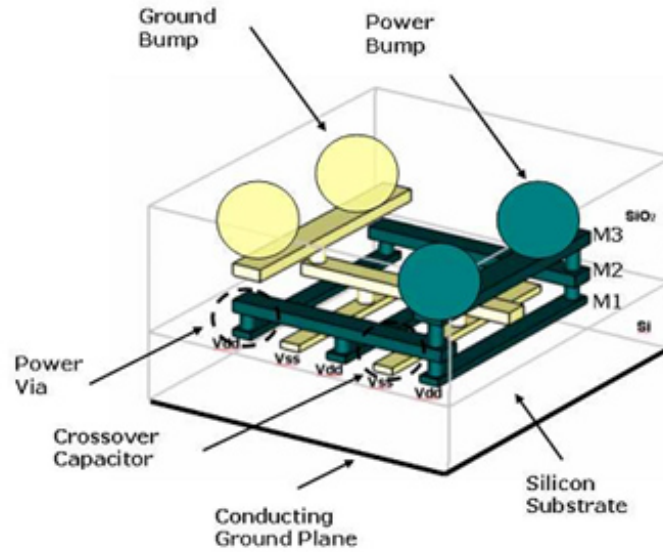


Figure 3.1: Simplified 3-D view of an on-chip power distribution network with three metal layers

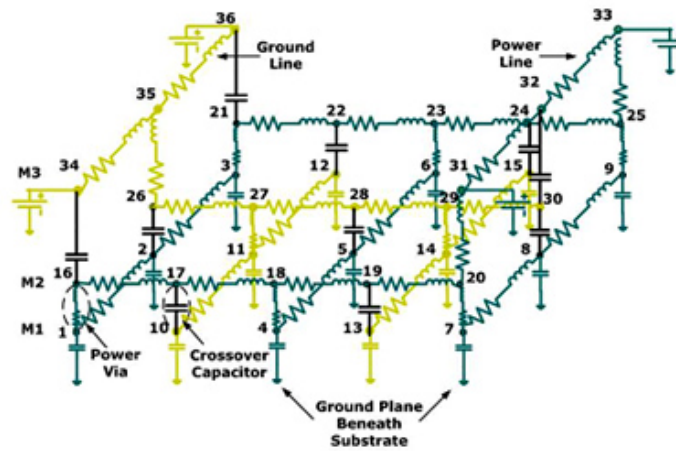


Figure 3.2: Equivalent circuit model of a PDN network in Fig. 3.1

There are two types of PDN analysis: static (steady state) and dynamic (transient). Static power rail analysis evaluates the IR drop caused by high average currents flowing through a design's resistive power rails. This type of power rail analysis has traditionally been used as sign-off analysis at technology nodes above 130 nm, where sufficient natural decoupling capacitance from the power network and non-switching logic tames most dynamic transients.

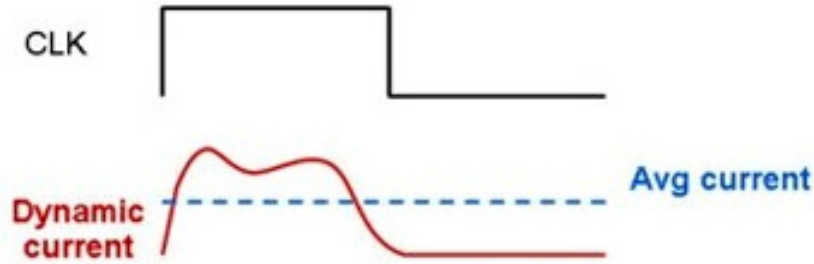


Figure 3.3: An average current through a power rail and a dynamic current waveform

Dynamic analysis evaluates the IR drop caused when large amounts of circuitry switch simultaneously, causing peak current demand on the power rails. This current demand can be highly localized and brief - within a single clock cycle as in (Fig. 3.3) - and can result in an IR drop that causes additional setup- or hold-time violations [19].

Analysis of a PDN is often done in two steps. The steady state solution is found first, and then the transient analysis is performed. The transient simulation is naturally carried out using the standard LIM formulation [6], [9]. In this dissertation, we focus on the first stage: the problem of computing steady state branch currents and node voltages. Due to the sheer size of contemporary PDNs, even the initial steady state analysis becomes a challenging problem and requires the use of efficient algorithms [3], [4], [5]. We show that the LIM, while being in essence a transient simulation technique, can be successfully applied to the steady state analysis problem [20], [21].

3.2 Overview of the Methodology

In case of the steady state (or DC) PDN analysis, all capacitors are open-circuited and inductors are short-circuited. Then, the model in Fig. 3.2 simplifies to the one shown in Fig. 3.4.

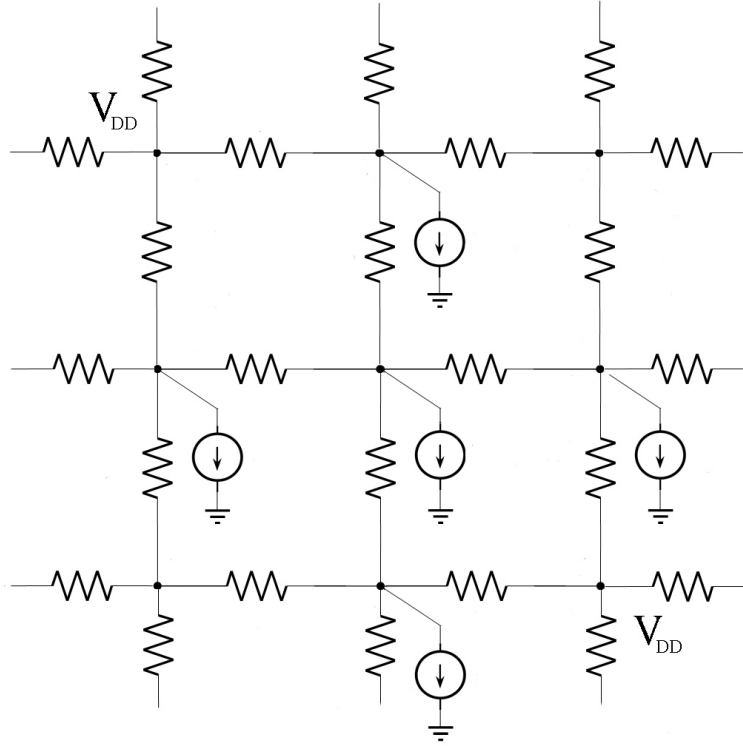


Figure 3.4: Circuit model for a steady state power distribution network

The power grid is then represented as a network of purely resistive branches and sources of two types: constant voltage sources (VDD pads) and constant current sources (currents drawn by the devices and functional blocks).

3.2.1 Simple Circuit Example

In order to demonstrate how the two methods, the random walk and the LIM, can be applied to the steady state analysis of a power grid, we can use a simple circuit example [3]. Figure 3.5 shows a simple circuit that can be viewed as the representation of a small section of a PDN.

If the supply voltage is chosen to be 1 V, the voltages at nodes A, B, C, and D can be easily calculated: 0.6 V, 0.8 V, 0.7 V, and 0.9 V, respectively.

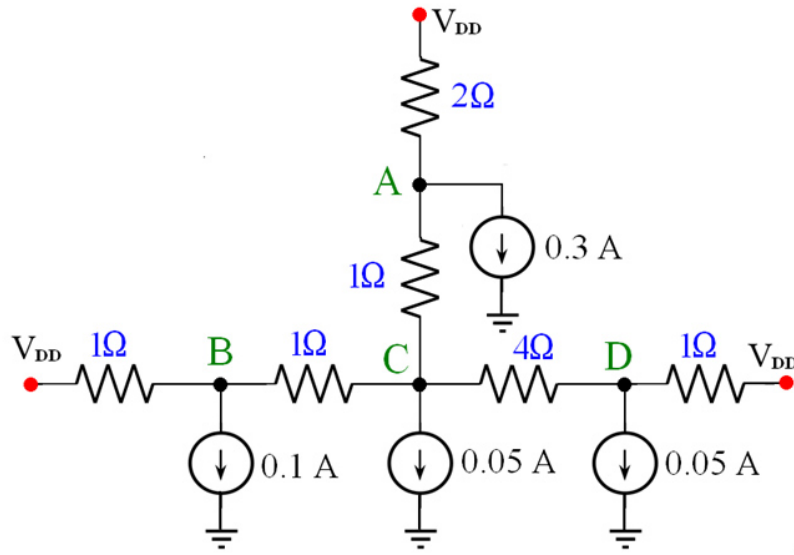


Figure 3.5: A simple circuit example

The random walk method uses a classical statistical approach, a random walk “game,” adopting it to the circuit analysis. The circuit is viewed as a street map. The walker starts from one of the nodes and moves to one of the adjacent nodes with probability p_{ij} , where i is the current node and $j = 1, 2, \dots, N_a$, where N_a is the number of nodes connected to node i . The walk ends when the walker reaches one of the “home” nodes, at which the voltages are known (either VDD or previously computed nodes). Each non-VDD node is assigned a certain “cost” m_i , which the walker “pays” at that node. At the VDD or other home node, the walker is “awarded” a certain amount equal to the voltage at the node. The walker starts with zero amount of “money.” The sum of money left at the end of the walk is taken as an estimate of the voltage at the starting node. As a result, the random walk method transforms the circuit in Fig. 3.5 into the grid shown in Fig. 3.6 with probabilities of moving between nodes calculated as

$$p_{i,j} = g_i / \sum_{k=1}^{N_a} g_k, \quad j = 1, 2, \dots, N_a \quad (3.1)$$

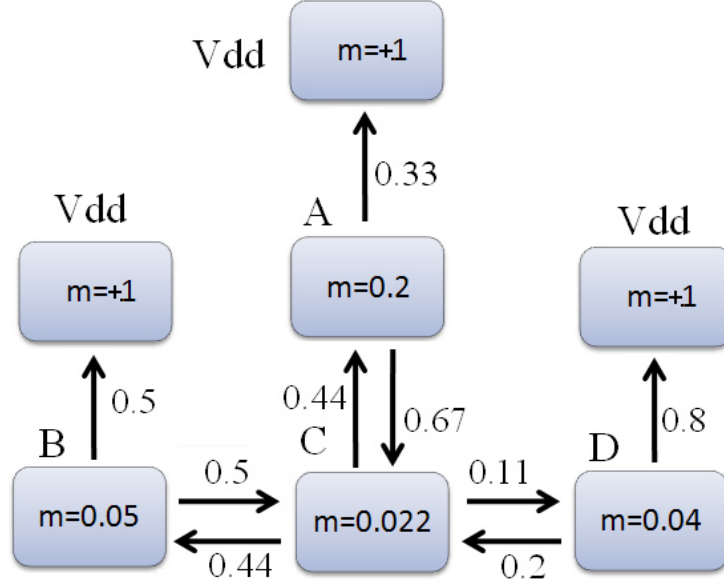


Figure 3.6: The random walk “game” representation of the circuit in Fig. 3.5

where g_i is the conductance of the branch connecting node i and j , and $\sum g_k$ is the sum of conductance of all branches connected to the node i . Costs associated with the nodes are computed as:

$$m_i = I_i / \sum_{k=1}^{N_a} g_k \quad (3.2)$$

where I_i is the current load at node i (I_i have negative values).

The problem of finding steady state voltages at nodes A, B, C, and D is very different from a typical LIM simulation, which normally involves modeling of signal propagation through the network similarly to the examples in Section 2.3. In this steady state case we will apply the LIM iteratively until the steady state solution is reached. We can still view VDD pads and current drains as sources of excitation. That excitation emerges instantaneously, its value stays constant at the source, and the effect of that excitation propagates through the network. However, before we can propagate anything in the LIM simulation, we need to address the issue of latency (more precisely, the lack of it) in the circuit model in Fig. 3.5.

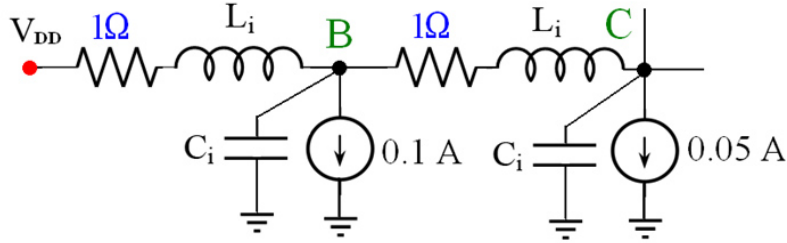


Figure 3.7: Augmented section of the example circuit in Fig. 3.5 with fictitious latency elements

3.2.2 Enabling LIM Formulation: Latency Insertion

The steady state PDN model does not contain latency. In order to enable LIM formulation for the circuit in Fig. 3.5, latency elements must be inserted in all branches and at all nodes. The resulting LIM-enabled augmented circuit corresponding to Fig. 3.5 is shown in Fig. 3.7.

Since the insertion elements are fictitious, they must be made as small as possible, so that the accuracy of the solution is not compromised. In general, the values of fictitious reactive elements are chosen to be much smaller than those of actual inductances and capacitances present in the circuit. On the other hand, the choice of values for fictitious reactive elements affects the stability of the simulation. In a transient LIM simulation, values of reactive elements determine the limit on the maximum possible time step of the simulation. Therefore, it is desirable to have a lot of latency in the circuit, which will allow using the largest time step and, hence, achieving the shortest run time.

However, in the case of the steady state analysis, the actual values of the insertion elements are not important, since only steady state voltages are computed and not transient waveforms. Therefore, values of insertion elements can be chosen arbitrarily. Considerations involved in the choice of the optimum time step are discussed in some detail in Section 3.2.3.

3.2.3 Choosing the Time Step of the LIM Simulation

In the case of the DC analysis, the duration of the simulation is determined by the number of iterations required for the solution to converge to the steady state values. The time step must be chosen so that the simulation remains stable; also, a fast rate of convergence is desirable. Condition (2.20) can be used for the calculation of the time step for the steady state PDN simulation. For the sake of simplicity, all insertion elements throughout the circuit can be assigned to a value of unity. Then, using (2.20), we can write the expression for the time step as

$$\Delta t = \sqrt{\frac{2}{\max_{i=1}^{N_n} (N_b^i)}} \quad (3.3)$$

It can be demonstrated using the circuit in Fig. 3.7 that the time step resulting from (3.3) is stable and closely approximates the optimum step for fastest convergence. Node C in the circuit has the highest number of connections $N_b^C = 3$; then, from (3.3) the time step for the circuit is $\Delta t = 0.8165$.

Another method for assessing stability can be used to verify the result of (3.3). In a method based on the amplification matrix, as described in Section 2.2, the time step of the simulation must be chosen so that the spectral radius of the amplification matrix is less than or equal to unity. Figure 3.8 shows the dependence of the spectral radius of the amplification matrix for the circuit in Fig. 3.7 on the value of the time step. It can be seen in Fig. 3.8 that the system becomes unstable as the value of the time step exceeds $\Delta t = 0.95$. The time step from (3.3) is, therefore, in the stable region.

It is also desirable to use the time step that results in fast convergence of the simulation. The simulation starts with zero initial conditions and then converges to a certain DC voltage level. A larger time step results in larger increments of voltage values at each step of the simulation. Therefore, in general, a larger time step allows the simulation to reach the steady state voltage value in fewer iterations. However, if the time step is too close to the

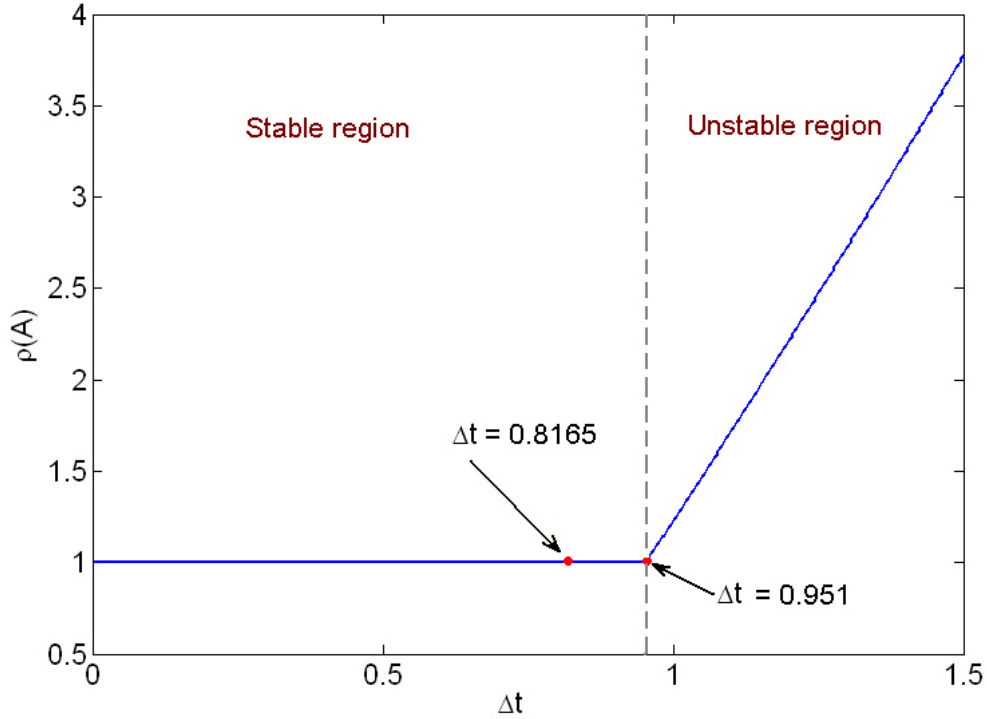


Figure 3.8: Spectral radius of the amplification matrix \mathbf{A} as a function of the time step

stability limit, voltages oscillate around their steady state values, which results in prolonged run times. The trend is shown in Fig. 3.9.

First, as the time step increases, the number of iterations required for convergence decreases. Then, as the value of Δt approaches 0.9, there is a steep increase in the number of steps required for convergence. The value of Δt predicted by (3.3) results in the near-optimum number of 17 iterations required for voltages at all nodes of the circuit example to converge.

3.2.4 Numerical Simulation Results for the Basic Circuit Example

After latency is created and the time step is determined, the augmented circuit can be simulated with LIM. Table 3.1 shows simulation results for the circuit in Fig. 3.5.

It took only 17 iterations to achieve relative error on the order of 0.001%. Figure 3.10 shows the convergence of the simulation for a single node of the basic circuit.

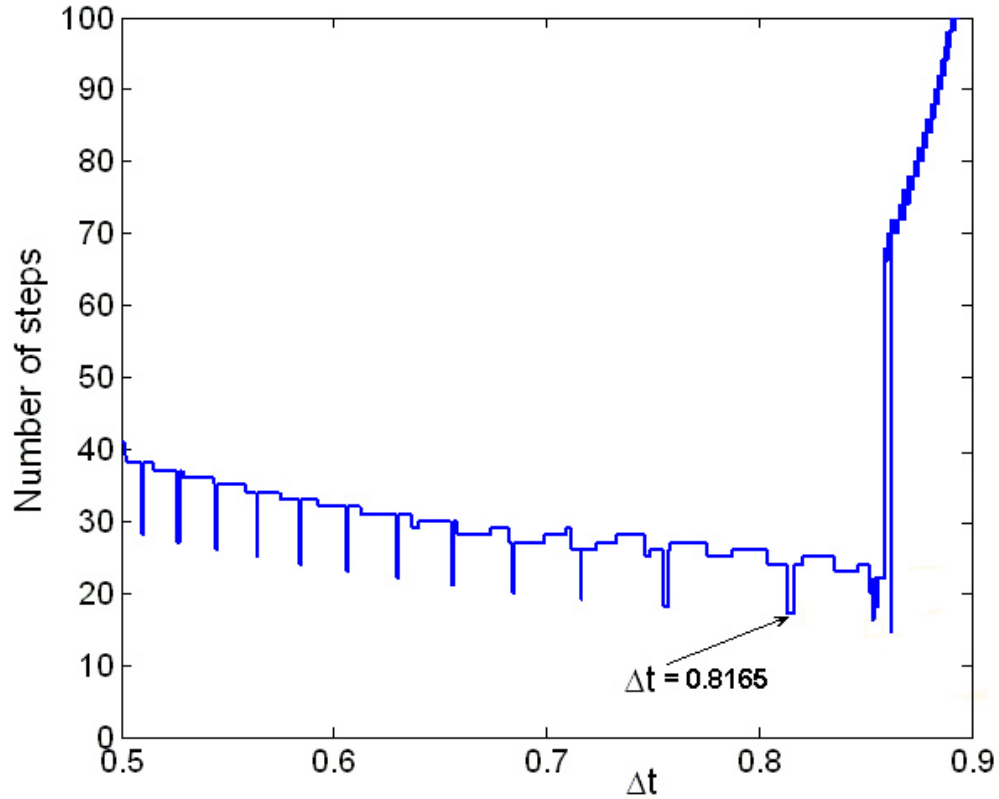


Figure 3.9: Convergence of the LIM simulation as a function of the time step

Table 3.1: Results for the basic circuit.

Node	Estimated voltage (V)	Actual voltage (V)
A	0.600007	0.6
B	0.800009	0.8
C	0.700010	0.7
D	0.900013	0.9

From Fig. 3.10 it can be seen that the LIM simulation quickly reaches the correct steady state value of the node voltage and stays stable at that value. For comparison, the same example was simulated using the random walk algorithm [3]. The solution converges to the correct value as more “walks” are taken from the node, and the results are averaged over a large number of iterations. However, even after a large number of walks the solution can still oscillate considerably. In Fig. 3.11 the convergence of the simulation performed using the random walk algorithm is shown.

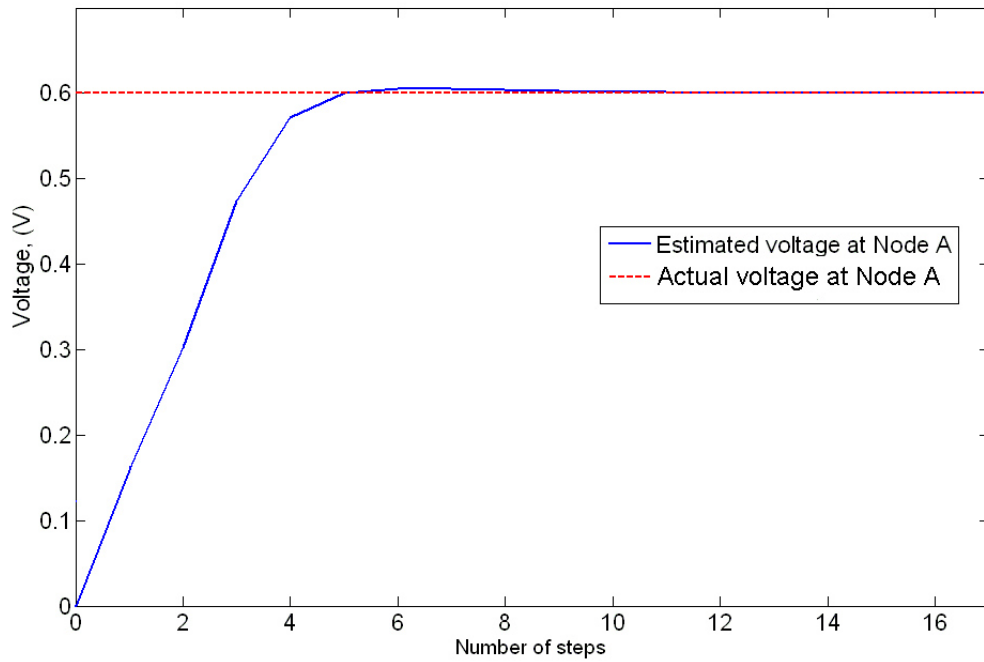


Figure 3.10: Convergence of the LIM simulation

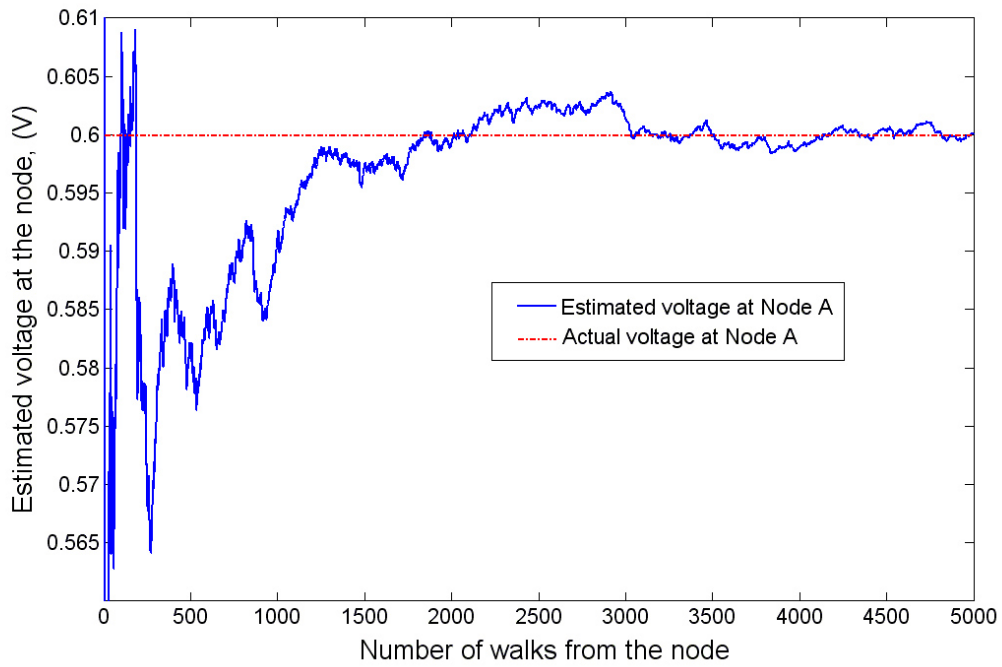


Figure 3.11: Convergence of the random walk method

Table 3.2: Run time results for large circuits.

Number of nodes	LIM (CPU sec)	random walk (CPU sec)
10 K	< 1	10
250 K	3	258
500 K	6	509
1 M	13	1126
2 M	28	2528

Figure 3.11 demonstrates that it is considerably harder to obtain high-precision solutions using the random walk algorithm due to the random nature of the method.

3.3 Experimental Results

As was mentioned above, the main advantage of iterative techniques in general and LIM in particular is the ability to efficiently analyze networks with very large numbers of nodes that are not suitable for matrix-based techniques (such as SPICE). Several large circuits were simulated using LIM and random walk methods. Run time results are summarized and compared in Table 3.2. Run time as a function of the number of nodes in the circuit is also shown in Fig. 3.12 on a semi-log scale.

Table 3.2 can be extrapolated because both algorithms have linear numerical complexity and run times scale nearly linearly. All computations were carried out on a Windows workstation with 2.4 GHz CPU and 2 GB of RAM.

In the above simulations, the error margin was set to 1 mV. It is possible to reduce the run time of the random walk method if high accuracy is not required, or if only a few node voltages need to be computed. Otherwise, the LIM method demonstrates almost two orders of magnitude speedup over the conventional random walk algorithm.

If the information about the mapping of the circuit model network on the physical layout of the IC is preserved, the results of the IR drop analysis can be represented in the form of an IR drop profile (Fig. 3.13).

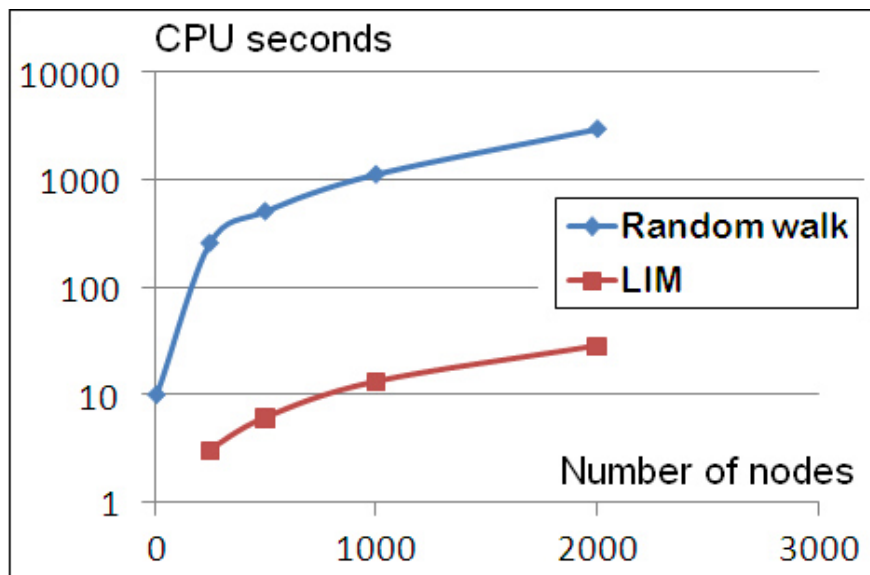


Figure 3.12: Run time as a function of the number of nodes

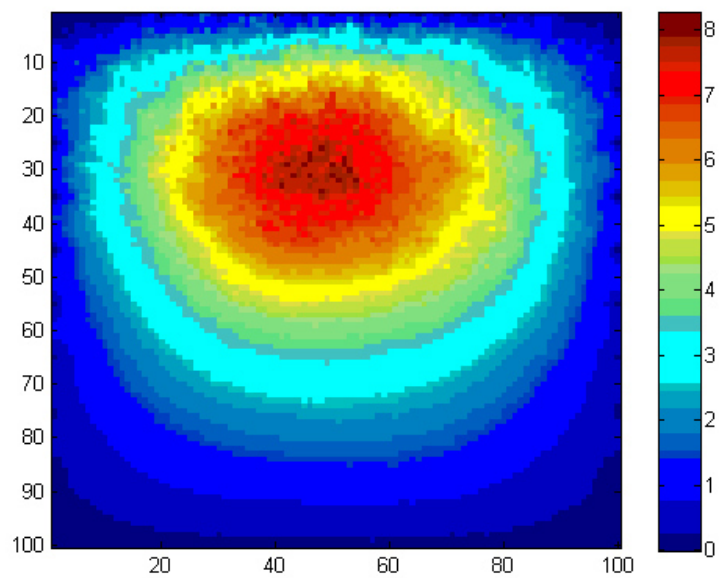


Figure 3.13: IR drop profile

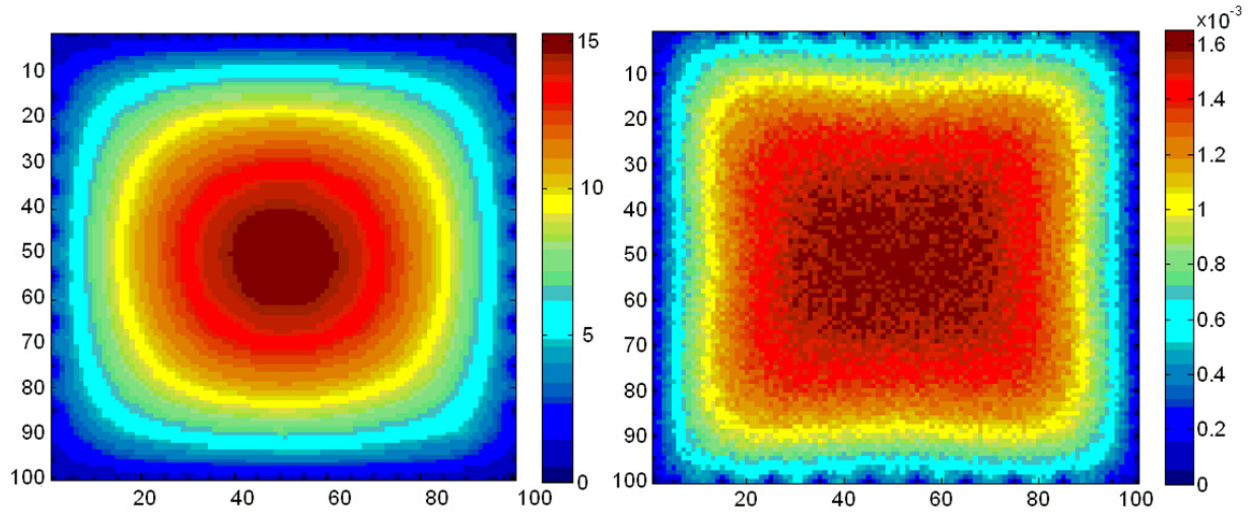


Figure 3.14: Voltage drop profile (left) and corresponding error profile (right)

In the random walk method, variance is used as the measure of convergence of a node voltage. LIM is not a statistical method and does not involve random number generation. In the LIM program, convergence is claimed when the difference between the new voltage value and the old one does not exceed a certain error margin (user specified). Normally the simulation runs until voltage values at all nodes stay within the error margin. A significant speedup can be achieved if the nodes at which convergence is registered are removed from the computation. Such nodes are assigned some special “untouchable” value and are not processed any more. A simulation is stopped once there are no more nodes left to process.

A certain error is introduced every time a node is considered converged and voltage at that node is fixed. However, the error can be kept small. Figure 3.14 demonstrates the error values for the 10 K node circuit (with HSPICE simulation results treated as exact) produced by the LIM simulation. The largest error is still less than 1.7 mV.

The color scale on the voltage drop plot (Fig. 3.14) corresponds to the percentage difference from the supply voltage. The color scale on the error profile plot shows the simple difference between the LIM simulation values and the ones produced by the HSPICE.

With all modifications, the LIM simulation of a circuit with 500 K nodes takes only 2 CPU seconds, while it takes 509 CPU seconds with the random walk method. The circuit has a favorable structure (large number of supply nodes), but it is favorable for both the LIM and the random walk method.

3.4 Complete (DC and Transient) Analysis of PDN

The static analysis of power distribution networks is very important in the early stages of the design process, and must be performed to generate robust power rails. Results of the static analysis are used to determine sufficient metal width, appropriate via sizes, etc. However, static analysis is not sufficient to produce a robust overall design, especially at technology nodes of 130 nm and below. Dynamic (transient) analysis must be performed to account for the effects of simultaneous switching noise (SSN) in order to optimize the insertion of de-coupling capacitance. Also, dynamic power-up analysis has to be used to optimize power switch sizes to control power-up ramp time. The PDN model in Fig. 3.2 is very suitable for the LIM. Therefore, transient simulation of a PDN can be performed by the LIM in a straightforward manner (as shown in [9]). By combining our steady state technique with the transient simulation, a complete analysis of on-chip (as well as off-chip) power distribution networks can be performed.

CHAPTER 4

APPLICATION OF THE LIM TO CDM ESD MODELING

4.1 Problem Formulation

Electrostatic discharge (ESD) has become a major consideration in the design and manufacture of integrated circuits (ICs). ESD impacts production yields, manufacturing costs, quality, and reliability of contemporary semiconductor devices. Problems related to ESD events increase as the trends continue toward higher speed and smaller device sizes. As nanoscale CMOS devices are fabricated with thinner gate oxides, ICs become more susceptible to oxide breakdowns. At the same time, more function blocks are being integrated into a single chip, which increases a die size and leads to larger amounts of static charge stored in the body of the circuit and, consequently, to higher ESD current and voltage spikes. To avoid device failures due to ESD, preventive measures must be taken. Such measures include providing safe paths for the discharge current to flow off-chip by equipping each IC with ESD protection circuitry. Design and placement of such circuitry requires identification of current and voltage distribution during ESD. Therefore, understanding mechanisms of ESD and the ability to test, analyze, and model ESD events are now crucial parts of the IC design process.

A number of different models have been developed to simulate different ESD conditions that cause device failures. These include the human body model (HBM), the machine model (MM), and the charged-device model (CDM). In this work we focus on CDM, which is currently the dominant ESD model because it is particularly effective at simulating damage induced by the automated equipment present in every modern manufacturing facility. Also,

as dielectric failure has become more prominent with scaled semiconductor devices, the CDM accounts for most ESD failure during chip manufacturing [22].

The CDM assumes that the IC package is charged either directly by triboelectric effect (i.e., by frictional contact with some material) or due to induction in the presence of an external electric field. One or more package pins (e.g., leads, solder balls) subsequently contact a conductive surface at or near ground potential. The charge stored on the metal parts of the device flows through in a very fast “spark discharge” and causes failures of junctions, dielectrics, and components along the discharge path [23].

Various hardware test setups have been developed and are used to physically replicate real-world CDM failures. In these tests, the electrostatic charge is induced on a device, and then a discharge path is created, emulating the ESD event. Although successful at stress-testing production devices, physical tests often do not provide enough information about the exact mechanisms of ESD. In particular, the effects of substrate resistivity, as well as failures occurring at internal nodes of a chip, are hard to investigate by test chip design and hardware testing. Circuit simulation presents an attractive time- and cost-effective alternative to physical testing.

Circuit simulations are particularly useful for modeling the failures at internal nodes of a chip, which occur due to ESD currents flowing through the substrate of an IC. In order to capture the effects of substrate conductivity, a suitable model has to be used to represent the substrate material in a simulation. Since on-chip devices are represented with their circuit models, it is desirable to use a circuit-level model for a substrate. Discrete-distributed resistive or resistive-capacitive models can be employed to model a substrate material with finite resistivity. However, such representation may result in very large networks since the number of substrate network nodes is related to the number of substrate taps (i.e., VSS bus connections to the substrate), and the number of substrate taps can be very big. The size of such models makes analysis of the circuit with conventional matrix-based methods inefficient or even prohibitive due to prolonged run times and excessive memory requirements. In this

dissertation we demonstrate that the latency insertion method is a suitable alternative that enables full-scale CDM ESD event simulation.

The LIM iteratively solves circuit equations for voltages and currents using a time-stepping scheme; therefore, it eliminates the need for large matrix equations employed in modified nodal analysis. The method is particularly efficient at modeling very fast transients, such as those occurring during CDM ESD events. The method enables the simulation of substrate models with very high levels of discretization in a reasonable amount of time; it provides access to voltage and current values at any point of a circuit model; and it allows mapping of the current distribution in the circuit with very high precision.

The LIM simulation of CDM ESD is based on the circuit representation of the industry-standard field-induced charged-device model (FICDM) ESD test setup [23]. Although the circuit model emulating the CDM event is uncommon for an LIM simulation, the method was found capable of efficient and accurate modeling of ESD transients.

There are several types of CDM test methods. The one on which we based our simulations is the field-induced method. In the actual test, the chip is placed pins-up (“dead bug” position) on a thin dielectric layer on top of a field-charging electrode. The chip capacitance is charged indirectly to a desired stress voltage by a high voltage source. Next, one of the package pins is touched with a pogo pin, which creates a connection to the top ground plane and triggers the discharge event. The test setup, shown in Fig. 4.1, can be represented with the equivalent circuit model (Fig. 4.2) [24], [25].

The model in Fig. 4.2 represents an IC in a wirebonded, leaded package. The circuit model for on-chip devices includes power busses, ESD protection circuitry, and decoupling capacitors. The substrate can be modeled as a three-dimensional resistive grid. The package pins as well as the pogo pin are represented by series combinations of resistance and inductance. The charge stored on the elements of the chip and the package is contained in capacitances formed between these elements and the two plates of the test setup. Using the model in Fig. 4.2, the CDM ESD event can be successfully simulated using SPICE-like tools. The

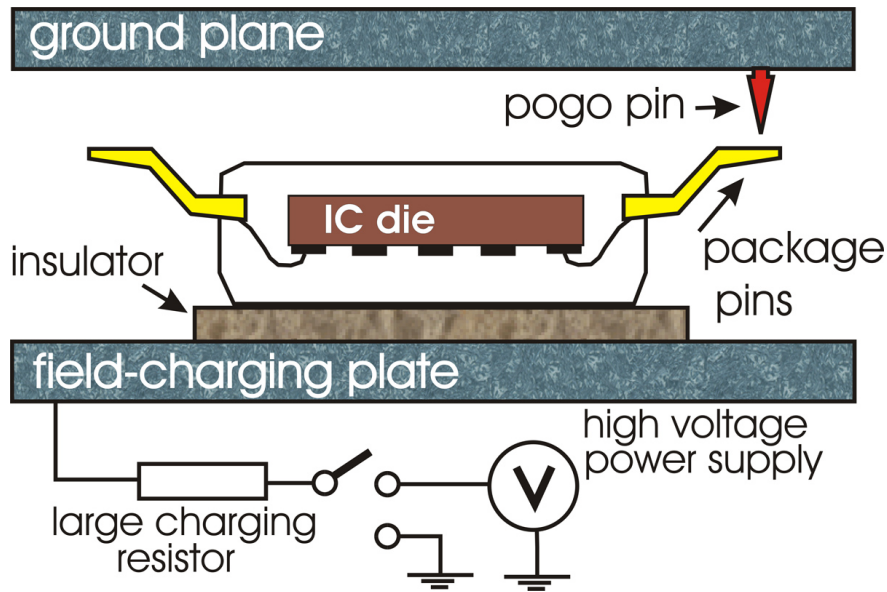


Figure 4.1: Field-induced CDM ESD test setup. Chip in the “dead bug” position

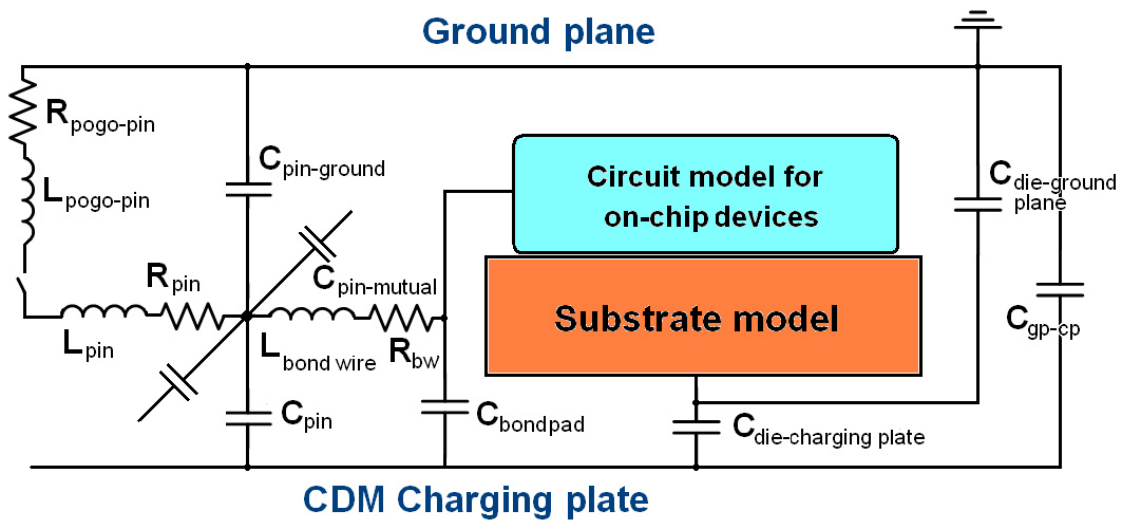


Figure 4.2: Circuit model of CDM test setup

simulation, however, becomes very time-consuming if the high-resolution substrate model is used. Another limiting factor is the size of the on-chip PDN. Using a high-resolution substrate model and large on-chip interconnect network may result in a circuit netlist with more than a million nodes. In such cases, conventional circuit simulators are ineffective, and the use of alternative methods (such as the LIM) is desirable [26].

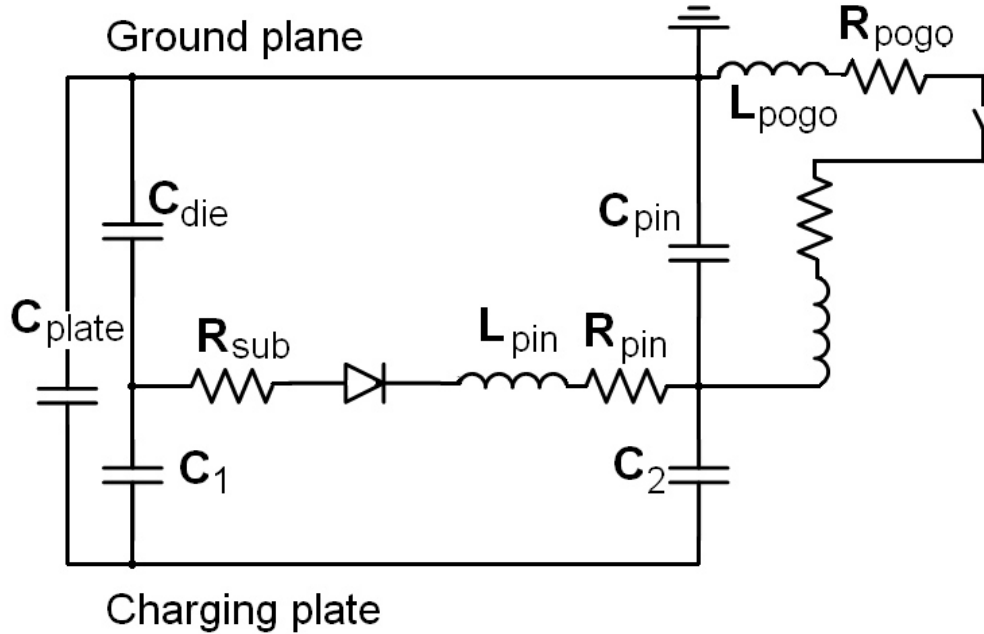


Figure 4.3: Simple circuit example

4.2 Overview of the Methodology

To illustrate how the LIM can be applied to the simulation of ESD events, we consider a simple circuit example shown in Fig. 4.3. The circuit in Fig. 4.3 represents a lumped model of all the key elements of the CDM test model in Fig. 4.2. The circuit essentially represents a single pin between the two plates of the test setup, with a single resistor R_{sub} modeling the substrate, and a diode representing on-chip circuitry. All nodes of the circuit are precharged to the desired stress voltage; then the switch is closed, the connection to the ground plane is established, and the discharge through the pogo pin occurs.

As mentioned earlier, the LIM algorithm requires that the circuit be represented as a network of nodes and branches containing latency elements. To enable the LIM simulation, the circuit in Fig. 4.3 must be augmented. The three nodes of the circuit have capacitors connecting them to the ground plane. Since in the LIM simulation the ground plane is treated as the ideal circuit ground with constant zero voltage, these capacitors can be treated as standard node capacitors in the LIM scheme. However, capacitors C_1 and C_2 are floating, and

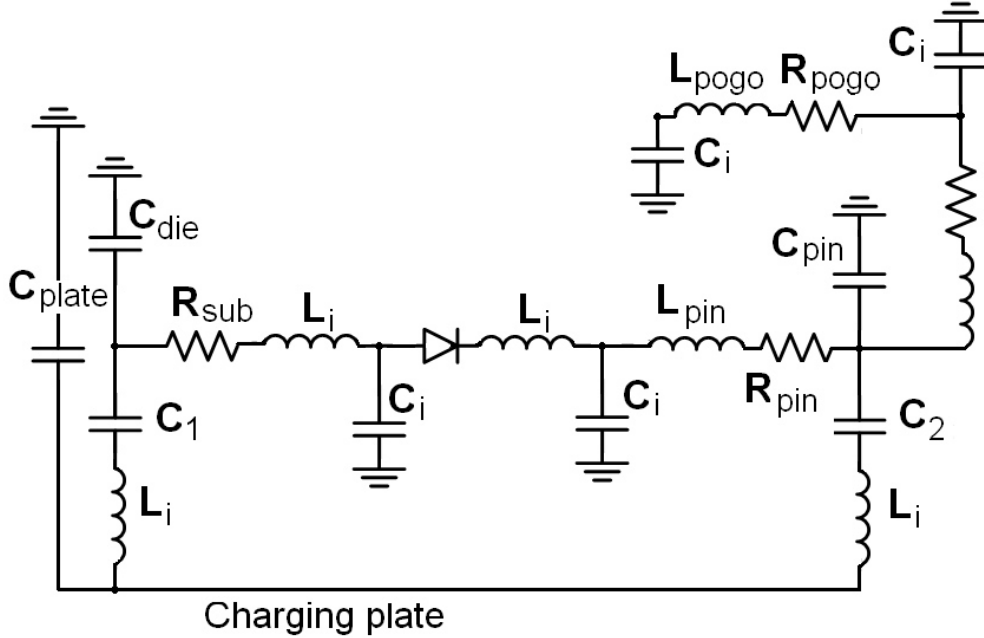


Figure 4.4: LIM-enabled simple circuit schematic

thus do not fit into the standard LIM format. Branch capacitors require special treatment in the LIM and are processed using the direct integration method [9]. Some nodes of the circuit do not have capacitance; small fictitious capacitors to ground are added at such nodes to create latency. Small fictitious inductors must also be inserted into some of the branches. The resulting augmented circuit is shown in Fig. 4.4.

Finally, the test circuit contains a nonlinear element, a diode. Nonlinear devices can be incorporated into the LIM algorithm as shown in [6]. The nonlinear equation for current flowing through the diode is solved iteratively at each step of the simulation.

Initially, voltage at all nodes of the circuit is set to the precharge value; then, voltage at the node representing the ground plane is forced to zero, and the simulation begins.

Results of the LIM simulation are shown in Fig. 4.5. The circuit was precharged to 300 V. Figure 4.5 shows a typical ESD transient current waveform with a very sharp spike occurring within the first 0.5 ns. The whole discharge event takes less than 3 ns. To verify the LIM results, the same simulation was performed using a modified nodal analysis-based

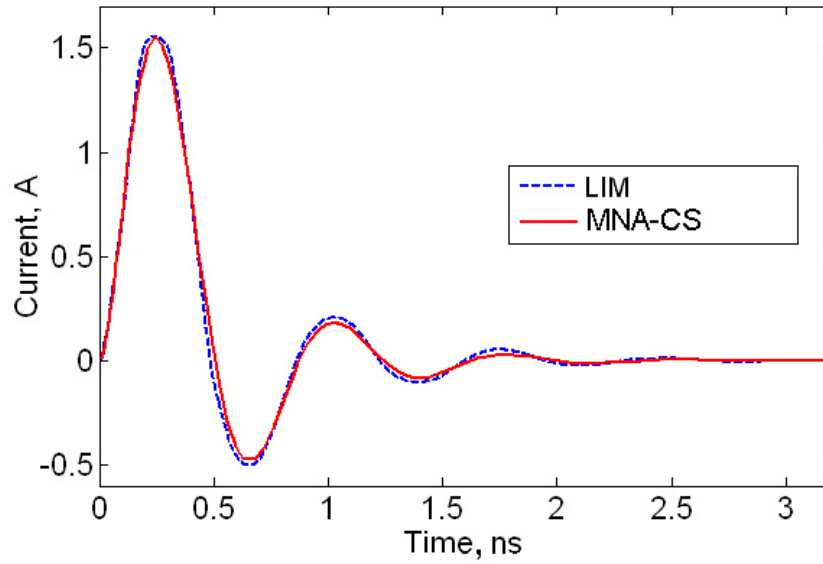


Figure 4.5: Small circuit ESD current waveform

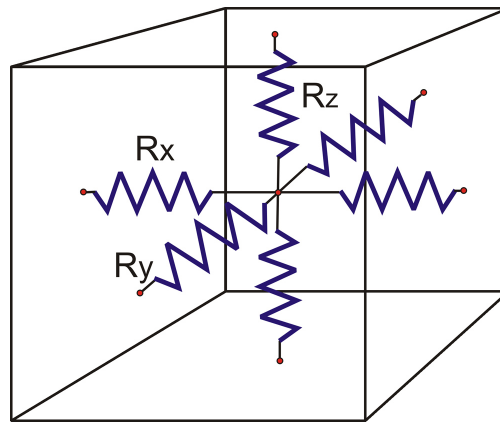


Figure 4.6: Unit cell of a lumped resistive substrate model

commercial circuit simulator (MNA-CS). As evident from Fig. 4.5, there is a good agreement with the LIM results.

The full chip model used for the simulation includes the discrete distributed resistive model for the substrate in the form of a three-dimensional resistive grid [24], [25]. The grid is composed of cells; each cell contains six resistors (as shown in Fig. 4.6).

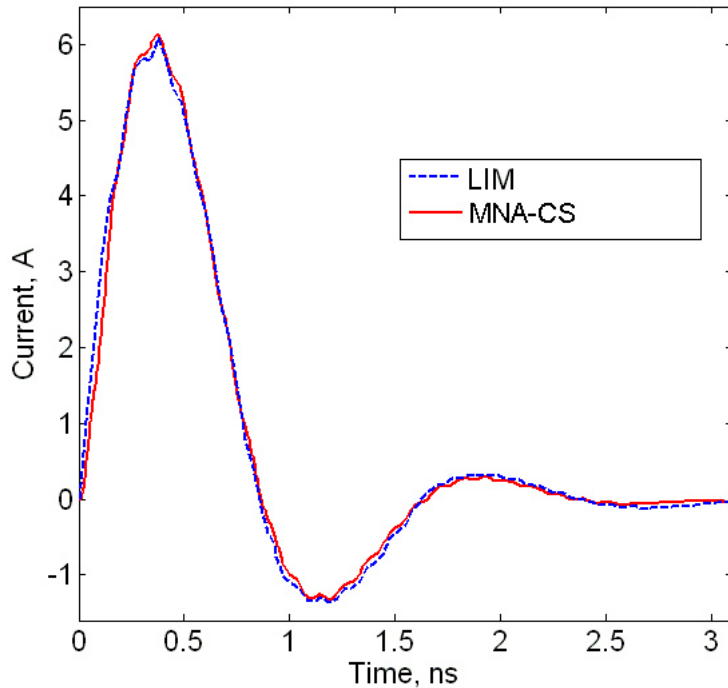


Figure 4.7: CDM ESD current waveform

4.3 Experimental Results

The model for the on-chip circuitry includes power buses, modeled with resistive chains. Diode-based ESD protection circuitry is also included, as well as decoupling capacitors. The circuit is precharged to 300 V, the discharge connection is realized via the low-resistance ($20\ \Omega$) pogo pin. The resulting transient current flowing through the pogo pin is shown in Fig. 4.7. MNA-CS is again used to verify the LIM results. Close agreement between LIM and MNA-CS results is observed.

The discharge current waveform observed at the pogo pin has the shape characteristic for the CDM ESD event. The first current spike exceeds 6 A, with a very fast rise time of roughly 0.5 ns. In the LIM simulation, we can probe voltage at any node and current through any branch in the circuit. For example, we can observe a transient voltage waveform at one of the nodes inside the substrate (Fig. 4.8).

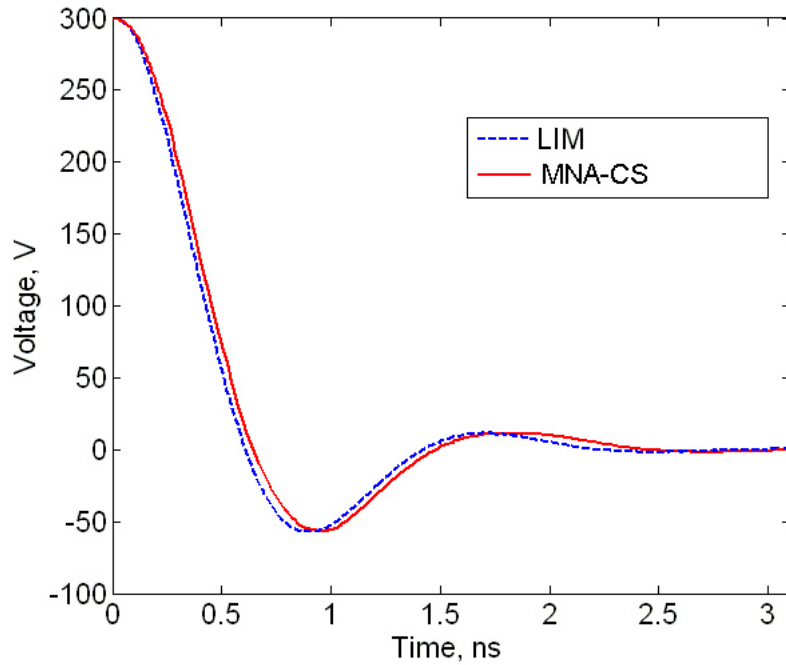


Figure 4.8: Transient voltage at one of the nodes inside the substrate

Table 4.1: Run time results.

Number of nodes	LIM (CPU sec)	MNA-CS (CPU sec)
65.5 K	7320 (122 min)	7800 (130 min)
82 K	9180 (153 min)	16020 (267 min)

Again, close agreement with the results of the simulation with the commercial tool can be observed.

Table 4.1 shows run time results for two model circuits. The difference between the two circuits is in the resolution of the resistive substrate model. The LIM run time results are compared with the commercial tool. All simulations are performed using severely under-powered LINUX workstation with 1 GB of RAM.

From Table 4.1 we can see that as the node count of the model increases, the LIM has a clear advantage over the commercial simulator. As the node count of a modern system on a chip (SoC) IC model can be well over a million, it is easy to see that the LIM is an attractive alternative to the conventional tools.

4.4 Enhanced Model

In the analysis described above, a purely resistive model of the chip substrate was employed. However, a more accurate model is a series resistive-capacitive one. A model that accounts for capacitive effects inside of chip substrate is cable of capturing transient effects emerging inside the substrate during the ESD event. A more accurate, enhanced substrate model can be incorporated into the LIM simulation. Although series capacitance requires special treatment in the LIM framework, in [9] it was demonstrated that series capacitors can be successfully modeled by the LIM algorithm.

Also, currently our model only includes diodes and disregards other nonlinear devices present on the chip. Recent advances in the simulation of nonlinear devices with the LIM [27] provide the means for incorporating transistors and logic blocks into the LIM simulation. We believe it is possible to perform a full-chip ESD event analysis with the LIM.

CHAPTER 5

TEMPERATURE-AWARE ANALYSIS USING LIM

5.1 Background

Driven by the aggressive scaling of modern integrated circuit (IC) technologies, functionality and speed improvements of IC designs are being achieved by increasing both the packing density and the clock frequency. As a result, IC power density has been rapidly increasing in accordance with Moore's law, and is now projected as a potential bottleneck for future performance improvements [28]. Not only are we challenged to distribute an increased power density, but we must also consider the equally daunting problem of removing the corresponding heat that is dissipated [29].

Elevated chip temperature can create many design challenges and reliability issues. First, high temperature and hot spots degrade the reliability of interconnects and transistors [30], [31]. Furthermore, on-chip thermal gradients can cause functional or timing failures through electrothermal coupling [30], [32]. Additionally, active power consumption and leakage are strong functions of the on-chip temperature profile, thereby making the prediction and minimization of power consumption problems inseparable from those of temperature analysis and control [33]. Moreover, the deployment of low-k dielectric materials to reduce capacitance (and hence, power dissipation) worsens the thermal transport due to the decrease in material thermal conductivities.

Recently, IC thermal effects have received a great deal of attention from the circuit design community, which has spawned several works. In [30], [32], and [34], the authors simulate the full-chip temperature profile by discretizing the partial differential equation (PDE) of

heat conduction using finite difference and finite element methods. The resulting discretized problem is then solved by adopting an equivalent circuit approach and a corresponding direct solution method. In [35], the full-chip thermal transients are solved in a similar manner using an alternating direction implicit (ADI) method for efficiency. The self-heating of multilevel IC interconnects and the corresponding impact on reliability and performance have been investigated in [32]. Of particular interest recently is the estimation of leakage power, which is exponentially dependent on device temperatures. In [36], leakage analysis was conducted for large industrial designs while considering power supply and temperature variations. The awareness of thermal effects has also spawned new research in design optimization. In [37] and [38], the thermal gradient was included as an optimization objective in cell-level placement. At the microarchitecture level, run time dynamic thermal management has been proposed as a means to regulate microprocessor operating temperature [39].

An efficient full-chip thermal analysis methodology is becoming increasingly important for the design and optimization of modern very large scale integrated (VLSI) systems. Such optimization problems include the packaging design, since the cost of the IC package and associated cooling [40] can dominate the IC product cost. However, considering the complex on-chip three-dimensional (3-D) multilayer structures, 3-D full-chip thermal analysis is a daunting task. The aforementioned discretized heat PDEs are often solved using direct methods or SPICE simulation by treating the heat equation as an equivalent resistance-capacitance (RC) circuit. While useful for small thermal problems, such an approach does not scale well with the complexity of the full-chip analysis [29].

5.2 Temperature-Dependent IR Drop Analysis

In Chapter 3 we have demonstrated that the LIM can be successfully applied to the analysis of IR drop in a steady state PDN. In our simulations, we assumed that interconnect resistance remains constant and independent of the amount of current transferred through

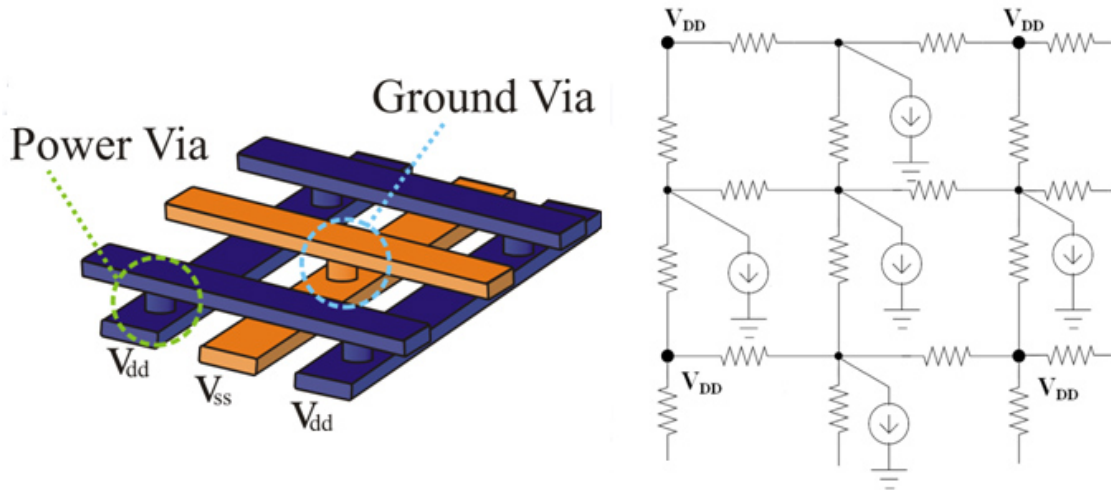


Figure 5.1: Power grid structure and a steady state equivalent circuit model

the wires. However, as the current densities of interconnects increase, the effect of self-heating (Joule heating) becomes more significant and cannot be ignored anymore. The temperature variation on the power grid can cause significant change in the interconnect resistances, and therefore can substantially increase the IR drops in the power grid. Since the IR drop and temperature are interdependent, a concurrent simulation has to be performed to achieve accurate results [41].

Figure 5.1 shows a multilayer power distribution network structure and its equivalent circuit representation. The model in Fig. 5.1 represents metal wires and vias as a resistive network. Independent current sources model average steady state currents, and independent voltage sources represent supply (VDD) pads.

Figure 5.2 illustrates the idea of a thermal-aware [41] IR drop analysis. Electrical and thermal simulations are run iteratively; the output of one simulation is used as an input for the other, and simulations are repeated until the solutions converge.

First an IR drop simulation can be performed. A certain temperature is assumed, and parameters (resistances) corresponding to that temperature are used. The electrical simulation determines the current distribution in the network.

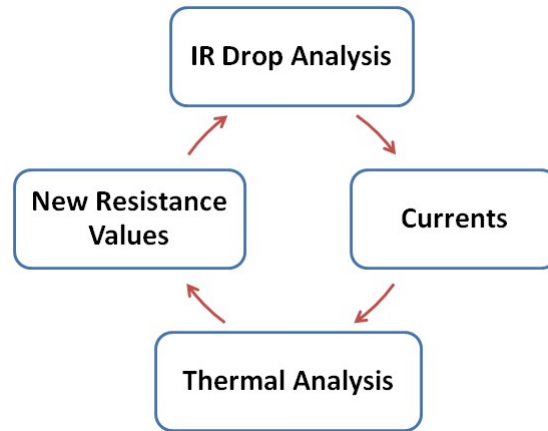


Figure 5.2: The flow of thermal-aware IR drop analysis

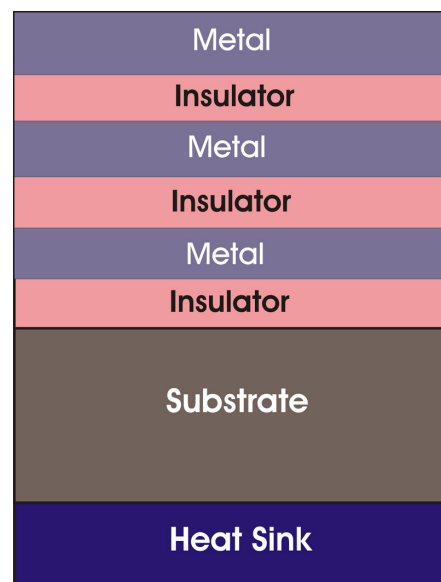


Figure 5.3: Multilayer structure of an IC

In the thermal simulation, two sources of heat are considered: heating from the substrate (due to the power dissipation in the devices) and self-heating of the wires (Joule heating). At first, for the sake of simplicity we can assume that the temperature profile of the substrate is available. An IC is represented with a multilayer structure of alternating metal and insulator layers, with a substrate and a heat sink at the bottom [41] (as in Fig. 5.3).

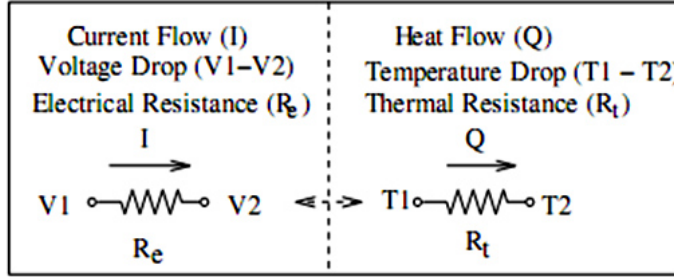


Figure 5.4: Analogy between thermal and electrical circuits

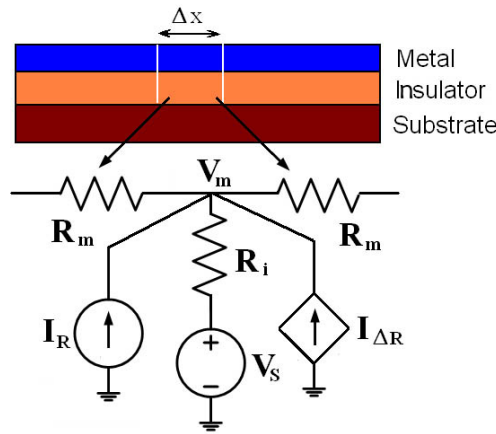


Figure 5.5: A lumped model of the interconnect thermal system

An analogy between thermal and electrical circuits (Fig. 5.4) can be employed for the construction of the equivalent thermal network to be used in the thermal simulation [32].

Metal interconnects are represented with their equivalent thermal resistance R_m . Interconnect temperature then becomes a node voltage V_m . Substrate temperature is represented with an independent voltage source V_s (Fig. 5.5).

Insulator thermal conductivity is also represented with an equivalent thermal resistor R_i . The Joule heating of the metal has two elements: I_R and $I_{\Delta R}$. I_R is the constant current (heat) source, the primary contributor to Joule heating, while $I_{\Delta R}$, represented by a voltage-controlled current source, accounts for additional heat due to increased resistivity caused by the interconnect temperature rise.

Once the thermal analogy is applied, the thermal problem essentially becomes an IR drop problem in the equivalent network. The equivalent network can be solved for node voltages using the same methodology that was initially applied to the electrical problem. The new voltage (temperature) distribution can then be obtained.

Once the new temperature profile is available, resistances in the electrical model can be recalculated at the new temperature. The electrical IR drop simulation will then be repeated. Electrical and thermal simulations will be iterated until the overall simulation converges to the steady state.

5.3 Analysis of Chip Substrate Temperature Profile

An interconnect temperature rise due to the heat conduction from the substrate (power dissipated by the active devices) is a significant factor and has to be accounted for in the thermoelectric interconnect analysis. In the previous section, we have assumed that the temperature profile of the chip substrate was known. We can further extend the LIM capabilities by assuming that only the heat sources are known. If the information about the power dissipated by transistors is available, the steady state temperature profile of the substrate can be obtained using the LIM.

Following [42], [30], and [32] and applying the analogy between a steady state thermal system and a resistive network, the thermal system representing the heat distribution in the substrate can be mapped onto a resistive network (as shown in Fig. 5.6).

Each heat source can be mapped onto a constant current source. Ambient temperature at the boundaries of the substrate can be modeled by constant voltage sources.

Once the equivalent thermal model is constructed, the problem of solving the steady state temperature becomes one of solving a resistive network. Such a network for contemporary devices is extremely large and not suitable for direct solvers. The LIM can be used to solve the mapped resistive network to obtain the steady state temperature profile.

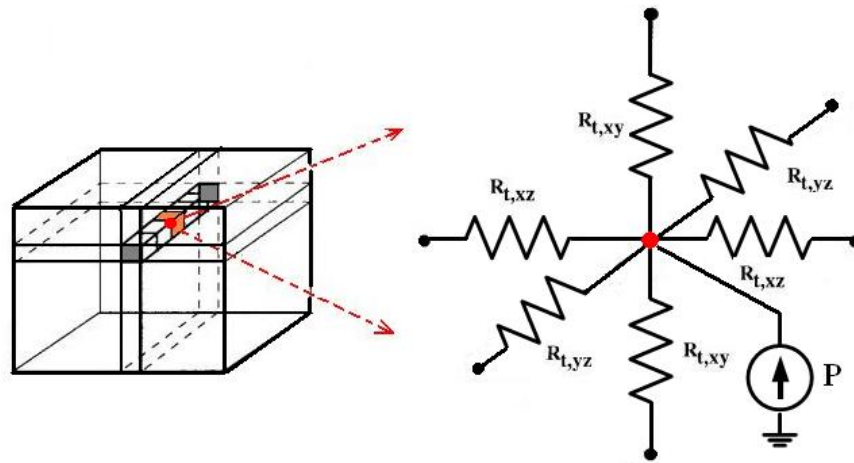


Figure 5.6: Analogy between the steady state thermal system and the resistive network

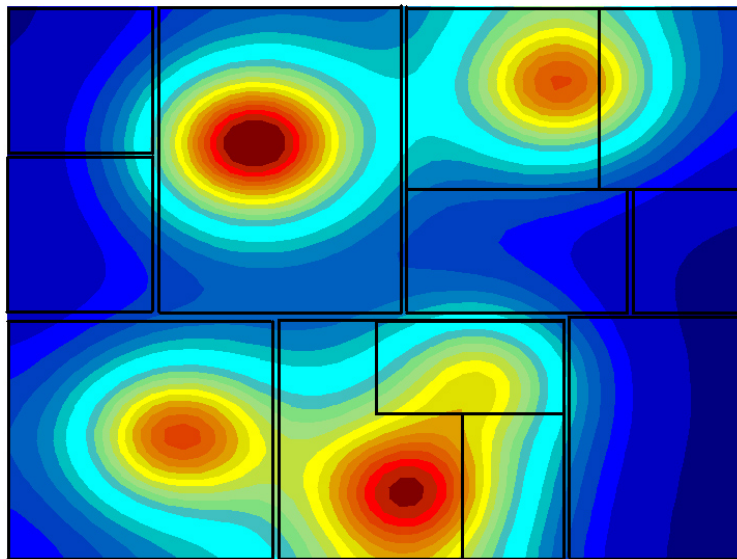


Figure 5.7: Temperature profile of a chip

Figure 5.7 shows an example of a thermal profile simulation with the LIM. The model structure represents the surface of an IC. It is assumed that the information about the floor plan of the IC is available, so the structure is divided into partitions with different thermal properties. Some partitions contain heat sources modeled with shunt current sources. It is also assumed that the surface of the IC is in contact with an ideal heat sink that maintains constant temperature.

<u>Thermal</u>			<u>Electrical</u>	
Temperature	T [K]	↔	Voltage	V [V]
Heat	Q [J]	↔	Charge	Q [C]
Heat transfer rate	q [W]	↔	Current	i [A]
Thermal resistance	R_T [K/W]	↔	Electrical resistance	R [V/A]
Thermal capacitance	C_T [J/K]	↔	Electrical capacitance	C [C/V]
Governing equations				
<u>Steady-State condition</u>				
Temperature Rise		↔	Voltage Difference	
$\Delta T = q R_T$			$\Delta V = i R$	
<u>Transient condition</u>				
Heat diffusion		↔	RC transmission line	
$\nabla^2 T = R_T C_T \frac{\partial T}{\partial t}$			$\nabla^2 V = RC \frac{\partial V}{\partial t}$	

Figure 5.8: Thermal-electrical analogous quantities

5.4 Analysis of the Transient Thermal Phenomena

In addition to the steady state on-chip temperature distribution, the thermal transients can also be of interest for various applications. For instance, in dynamic thermal management, the dynamic variation of on-chip temperature is used to adjust the operation of the chip so that the leakage power and the peak chip temperature can be properly controlled.

The analogy between thermal and electrical phenomena can still be employed; however, a more detailed (Fig. 5.8) mapping is required since effects of the thermal capacitance must be taken into account.

To solve the thermal transient analysis problem, one can model the thermal system as an equivalent RC circuit. Then a circuit simulation technique can be applied to the equivalent RC circuit to provide the thermal transient response. However, just as for the thermal steady state analysis, a direct solution method does not scale well with the size of the equivalent

circuit model. Again, the LIM can present an efficient alternative to the traditional methods. Using RC segments (similar to a transmission line model) in the equivalent thermal system, the thermal problem can be solved within the LIM framework.

CHAPTER 6

APPLICATION OF THE LIM TO THE ELECTRO-THERMAL CIRCUIT ANALYSIS AT THE EARLY DESIGN STAGES

6.1 Motivation

In previous chapters we have shown that the LIM can be successfully applied to the problem of the DC IR drop analysis as well as to the transient simulation of on-chip and off-chip interconnect networks. We also suggested several applications of the LIM to the problems of electro-thermal circuit analysis. In this chapter we will discuss application of the LIM to electro-thermal problems in greater detail. We also consider some examples derived from actual industry problems and demonstrate application of the method to the problems specific to the early, pre-layout design stages. We believe that the method can be particularly useful for the early trade-off and feasibility studies of on-chip and off-chip interconnect systems as well as entire chip and package structures. In particular, we demonstrate the capability of LIM to perform steady state and transient thermal analysis of a 3-D IC model that represents an actual industry design.

6.2 Problem Formulation

Advances in the silicon industry led to a variety of signal and power integrity issues. As designs get more complicated and design margins shrink, those issues have to be addressed at the earliest (at architecture or floor planning) stages of the design flow. The pre-layout stage becomes extremely important because the errors introduced at that stage are fundamental, are likely to affect the performance and placement of all system components, and

can potentially be fatal for the product, thus requiring a complete redesign and leading to unacceptable time-to-market delays.

At the pre-layout stage, detailed information about system components and their characteristics is not available. Therefore, the models used in the analysis can be relatively simple. However, with the ever-increasing complexity of integrated circuits, even these simplified models require very large computational effort. The large number of elements in the equivalent circuit model creates computational challenges that are not handled well by most of the industry-standard, matrix-based CAD tools. Also, at that stage of the design process, there is a need for quick estimates, and simulations have to be run multiple times with minor adjustments to the model. Hence, tools are needed that can perform fast and efficient analysis of relatively simple but potentially very large networks. It is desirable to simultaneously address multiple interdependent issues in various components of the system, ensuring electrical, thermal, and mechanical reliability of the future product.

We propose the latency insertion method (LIM) as the tool that has the power and versatility necessary for performing a complex electrical and thermal analysis of the systems at early design stages. Our focus is on the modeling of on-chip interconnect networks and phenomena that affect their performance in tightly packed, very dense modern ICs. In particular, we consider on-chip power delivery system, with such effects as the steady state IR drop and its dependence on temperature.

6.3 Pre-Layout Steady State IR Drop Analysis Using LIM

At the pre-layout stage, information about the circuit is available in the form of a floor plan, pad out, and current and voltage budgets. The model for power integrity analysis is then built based on the process parameters, supply voltage values, required current loads, general geometry, and the floor plan.

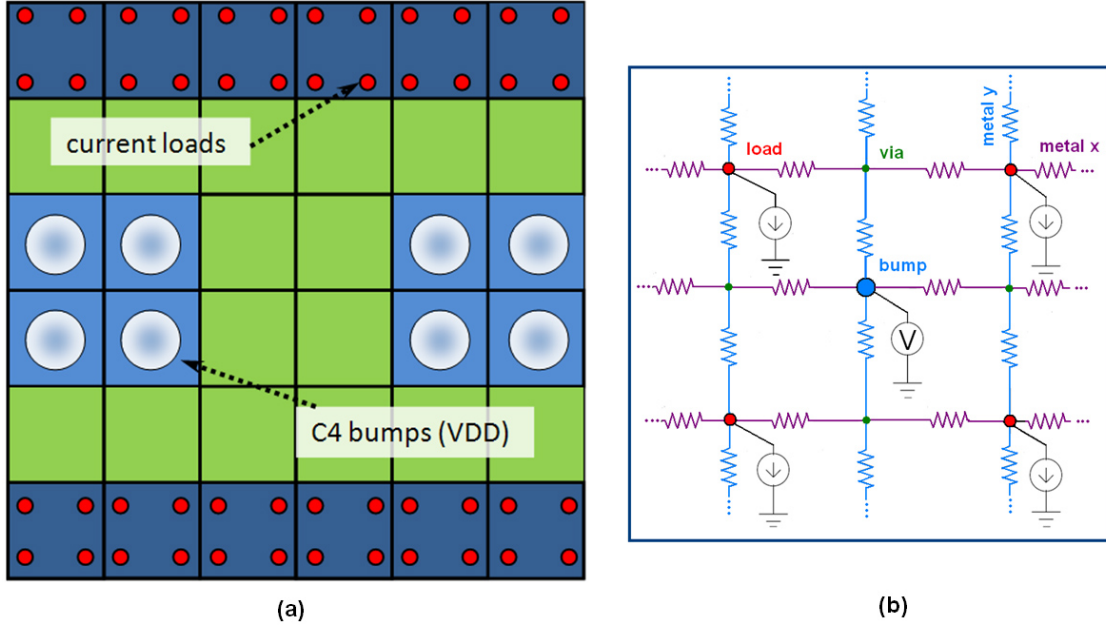


Figure 6.1: Model structure for IR drop analysis

We consider an example based on the structure in Fig. 6.1, which represents a segment of a power delivery network. The subject of the analysis is the VDD net in the top two metal layers. The floor plan is divided into six rows and six columns. The connections to the voltage supply (C4 bumps) are on the top level in the two middle rows (as shown in Fig. 6.1(a)). The power is drawn from the second metal layer at the top and bottom rows. Each metal segment (a square in Fig. 6.1(a)) is represented with a resistive grid of the form shown in Fig. 6.1(b). Metals in two layers are orthogonal to each other and interconnected by vias. If the segment has a connection to the power supply, a constant voltage source is inserted in the center (a bump). If the segment has connections to the active devices, the power drawn by those devices is modeled with constant current sources (current loads).

The resulting resistive network has the structure common for steady state IR drop problems and can be efficiently analyzed with the LIM. The results of the LIM simulation are verified by simulating the same structure in HSPICE. The results are shown in Fig. 6.2.

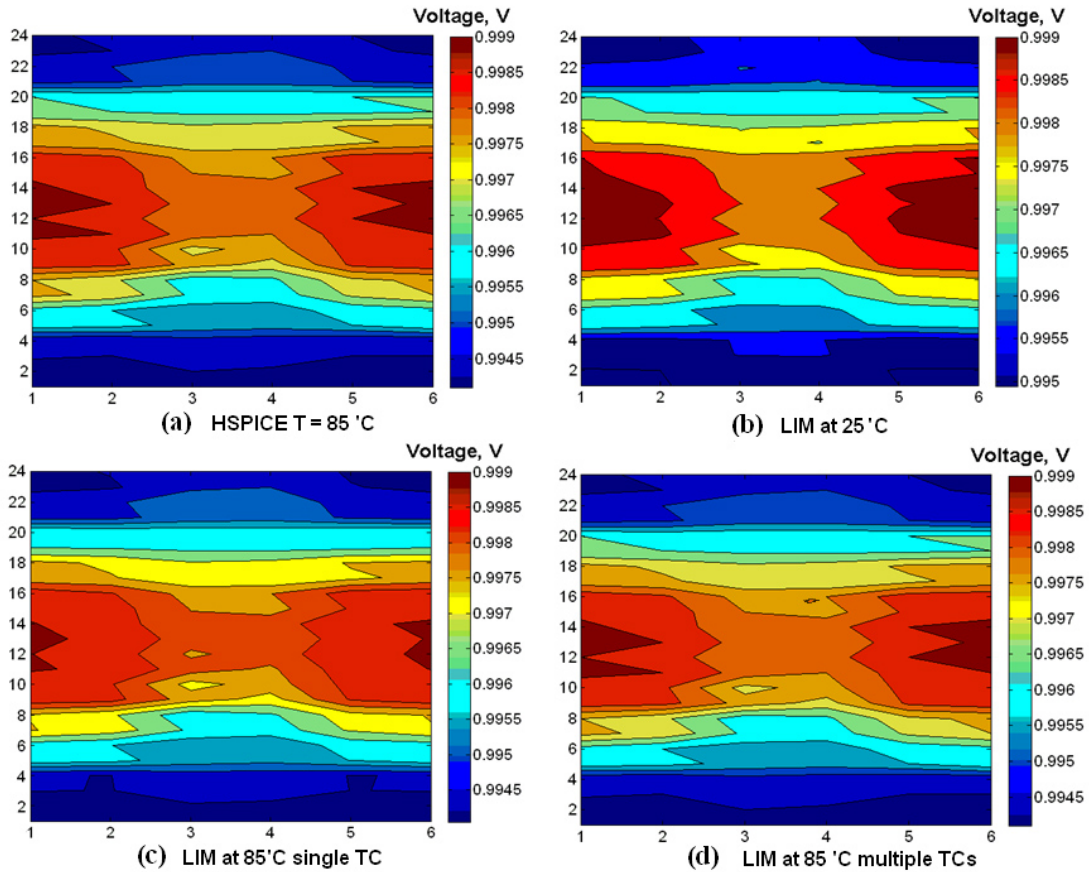


Figure 6.2: Pre-layout IR drop analysis results

The supply voltage level in this example is 1 V. Figure 6.2 displays voltage at four load points of each segment. It is clear that voltage is higher in the midsection at the locations of the supply connections. Voltage levels drop toward the top and the bottom parts of the plots, which correspond to the rows where current is drawn from the supply.

In general, in this example the voltage across the whole structure stays very close to the nominal supply value, and the range of values is only 6 mV. The benchmark HSPICE simulation (Fig. 6.2(a)) assumes the temperature of 85 °C. The first simulation of the structure with the LIM is performed using nominal values of resistances (at 25 °C). It can be seen from Fig. 6.2(b) that voltage values predicted by the simulation are overestimated. In the next LIM simulation, the temperature is taken into account and set to 85 °C.

For simplicity, a single value of the temperature coefficient was used for resistances in all metal layers. The voltage distribution in Fig. 6.2(c) very closely matched that of the benchmark. Finally, with differences in the thermal coefficient of resistance of the three metals taken into account, Fig. 6.2(d) is almost identical to the results from the HSPICE benchmark. It can be seen from Fig. 6.2 that LIM has enough precision to correctly resolve voltage values in the narrow range of 6 mV.

6.4 Electro-Thermal Analysis Using the LIM

The differences in Fig. 6.2(a), (b), and (c) demonstrate the well-known fact that resistance depends on temperature. To obtain the correct estimate of the IR drop profile in the power grid, the operating temperature of the interconnect network must be factored in.

There are two main sources of interconnect heating: the self-heating (or Joule heating) due to current passing through resistive wires, and the raised temperature of the substrate due to power dissipated by the active devices.

The problem of self-heating can be modeled by creating an equivalent electro-thermal network based on a thermal-electrical analogy. In such a network, thermal resistances account for the thermal conductivity of the material; heat flux is then represented by the electrical current, and voltage is equivalent to temperature [32]. The equivalent steady state thermal problem essentially becomes a problem of steady state IR drop analysis in the equivalent network. The resulting network can be solved using an electrical simulator (such as SPICE). However, the electrical and thermal simulations are interdependent and therefore have to be iterated until convergence is achieved. The number of elements in such electro-thermal problems makes traditional matrix-based direct solvers inefficient, since the complexity of these methods does not scale well with the size of the model. A number of alternative techniques have been proposed to address this issue [41], [29]. We suggest that the LIM can be successfully applied to the electro-thermal analysis of large on-chip and off-chip intercon-

nect systems. The added advantage of the LIM is its ability to efficiently model transient behavior, whereas most other techniques targeting the steady state IR-drop analysis cannot perform transient simulations.

Another source of interconnect heating (and the dominant one) is the heat generated by the active devices, which raises the overall temperature of the substrate. While the problem of self-heating of interconnects is addressed in [5], the substrate temperature is assumed to be known. In reality, particularly at the early design stages, the temperature profile of the substrate is not readily available and needs to be calculated. Therefore, the problem of thermal analysis of the substrate containing the interconnect network is added to the problem of electro-thermal analysis of the interconnect system itself. The number of elements in the resulting problem almost inevitably exceeds the capabilities of traditional methods. In this dissertation we show that the LIM can be employed to obtain the thermal solution for the system and hence, potentially, the LIM can be used for the complete analysis of on-chip power integrity problems.

6.4.1 2-D Benchmark Test Example

To demonstrate the application of the LIM to thermal analysis, we consider an example of 2-D heat transfer problem with convection shown in Fig. 6.3, for which the analytical solution exists [43]. The structure in Fig. 6.3 represents a solid block of a material with the thermal conductivity of $52 \text{ }^\circ\text{C}/\text{W}$. The source of heat is the edge AB with constant temperature of $100 \text{ }^\circ\text{C}$. The heat can only escape through edges BC and CD via convection to the ambient temperature of $0 \text{ }^\circ\text{C}$. The surface convective heat transfer coefficient on edges BC and CD is $h = 750 \text{ W}/(\text{m}^2 \text{ }^\circ\text{C})$. The target solution for the temperature at point E is $18.3 \text{ }^\circ\text{C}$.

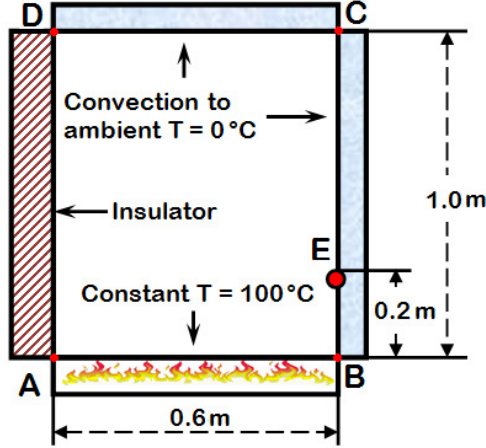


Figure 6.3: Test example setup

We employ the analogy between thermal and electrical circuits. The bulk of the material is modeled with 2-D resistive grid with uniform square cells. Values of the grid resistors are calculated based on the thermal conductivity of the material and the unit cell size as

$$R_t = \frac{\Delta x}{kL} \left[\frac{^\circ\text{C}}{\text{W}} \right] \quad (6.1)$$

where k is the thermal conductivity and L is the size of a unit cell. The fixed raised temperature on the edge AB is introduced via constant voltage sources along the boundary. Boundary conditions on the convective edges are modeled with resistances connected from the grid nodes on the edges to constant voltage sources (the ambient temperature) [32], [42]. Values of the interface resistors are derived from the heat transfer coefficient h as

$$R_h = \frac{1}{hA_e} \left[\frac{^\circ\text{C}}{\text{W}} \right] \quad (6.2)$$

where A_e is the effective area of a unit cell of the grid.

Adiabatic boundary conditions are assumed on the AD side of the structure, so there is no heat transfer through that edge. The general cell structure of the equivalent network is shown in Fig. 6.4.

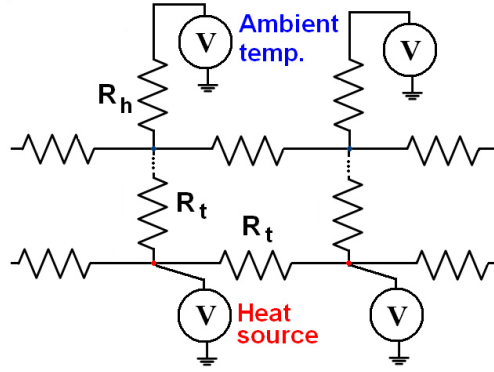


Figure 6.4: Thermal-electrical model structure

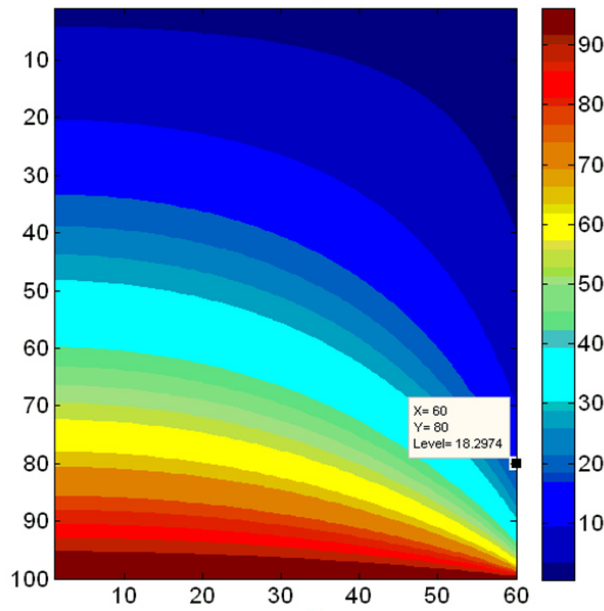


Figure 6.5: Simulation results for the benchmark example

Results of the simulation using equivalent resistive circuit representation for the model and the LIM-based solver are shown in Figs. 6.5 and 6.6. As can be seen in Fig. 6.6, results of the LIM simulation exactly match the target analytic solution.

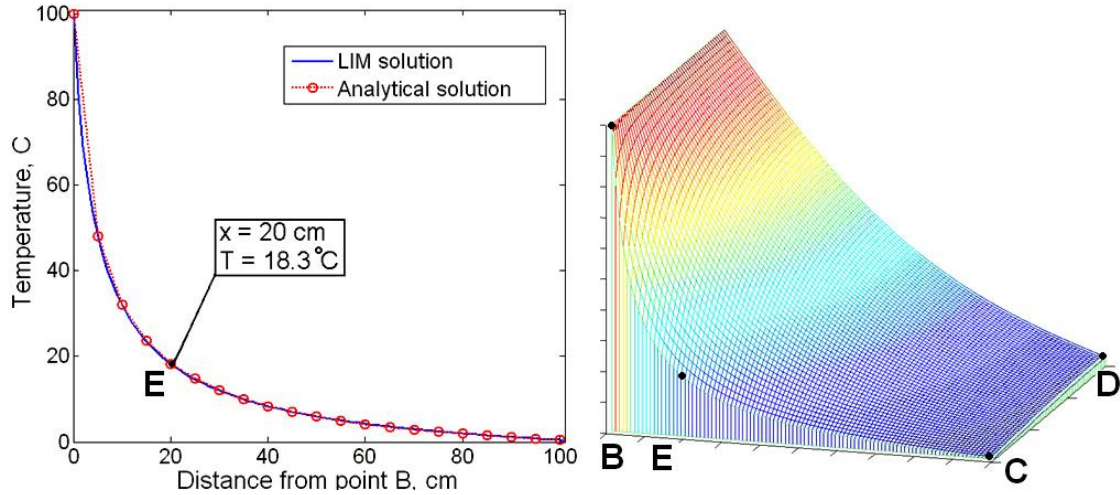


Figure 6.6: Comparison of the results for the benchmark example

6.4.2 Temperature Distribution Analysis in a Stacked 3-D IC

Next, we consider the 3-D model of the structure in Fig. 6.7, which consists of two silicon chips with a composite interposer substrate. On the top is a controller chip and on the other side of the interposer is a memory chip. The two chips are connected through vertical connections. Solder ball grid arrays with under-fill are used to attach both chips to the substrate.

Placing controller and memory chips on top of each other (rather than side by side) shortens the length of interconnects. Reduction in the length of interconnects lowers signal propagation delay and mitigates various signal integrity issues related to signal interference and attenuation. These improvements allow faster communication between the two chips and, therefore, increase performance of the system.

Close proximity of the two chips, beneficial in terms of chip-to-chip communication, has certain downsides. Because of the unconventional architecture, the system requires a complex multistage manufacturing process, which increases the cost of the device and lowers the yield. Another issue, the one that we will address, is thermal coupling between the chips. While the memory on its own does not dissipate a lot of power, it can get heated by the controller

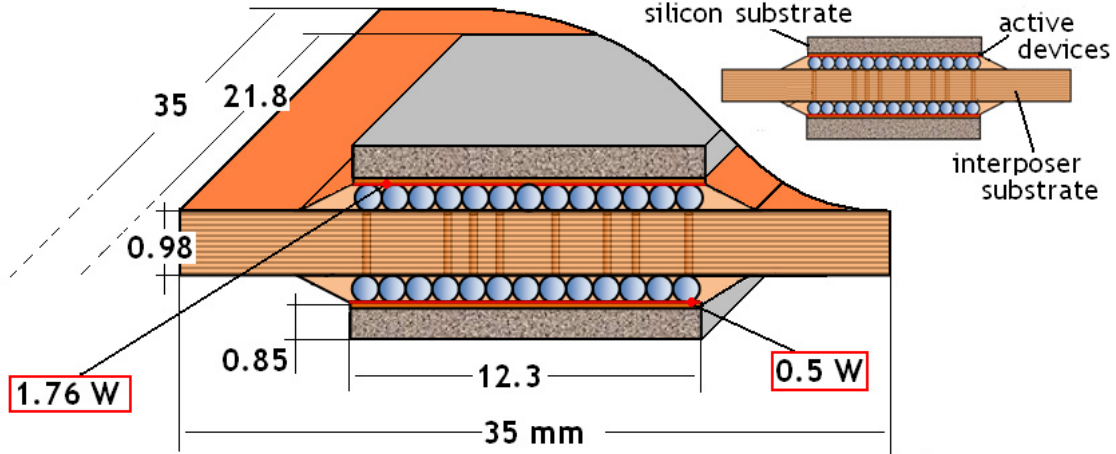


Figure 6.7: System of two chips with an interposer substrate (cross section)

via heat conduction through the interposer substrate. Raised temperature increases leakage in memory cells, which lowers memory retention time (and therefore increases the required refresh rate). Since leakage is exponentially dependent on temperature, the impact can be catastrophic.

It becomes crucially important to perform early architectural studies of 3-D ICs. Given the material composition of the system and approximate power numbers, the operating temperature of the system has to be determined and the feasibility of the design has to be assessed.

We again use a resistive grid to model the bulk of the structure, but now we add the third dimension (Fig. 6.8).

The equation (6.1) for resistors in the x -direction becomes (6.3), where A is the area of a unit cell.

$$R_x = \frac{\Delta x}{kA} \left[\frac{^\circ\text{C}}{\text{W}} \right] \quad (6.3)$$

For our example we use uniform mesh with cubic cells. Then the unit resistance simply becomes

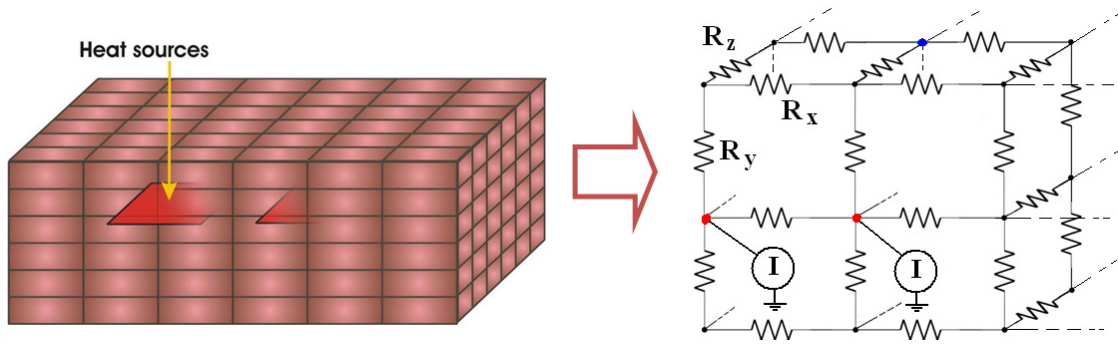


Figure 6.8: 3-D resistive grid model

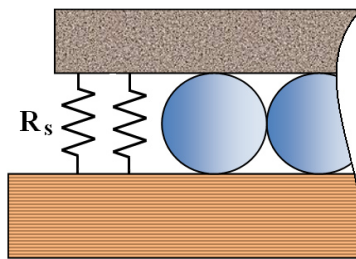


Figure 6.9: 1-D model for solder ball grid array

$$R = \frac{1}{k\Delta x} \left[\frac{^{\circ}\text{C}}{\text{W}} \right] \quad (6.4)$$

We model the solder ball grid array with 1-D resistors since there is little lateral heat transfer (Fig. 6.9). It is possible to model every single solder ball, but that approach is impractical and does not account for the thermal conductivity of the under-fill material. We find the average thermal conductivity of the combination of solder balls and the under-fill epoxy and calculate values of individual resistors based on the mesh density (the number of nodes that fit in the area under the chip). These 1-D resistors are placed between the nodes of the mesh that models the silicon chip and the one that represents the interposer, “stitching” the two together.

Exposed surfaces of the structure require special treatment because they are cooled not by conduction but by the combination of convective heat transfer and radiation. As before,

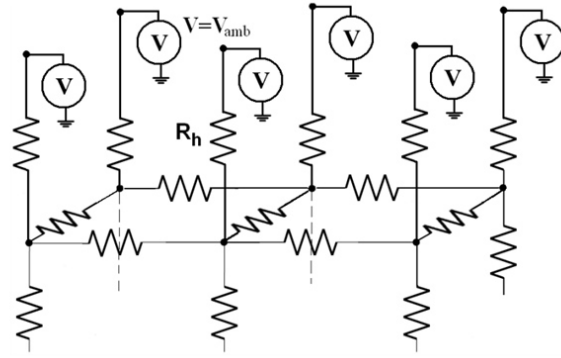


Figure 6.10: Model for convective boundary

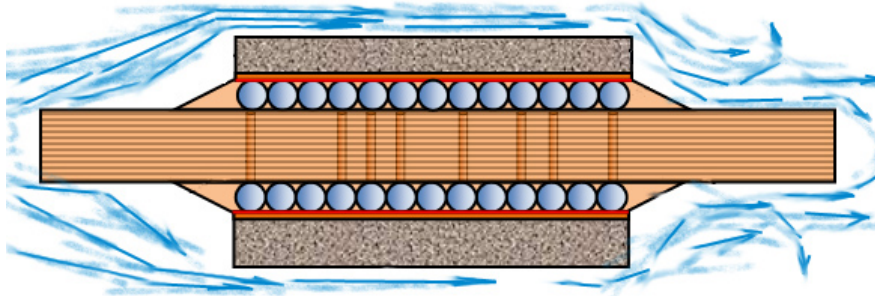


Figure 6.11: Forced air cooling

we will use the effective thermal resistance “trick” and define R_h based on the values of the effective heat transfer coefficient h_e (also called film coefficient). Effective resistances R_h are connected between the nodes on the surfaces of the structure and constant voltage sources that represent the ambient temperature (Fig. 6.10).

We can assume that the system is cooled by natural convection; then the approximate value of the effective heat transfer coefficient is $5 \text{ W}/(\text{m}^2 \text{ }^\circ\text{C})$. However, such an assumption is not realistic, since the structure will melt. A somewhat better assumption is forced air cooling (Fig. 6.11) with air streaming at 2 m/s (which corresponds to a good computer case fan).

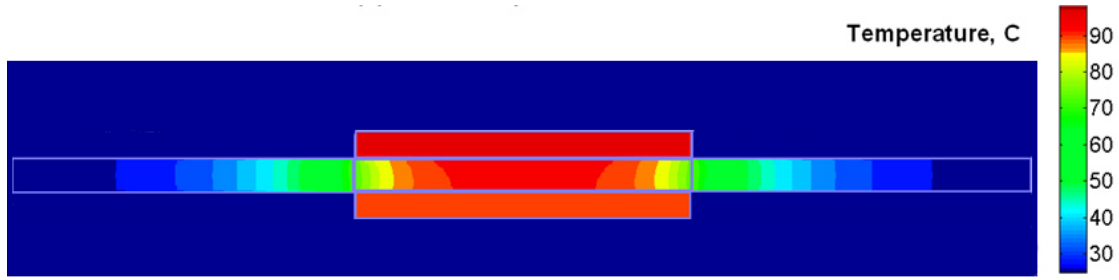


Figure 6.12: Temperature distribution within the structure (cross section)

A reasonable estimate for the heat transfer coefficient in that case is $100 \text{ W}/(\text{m}^2 \text{ }^\circ\text{C})$. We use this value in our calculations of the effective thermal resistance at the exposed surfaces of the structure.

Sources of heat in our model are represented with constant current sources. We use total power dissipation numbers for each chip to calculate the value of a unit source. We then distribute the sources evenly so that each node at the bottom of the controller chip and at the top of the memory chip has a current source connected to it.

The resulting thermal-electrical model is simulated with the LIM, and the results are shown in Fig. 6.12.

In Fig. 6.12, the chip on top dissipates almost four times more power than the one at the bottom and reaches temperatures close to $100 \text{ }^\circ\text{C}$. The ambient temperature is set to $25 \text{ }^\circ\text{C}$. The organic interposer substrate intrinsically has low thermal conductivity. However, the substrate is interlaid with copper planes, which increase heat spreading in lateral directions. Top and bottom chips have vertical interconnect structures between them, which increase heat transfer across the substrate.

For verification, the same structure is modeled in the commercial CAD tool ANSYS Icepak. The results are shown in Fig. 6.13.

In the case of the Icepak simulation, forced air cooling is assumed again at 2 m/s , with airflow from left to right. As can be seen in Fig. 6.13, the resulting temperature distribution is not symmetric. The heat transfer is more efficient at the leading (left) edge of the structure,

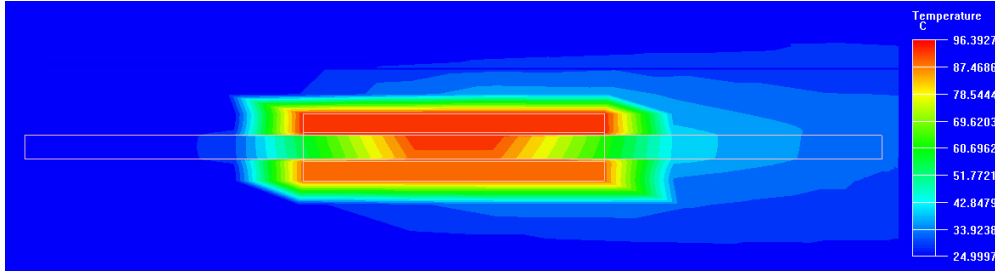


Figure 6.13: Temperature distribution within the structure. ANSYS Icepak

where the air is the coolest and the boundary layer has not yet been formed. The cooling is much less efficient on the right side of the structure. In our first LIM simulation, we use uniform values of the effective heat transfer coefficient to calculate the value of the resistance R_h that models the conjugate heat transfer, which results in a symmetric temperature profile (as in Fig. 6.12). We can account for the difference in heat transfer rates by assigning different values of R_h to different locations on the surfaces of the structure. For purposes of this example, we want to keep our model as simple as possible. For that reason, uniform mean values of R_h are used in the LIM model for the left and right parts of the structure (higher R_h on the right). In general, a good agreement of the LIM and Icepak results can be observed in Fig. 6.14. We are not concerned with the detailed modeling of the heat transfer between the chip and the ambient air. Our goal is to get a quick and reasonably accurate estimate of the temperature profile of the chip and use that information in the subsequent power integrity analysis. For those purposes, the LIM performance is more than adequate. Figure 6.14(b) demonstrates the correct temperature range and distribution when compared to the results obtained with the commercial CAD tool.

In fact, temperature distribution within a single cross section of the structure can be obtained from a 2-D simulation with reasonable accuracy since in our example the structure is very symmetric. However, a full 3-D simulation gives us the ability to look at any “slice” of the system. For example, Fig. 6.15 shows a top view of the structure. Also, if we obtain the detailed information about the power distribution profiles of the chips and/or the via

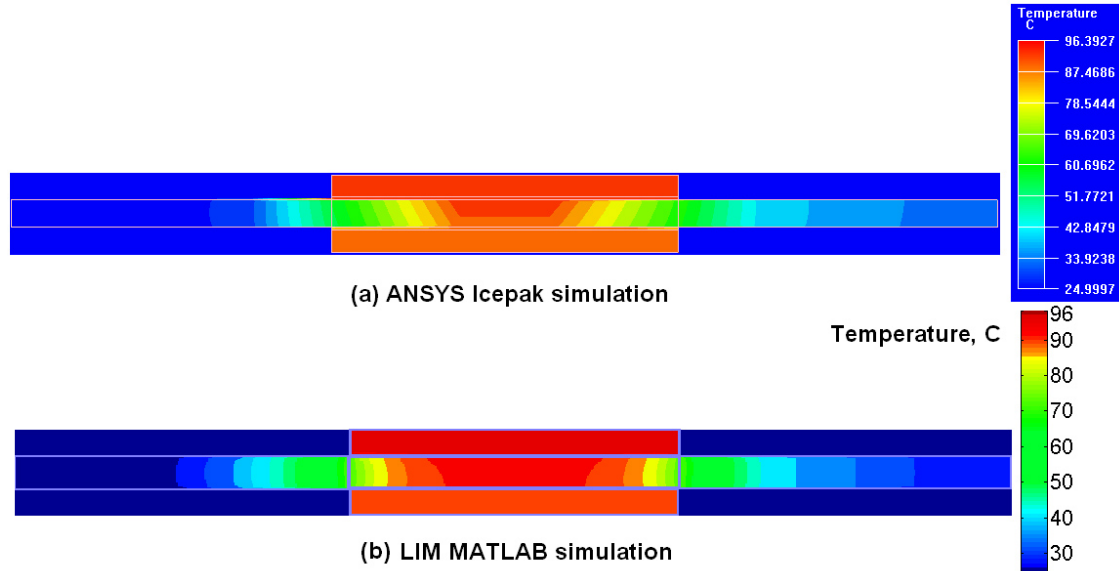


Figure 6.14: Temperature distribution within the structure (cross section) comparison

density inside of the interposer substrate and want to use that knowledge, a 3-D simulation will be required.

As was mentioned previously, one of the major concerns with the chip-interposer-chip design is the heating of the memory chip by heat conduction from the controller. In our model, we can easily analyze the two scenarios: one with the controller chip powered down and the other with the controller switched on. As can be seen from Figs. 6.16 and 6.17, the controller is indeed the main contributor of the heat in the system. It is evident that the assumed forced air cooling solution will not be sufficient to keep the memory chip at a reasonable temperature.

6.4.3 Model Size Issues

In general, once the equivalent electrical model for the thermal problem is set up, the resulting circuit can be analyzed using any circuit solver, such as HSPICE, for example. However, the equivalent network tends to be very large, which renders direct solvers inefficient.

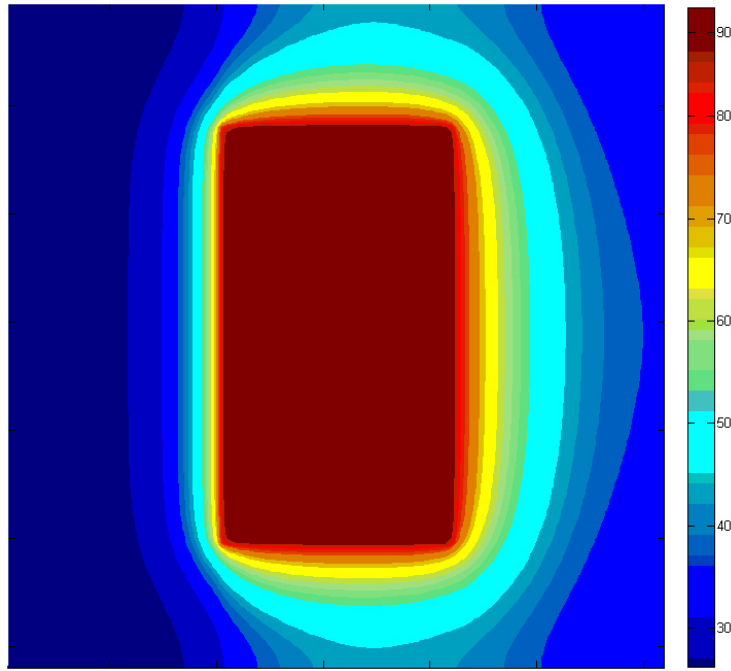


Figure 6.15: Top view of the structure

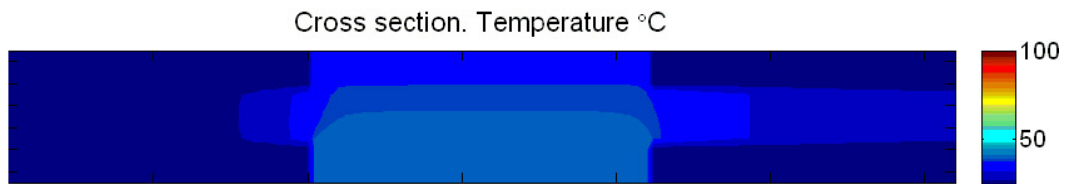


Figure 6.16: Controller off

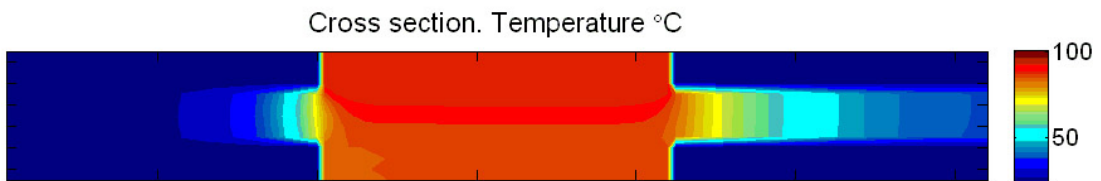


Figure 6.17: Controller on

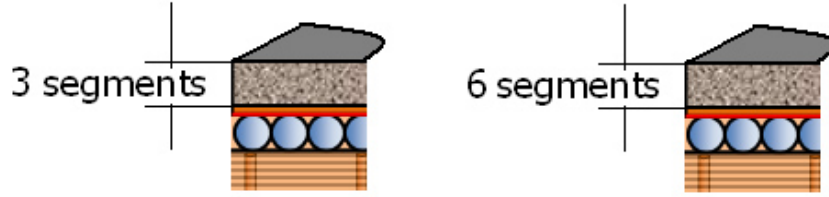


Figure 6.18: Two levels of discretization

Table 6.1: Run time results.

Number of elements	LIM (C++) (CPU sec)	HSPICE (CPU sec)
461,257	7	27
3,514,400	65	949

The number of elements in the model depends on the mesh density. The unit cell dimensions are determined by the smallest feature size that needs to be resolved in the simulation. In our first simulation of the 3-D stacked design, we used a very simple consideration for choosing the largest unit cell dimensions. We wanted to see some thermal distribution in the vertical direction, so we had to have at least several unit cells modeling the thickness of the two chips. We only used three cells for the thickness of silicon chips and four cells for the interposer. Since the horizontal dimensions of our system are much larger than the vertical ones (which is generally the case with IC components), we still ended up with nearly half a million elements in our model.

HSPICE is still efficient enough for model sizes below a million of elements. However, if we want to increase our resolution to the thickness of six cells per chip (Fig. 6.18), which is still arguably a fairly low level of discretization, the model size exceeds three million elements. Run time results in Table 6.1 show that even the low-resolution model takes a while to analyze using HSPICE; any further increase in mesh density makes the use of HSPICE unfeasible. The LIM run time scales linearly with the model size, making the LIM an attractive alternative to direct solvers.

6.5 Analysis of Thermal Transients Using the LIM

For certain applications, such as run time dynamic thermal management at the microarchitecture level to regulate microprocessor operating temperature [39] and dynamic leakage estimation, information about the transient thermal behavior of the system is of interest.

Latency in heat conduction is characterized by the quantity called heat capacity. An object's heat capacity C is defined as the ratio between the amount of heat energy Q transferred to the object and the resulting increase in temperature of the object ΔT , as in (6.5).

$$C = \frac{Q}{\Delta T} \quad (6.5)$$

Material properties are usually characterized by the specific heat capacity (specific heat), the heat capacity per unit mass:

$$c_p = \frac{Q}{\Delta T \cdot m} \left[\frac{\text{J}}{\text{kg} \cdot \text{K}} \right] \quad (6.6)$$

It is easy to extend the electro-thermal analogy that we used previously by drawing the analogy between the heat capacity and the electrical capacitance. Then, if thermal transients need to be accounted for in the simulation, we can translate our thermal problem into the equivalent RC network instead of a purely resistive one, as in Fig. 6.19.

In the equivalent model, capacitors are connected between every node in the circuit and the common ground. Equivalent electrical capacitance can be calculated based on the specific heat and the size of a unit cell. Once the unit size Δx is determined, we can find the volume of that cell. Assuming a uniform mesh with cubic cells is used, the unit cell volume is $\Delta V = \Delta x^3$. Then, the mass of the unit cell is $\Delta m = \rho \Delta V$. Finally, the capacitance associated with a unit cell can be found as

$$C_{eq} = \Delta m \cdot c_p \left[\frac{\text{W} \cdot \text{s}}{\text{K}} \right] \quad (6.7)$$

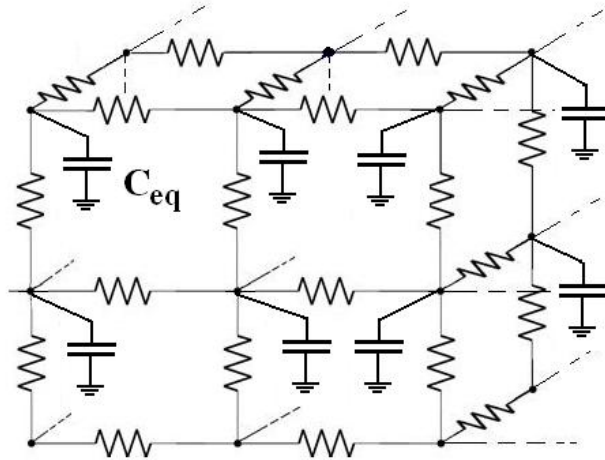


Figure 6.19: A segment of an equivalent model for transient thermal simulations

The LIM simulation requires presence of latency elements; therefore, any LIM simulation is inherently transient. In the case of the purely resistive equivalent electro-thermal model, we insert arbitrary (or, for simplicity, unitary) values of latency elements. Now if we use the meaningful values of capacitance from (6.7) in our simulation for the system in Fig. 6.7, the transient phase of the simulation reflects the transition of the system from the initial ambient temperature to the steady state operational temperature. Temperature of a node inside the controller chip as the function of time is shown in Fig. 6.20. The curve in Fig. 6.20 has the expected exponential shape. The system heats up and reaches the steady state. Temperature distribution profiles of the system at various time points can also be obtained.

Because of the nature of the LIM algorithm, we always have to include inductance in our model. The analogy between the thermal and electrical phenomena is not complete because there is no such thing as “thermal inductance.” If there were such a phenomenon, that would imply that the heat could be transferred from a hot object to a cold one. That would violate the second law of thermodynamics and has never been observed experimentally. Therefore, our LIM model is inevitably nonphysical. However, as long as we only care about the steady state thermal solution, we do not have to worry about the fact that our model contains nonphysical elements. The simulation converges to the steady state values that do

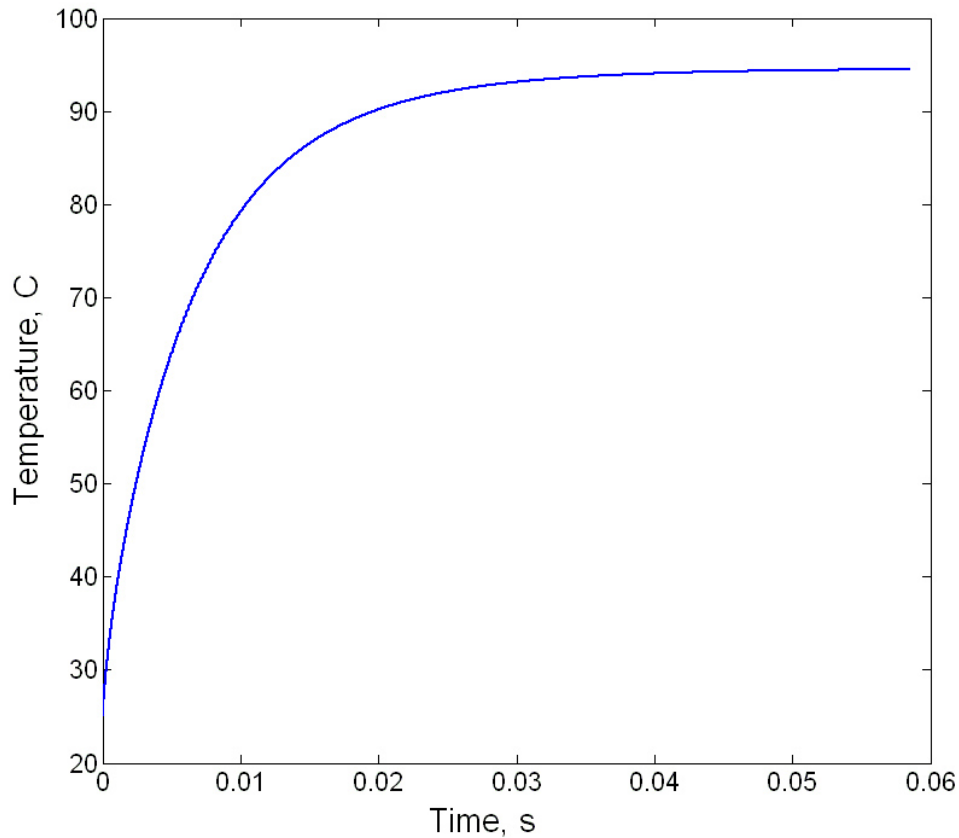


Figure 6.20: Transient thermal behavior of the system in Fig. 6.7

not depend on the inserted fictitious latency elements. However, if we want to model the transient behavior of the system, we have to make sure that our fictitious inductances are small enough and do not affect the transient solution. If fictitious inductance is too large, the solution oscillates and can look like the one in Fig. 6.21.

The issue of choosing small enough fictitious inductance values requires further investigation and is not within the scope of this work. For the time being, we use practical considerations in our models for transient simulations. Since the model for transient thermal analysis includes actual capacitance, fictitious inductance must be small enough, relative to that capacitance, so that capacitive effects dominate. At the same time, we need fictitious inductance to be large enough to enable a reasonable time step of the simulation, such that

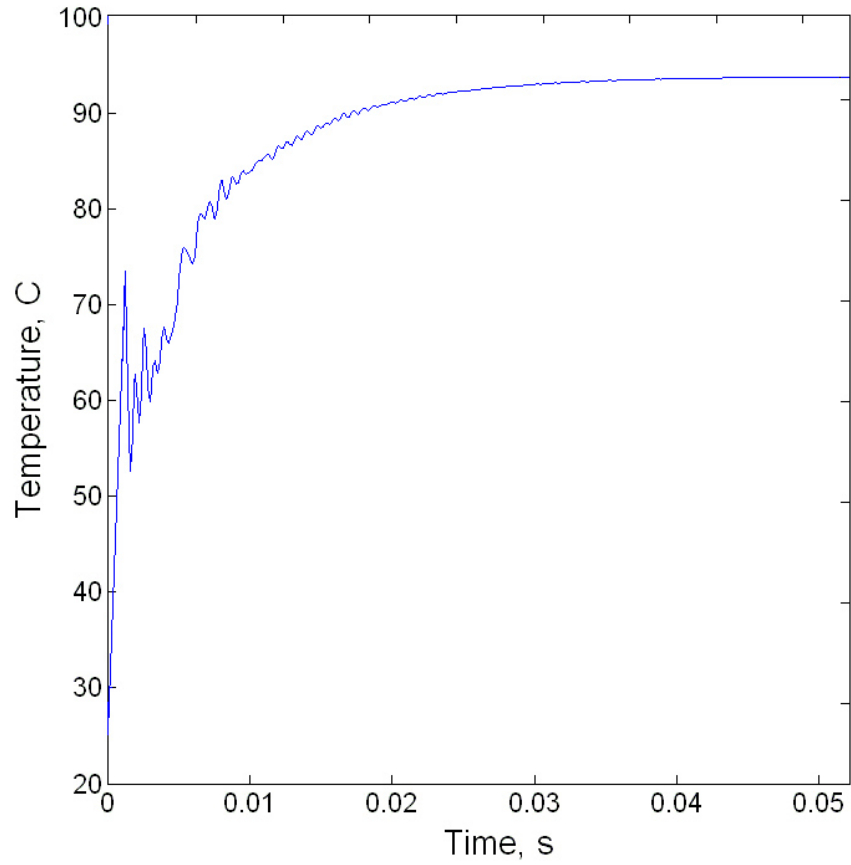


Figure 6.21: Nonphysical transient response of the electro-thermal model

it will result in fast-enough run time. The transient solution in Fig. 6.20 was verified by an HSPICE simulation using a purely RC model.

CHAPTER 7

CONCLUSIONS AND FUTURE DEVELOPMENT

7.1 Conclusions

The overall goal of the latency insertion method (LIM) project is the development of the general purpose circuit simulator based on the LIM algorithm that is superior to the traditional techniques that are in use today. To achieve the objective, we need to develop methodologies that will allow the LIM-based simulator to handle various types of circuit analysis and model a variety of circuit structures.

In this dissertation we extended the LIM method to handle steady state analysis of purely resistive networks with constant power sources. Steady state resistive networks are used to model various phenomena. Among others, DC IR drop analysis of on-chip power distribution networks can be performed using resistive grid models. Although the LIM was initially developed as a purely transient technique, we demonstrated that it could be successfully applied to steady state problems. In fact, in the case of steady state analysis, the issues of latency insertion and conditional stability of the method become less critical. We discovered that the steady state solution did not depend on the fictitious latency values. Therefore, in the absence of actual latency, the time step of the LIM simulation was not limited and the stability of the simulation could always be insured.

In Chapter 6, we proposed the LIM for the complex analysis of integrated circuits and systems at the prelayout stages. We demonstrated that the LIM can be successfully used to address both electrical and thermal aspects of circuit design and potentially combine the two

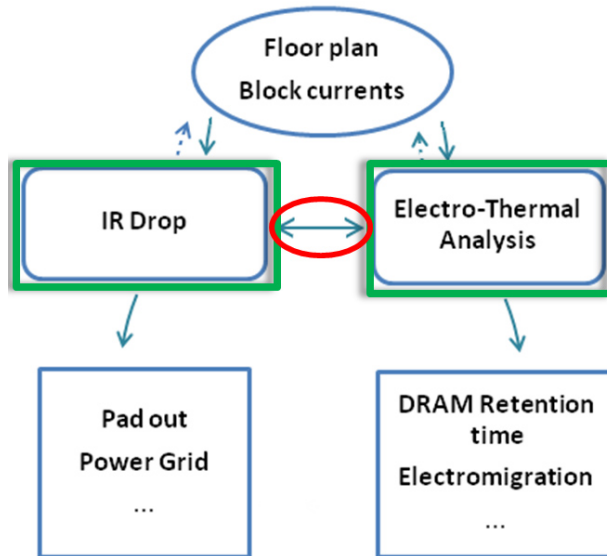


Figure 7.1: Prelayout power integrity analysis flow

and perform complete electro-thermal analysis of on-chip and off-chip interconnect networks as shown in the diagram in Fig. 7.1.

The number of elements in the full electro-thermal model of the integrated circuit can be prohibitively large for traditional circuit simulators. Such a model becomes very large even at the earliest stages of design when quick assessments and repeated simulations with minor adjustments to the model are needed. The LIM demonstrates linear numerical complexity and hence scales well with the size of the model, unlike the traditional matrix-based methods. The LIM provides the capability of performing both steady state and transient analysis within the scope of the single method.

7.2 Further Development

7.2.1 Electrical and Thermal Co-Analysis

As was mentioned earlier, electrical and thermal phenomena are tightly coupled in IC structures. We believe that the LIM can be used to provide complete power and signal

integrity as well as reliability analysis of on-chip and off-chip (package- and board-level) structures. Such analysis should take into account complex electrical, thermal, and potentially even mechanical effects and their interdependence.

Of particular interest is combining IR drop analysis with thermal simulation. At the early design stages, IR drop analysis can be performed based on the information about the floor plan, pad out, and current and voltage budgets for the IC. Results of such analysis can be used to estimate power distribution in the circuit. Power distribution data can then be fed into the thermal simulator that will determine the temperature profile of the structure and identify potential hot spots. Two simulations can be run iteratively with adjustments made to the floor plan of the IC or the structure of the power delivery network.

7.2.2 New Extensions of the LIM

Currently effort is being made to extend the LIM to modeling of nonlinear devices, which will enable transistor-level simulation of digital and analog circuits and systems. In particular, such structures as phase locked loops (PLL), used in clock recovery circuitry, are being modeled with the LIM.

Recently, the LIM was successfully combined with vector-fitting macromodeling techniques, further extending the capabilities of the method. A “black box” model characterized by a set of frequency domain S-parameters can be incorporated into the model network and efficiently analyzed with the LIM.

7.2.3 Enhancement of the Algorithm

The main factor limiting the performance of the LIM transient simulation of complex systems is the requirement for fictitious latency. It is, therefore, crucially important to develop an efficient, robust methodology for determining the size of such fictitious elements. It is desirable to have as much latency in the model network as possible, enabling larger

time step and shorter run time. However, any fictitious latency introduced in the network must be much smaller than any real one present in the model. Inserted latency must be small enough to have no significant impact on the result of the simulation. Essentially, the question of “how small is small enough?” has to be answered.

REFERENCES

- [1] M. Zhao, R. Panda, S. Sapatnekar, and D. Blaauw, “Hierarchical analysis of power distribution networks,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 21, no. 2, pp. 159–168, 2002.
- [2] J. Kozhaya, S. Nassif, and F. Najm, “A multigrid-like technique for power grid analysis,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 21, no. 10, pp. 1148–1160, 2002.
- [3] H. Qian, S. Nassif, and S. Sapatnekar, “Random walks in a supply network,” in *Proceedings of the 40th Annual Design Automation Conference*. ACM, 2003, pp. 93–98.
- [4] Y. Zhong and M. Wong, “Fast algorithms for IR drop analysis in large power grid,” in *ICCAD*. IEEE, 2005, pp. 351–357.
- [5] J. Xie and M. Swaminathan, “DC IR drop solver for large scale 3D power delivery networks,” in *Electrical Performance of Electronic Packaging and Systems (EPEPS), 2010 IEEE 19th Conference on*. IEEE, pp. 217–220.
- [6] J. Schutt-Ainé, “Latency insertion method (LIM) for the fast transient simulation of large networks,” *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, vol. 48, no. 1, pp. 81–89, 2001.
- [7] K. Yee, “Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media,” *Antennas and Propagation, IEEE Transactions on*, vol. 14, no. 3, pp. 302–307, 1966.
- [8] R. Booton, *Computational Methods for Electromagnetics and Microwaves*, vol. 1. New York, NY: Wiley-Interscience, 1992.
- [9] S. Lalgudi, M. Swaminathan, and Y. Kretchmer, “On-chip power-grid simulation using latency insertion method,” *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 55, no. 3, pp. 914–931, 2008.
- [10] A. Taflove and S. Hagness, *Computational Electrodynamics*. Boston, MA: Artech House, 2005.
- [11] Z. Deng and J. Schutt-Aine, “Stability analysis of latency insertion method (LIM),” in *Electrical Performance of Electronic Packaging, 2004 IEEE 13th Topical Meeting on*. IEEE, 2004, pp. 167–170.

- [12] R. Courant, K. Friedrichs, and H. Lewy, “Über die partiellen Differenzgleichungen der mathematischen Physik,” *Mathematische Annalen*, vol. 100, no. 1, pp. 32–74, 1928.
- [13] J. Schutt-Ainé, “Stability analysis for semi-implicit LIM algorithm,” in *Microwave Conference, 2009. APMC 2009. Asia Pacific*. IEEE, pp. 1270–1272.
- [14] P. Goh, J. Schutt-Ainé, D. Klokov, J. Tan, P. Liu, W. Dai, and F. Al-Hawari, “Partitioned latency insertion method with a generalized stability criteria,” *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 1, no. 9, pp. 1447–1455, 2011.
- [15] S. Lalgudi and M. Swaminathan, “Analytical stability condition of the latency insertion method for nonuniform GLC circuits,” *Circuits and Systems II: Express Briefs, IEEE Transactions on*, vol. 55, no. 9, pp. 937–941, 2008.
- [16] Z. Deng, J. Schutt-Ainé, J. Tan, C. Kumar, and D. Milosevic, “Latency insertion method with frequency dependent effect,” in *Electronics Systemintegration Technology Conference, 2006, 1st*, vol. 1. IEEE, pp. 503–506.
- [17] D. Klokov and J. Schutt-Ainé, “Transient simulation of lossy interconnects using the latency insertion method (LIM),” in *Electrical Performance of Electronic Packaging, 2008 IEEE-EPEP*. IEEE, pp. 251–254.
- [18] M. Gowan, L. Biro, and D. Jackson, “Power considerations in the design of the alpha 21264 microprocessor,” in *Proceedings of the 35th Annual Design Automation Conference*. ACM, 1998, pp. 726–731.
- [19] P. McCrorie. (2006, June 26). Using dynamic and static power rail analysis to maximize results with minimum effort. *SOCentral*. [Online]. Available: <http://www.socentral.com/results.asp?EntryID=19453>.
- [20] D. Klokov and J. Schutt-Ainé, “Extensions of the latency insertion method (LIM) to DC analysis of power supply networks and modeling of circuit interconnects with frequency-dependent parameters,” in *Electronic Components and Technology Conference, 2009. ECTC 2009. 59th*. IEEE, pp. 1624–1629.
- [21] D. Klokov, P. Goh, and J. Schutt-Ainé, “Latency insertion method (LIM) for DC analysis of power supply networks,” *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, to be published.
- [22] J. Lee, K. Kim, Y. Huh, P. Bendix, and S. Kang, “Chip-level charged-device modeling and simulation in CMOS integrated circuits,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 22, no. 1, pp. 67–81, 2003.
- [23] *ADI Reliability Handbook*. Norwood, MA: Analog Devices, 2000.

- [24] M. Sowariraj, T. Smedes, C. Salm, T. Mouthaan, and F. Kuper, “Role of package parasitics and substrate resistance on the charged device model (CDM) failure levels - an explanation and die protection strategy,” *Microelectronics Reliability*, vol. 43, no. 9-11, pp. 1569–1575, 2003.
- [25] V. Shukla, N. Jack, and E. Rosenbaum, “Predictive simulation of CDM events to study effects of package, substrate resistivity and placement of ESD protection circuits on reliability of integrated circuits,” in *Reliability Physics Symposium (IRPS), 2010 IEEE International*. IEEE, 2010, pp. 485–493.
- [26] D. Klokov, V. Shukla, J. Schutt-Aine, and E. Rosenbaum, “Application of the latency insertion method (LIM) to the modeling of CDM ESD events,” in *Electronic Components and Technology Conference (ECTC), 2010 Proceedings 60th*. IEEE, pp. 652–656.
- [27] T. Sekine and H. Asai, “CMOS circuit simulation using latency insertion method,” in *Electrical Performance of Electronic Packaging, 2008 IEEE-EPEP*. IEEE, pp. 55–58.
- [28] A. Allan, D. Edenfeld, W. Joyner, A. Kahng, M. Rodgers, and Y. Zorian, “International technology roadmap for semiconductors,” *IEEE Computer Society*, vol. 35, pp. 42–53, 2002.
- [29] P. Li, L. Pileggi, M. Asheghi, and R. Chandra, “IC thermal simulation and modeling via efficient multigrid-based approaches,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 25, no. 9, pp. 1763–1776, 2006.
- [30] Y. Cheng, P. Raha, C. Teng, E. Rosenbaum, and S. Kang, “ILLIADS-T: An electrothermal timing simulator for temperature-sensitive reliability diagnosis of CMOS VLSI chips,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 17, no. 8, pp. 668–681, 1998.
- [31] T. Chiang, K. Banerjee, and K. Saraswat, “Compact modeling and spice-based simulation for electrothermal analysis of multilevel ULSI interconnects,” in *Proceedings of IEEE/ACM International Conference on Computer Aided Design*, 2001, pp. 165–172.
- [32] Y. Cheng, C. Tsai, C. Teng, and S. Kang, *Electrothermal Analysis of VLSI Systems*, Boston, MA: Kluwer Academic Publishers, 2000.
- [33] L. He, W. Liao, and M. Stan, “System level leakage reduction considering the interdependence of temperature and leakage,” in *Proceedings of the 41st annual Design Automation Conference*. ACM, 2004, pp. 12–17.
- [34] Z. Yu, D. Yergeau, R. Dutton, S. Nakagawa, and J. Deeney, “Fast placement-dependent full chip thermal simulation,” in *VLSI Technology, Systems, and Applications, 2001. Proceedings of Technical Papers, 2001 International Symposium on*. IEEE, 2001, pp. 249–252.

- [35] T. Wang and C. Chen, “3-D thermal-ADI: A linear-time chip level transient thermal simulator,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 21, no. 12, pp. 1434–1445, 2002.
- [36] H. Su, F. Liu, A. Devgan, E. Acar, and S. Nassif, “Full chip leakage estimation considering power supply and temperature variations,” in *Proceedings of the 2003 International symposium on Low Power Electronics and Design*. ACM, 2003, pp. 78–83.
- [37] C. Tsai and S. Kang, “Cell-level placement for improving substrate thermal distribution,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 19, no. 2, pp. 253–266, 2000.
- [38] B. Goplen and S. Sapatnekar, “Efficient thermal placement of standard cells in 3D ICs using a force directed approach,” in *Proceedings of the 2003 IEEE/ACM International Conference on Computer-Aided Design*. IEEE Computer Society, 2003, p. 86.
- [39] K. Skadron, M. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan, “Temperature-aware microarchitecture,” in *ACM SIGARCH Computer Architecture News*, vol. 31, no. 2. ACM, 2003, pp. 2–13.
- [40] S. Gunther, F. Binns, D. Carmean, and J. Hall, “Managing the impact of increasing microprocessor power consumption,” *Intel Technology Journal*, vol. 5, no. 1, pp. 1–9, 2001.
- [41] Y. Zhong and M. Wong, “Thermal-aware IR drop analysis in large power grid,” in *Quality Electronic Design, 2008. ISQED 2008. 9th International Symposium on*. IEEE, 2008, pp. 194–199.
- [42] C. Teng, Y. Cheng, E. Rosenbaum, and S. Kang, “iTEM: A temperature-dependent electromigration reliability diagnosis tool,” *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 16, no. 8, pp. 882–893, 1997.
- [43] G. Davies, R. Fenner, and R. Lewis. (1992, June 1). NAFEMS background to benchmarks. *NAFEMS*. [Online]. Available: http://www.nafems.org/publications/browse_buy/benchmark/r0006.