

UNIVERSITY OF CALIFORNIA  
SANTA CRUZ

**Transient Analysis of Coupled Transmission Lines Characterized with  
Frequency-Dependent Losses or Measured Scattering-Parameter Data  
and Optimal Design of Self-Damped Interconnects**

A dissertation submitted in partial satisfaction  
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER ENGINEERING

by

Jimmy Shinn-Hwa Wang

June 1995

The dissertation of Jimmy Shinn-Hwa Wang  
is approved:

---

Wayne Wei-Ming Dai

---

Patrick Mantey

---

Peter Saviz

---

Dean of Graduate Studies and Research

Copyright © by  
Jimmy Shinn-Hwa Wang  
1995

# Contents

<b>Abstract</b>	<b>viii</b>
<b>Acknowledgments</b>	<b>x</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. Transient Analysis of Coupled Transmission Lines</b>	<b>6</b>
2.1 Congruence Transformation of the Coupled Transmission Lines . . . . .	7
2.2 Scattering Parameter Macromodel of a Congruence Transformer . . . . .	9
2.3 Scattering Parameter Macromodel of a Decoupled Transmission Line . . . .	14
2.4 Experimental Result of the Transient Analysis of Coupled Transmission Lines	15
<b>3. Transient Analysis of Coupled Transmission Lines Characterized with the Frequency-Dependent Losses</b>	<b>18</b>
3.1 Representing the Frequency-Dependent Coupled Transmission Lines Using the S-Parameter Macromodel . . . . .	19
3.2 S-Parameter Macromodel of the Frequency-Dependent Single Transmission Line . . . . .	20
3.3 S-Parameter Macromodel of the Frequency-Dependent Coupled Transmission Line . . . . .	25
3.4 Experimental Results for Frequency-Dependent Transmission Lines . . . . .	30
3.4.1 Example 1: A Single Transmission Line . . . . .	31
3.4.2 Example 2: Two Coupled Transmission Lines . . . . .	31
3.4.3 Example 3: Two Cascade Sections of Coupled Transmission Lines . .	33
3.4.4 Example 4: Three Coupled Transmission Lines . . . . .	36

<b>4. Transient Analysis of Interconnects Characterized with Measured S-Parameter Data</b>	<b>40</b>
4.1 Motivations . . . . .	42
4.2 Proposed Method . . . . .	47
4.3 Converting Measured S-Parameter Data into Parasitic Functions . . . . .	47
4.4 Finding the Moments of the Four Parasitic Functions . . . . .	51
4.5 Creation of the S-Parameter Macromodel . . . . .	52
4.6 Experimental Results for Measured Scattering Parameter Data . . . . .	52
4.6.1 Example 1: A Step Structure . . . . .	53
4.6.2 Example 2: An Uniform Microstrip Structure . . . . .	55
4.6.3 Example 3: A Bend Structure . . . . .	55
<b>5. Optimal Design of Self-Damped Interconnects for Multichip Modules</b>	<b>57</b>
5.1 Wire-Sizing . . . . .	58
5.2 Performance Criteria . . . . .	59
5.3 Levenberg-Marquardt Method . . . . .	59
5.4 Previous Work . . . . .	60
5.5 S-Parameter Macromodel and the Time-of-Flight Extraction . . . . .	61
5.6 Transfer Function Approximation . . . . .	62
5.7 Formulation of the Optimal Self-Damped Design . . . . .	65
5.8 Proof of the Correctness of the Least-Squares Estimation Transformation . . . . .	67
5.9 Implementation of the Least-Square Estimation Optimization . . . . .	69
5.10 Experimental Results for the Optimal Self-Damped Design . . . . .	70
5.10.1 Example 1 A Tree Network . . . . .	70
5.10.2 Example 2 A Network with a Loop . . . . .	71
5.10.3 Example 3 A Clock Tree Network . . . . .	73

<b>6. Transformation of Min-Max Optimization to Least-Square Estimation and Its Application to Interconnect Design Optimization</b>	<b>75</b>
6.1 Transform the Min-Max Problem to One Least-Square Estimation Problem	77
6.1.1 Physical Example . . . . .	77
6.1.2 Problem Definition . . . . .	78
6.1.3 Transformation Formulation . . . . .	79
6.1.4 Physical Meaning . . . . .	80
6.1.5 Practical Considerations . . . . .	82
6.2 Novel Transformation of the Min-Max Problem into Two Least-Square Estimation Problems . . . . .	83
6.2.1 Problem Definition . . . . .	84
6.2.2 Transformation Formulation . . . . .	84
6.2.3 Physical Meaning . . . . .	87
6.2.4 Practical Considerations . . . . .	89
6.3 Implementation of the Least-Square Estimation Optimization . . . . .	90
6.4 Experimental Results for Better Optimal Self-Damped Design . . . . .	90
6.4.1 Example 1 A Tree Network . . . . .	93
6.4.2 Example 2 A Network with a Loop . . . . .	93
6.4.3 Example 3 A Clock Tree Network . . . . .	94
<b>7. Conclusions</b>	<b>97</b>
<b>8. Future Work</b>	<b>99</b>
<b>A. Proof of the Tighter Bound</b>	<b>100</b>
<b>References</b>	<b>105</b>

## List of Figures

2.1	Coupled Transmission Lines . . . . .	8
2.2	The Separation of Wave Components . . . . .	12
2.3	Circuit Schematic of the Two Groups of Three Coupled Lossless Transmission Lines . . . . .	16
2.4	Enlarged Simulation Waveforms of the Two Groups of Three Coupled Lossless Transmission Lines . . . . .	16
2.5	Simulation Waveforms of the Two Groups of Three Coupled Lossless Transmission Lines . . . . .	17
3.1	Simulation Waveforms of the Single Transmission Line Characterized with Frequency-Dependent Losses . . . . .	32
3.2	Simulation Waveforms of the Coupled Transmission Line Characterized with and without Frequency-Dependent Losses . . . . .	34
3.3	Schematic of the Coupled Transmission Line Characterized with Frequency-Dependent Losses . . . . .	36
3.4	Simulation Waveforms of the Coupled Transmission Line Characterized with Frequency-Dependent Losses . . . . .	37
3.5	Simulation Waveforms of the Coupled Transmission Line Characterized with Frequency-Dependent Losses . . . . .	39
4.1	A Particular Curve-Fitting Example . . . . .	43
4.2	AWEsim and Rational function Curve-Fitting Comparison . . . . .	43
4.3	Indirect Curve-Fitting Example . . . . .	44
4.4	The Introduced Conversion Error . . . . .	46
4.5	Test Circuit Schematic . . . . .	53
4.6	Simulation Waveforms of the Step Structure Characterized with Measured S-Parameter Data . . . . .	54

4.7	Enlarged Far End Simulation Waveforms . . . . .	54
4.8	Simulation Waveforms of the Uniform Microstrip Characterized with Measured S-Parameter Data . . . . .	55
4.9	Simulation Waveforms of the Bend Structure Characterized with Measured S-Parameter Data . . . . .	56
5.1	Simulation Waveforms of Different Line Widths . . . . .	57
5.2	Root Locus of a Second-Order System . . . . .	63
5.3	A Tree Network and the Simulation Waveforms of the Receiver 3 . . . . .	72
5.4	A Network with a Loop and the Simulation Waveforms of the Receiver 4 . .	72
5.5	A Clock Tree Network and its Simulation Waveforms . . . . .	74
6.1	A Physical Design Example . . . . .	78
6.2	A Positive Function . . . . .	81
6.3	Relationship between the Maximum, Minimum, Average and RMS Values .	81
6.4	A Positive Function . . . . .	88
6.5	Relationship between the Maximum, Minimum, Average and RMS Values .	88
6.6	Relationship between the Values of the New Linear Bound . . . . .	89
6.7	A Tree Network and the Simulation Waveforms of the Receiver 3 . . . . .	92
6.8	A Network with a Loop and the Simulation Waveforms of the Receiver 4 . .	93
6.9	A Clock Tree Network and its Simulation Waveforms . . . . .	95

# **Transient Analysis of Coupled Transmission Lines Characterized with Frequency-Dependent Losses or Measured Scattering-Parameter Data and Optimal Design of Self-Damped Interconnects**

*Jimmy Shinn-Hwa Wang*

## **ABSTRACT**

As the packaging density and the clock speed of the multichip modules (MCMs) increase, crosstalk induced noise can become one of the major sources of noise which will limit the performance of the high-speed digital systems. In addition, as the magnitudes of the harmonics of the transmitted signals above 1 GHz become more significant, the interconnects exhibit more frequency-dependent conductor losses (skin effect) and the dielectric losses (dispersion). In order to design the interconnects more accurately and more efficiently, the S-parameters macromodels of the coupled lossless and lossy transmission lines with frequency-dependent losses are constructed in this thesis. Since the discontinuities in interconnects are hard to describe with close form equations, they are better characterized with measured S-parameter data. Thus, it is equally important to include the measured S-parameter macromodel in the transient analysis.

The task of designing interconnections does not stop at viewing simulation waveforms, it requires the design optimization as well. For today's Computer-Aided Design tools, it usually takes several iterations between the layout and the simulation tools to find the optimal design of the interconnects. A simple and robust method of designing the lossy-transmission-line interconnects in a network for multichip modules has been developed. It uses wire-sizing to solve the problems encountered in propagating high-speed signals through unterminated lossy transmission lines on the substrates of multichip modules. The optimal self-damped design concept is used to relate the layout parameter (line width) and the transfer function (damping ratio, and natural undamped frequency) to the signal propagation delay. This method can produce a fast and a more stable signal propagation for



single-source multi-receiver networks on multichip modules, without using termination resistors. There are further improvements that can be gained by incorporating an incremental simulation and a sensitivity analysis into the optimal design to improve the efficiency.

Building these scattering parameter macromodels greatly improves the accuracy and the efficiency of the transient analysis. Together with the optimal design, they provides a faster turn around time in producing high-performance MCM designs.

**Keywords:** Coupled Transmission Lines, Transient Analysis, Congruence Transformation, Scattering Parameter, Macromodel, S-Parameter Based Macromodel Simulator, Measured Data, Frequency-Dependent Losses, S-Parameter, Measured S-Parameter Data, Frequency-Dependent Parasitic, Optimization, Min-Max, Least-Square Estimation, Linear Bound, Root-Mean-Square

## Acknowledgments

I want to thank my thesis advisor, Dr. Wayne Dai, whose advise and guidance have been essential to my work. I am most grateful to him for a privilege to participate in his vision to pioneer the research work of Signal Integrity before Integrated Circuit designer have a chance to realize its potential. I would like to thank Dr. Patrick Mantey for his invaluable suggestions based upon his indepth knowledge on the optimal design of control systems, and to thank Dr. Peter Saviz for his inspirations on the subjects of circuit simulation. I would like to thank Dr. J. C. Liao of Intel for both his inspirations and his assistance during the course of this research. I would like to thank Dr. Norman H. C. Chang of HP for his assistance. I am deeply indebted to the fellow students in the Scattering-Parameter Macromodel Based Simulator project, especially Haifang Liao. I am also indebted to the excellent faculty members at U.C. Santa Cruz, especially Dr. Martine Schlag for her kind assistance. I would like to thank Dr. Anthony Wu for painstakingly proofreading all of my papers, for whom I am eternally grateful. Finally, I want to thank my wife, Wendy Lin; without her encouragement and financial support, I will never be able to complete this degree, for which I am forever indebted.

# CHAPTER 1. Introduction

Multichip modules (MCMs) are one of the most recent advances in the high-performance packaging technology. They are usually constructed with multiple bare dice mounted and interconnected on a thin-film substrate. As the packaging density and the clock speed of the MCMs increases, crosstalk-induced noise, as well as some of the frequency-dependent losses, have become some of the major issues in the designing of the interconnect networks. As a result, more accurate and more efficient circuit simulators are necessary for the building of advanced MCMs.

The scattering-parameter (S-parameter) macromodel based simulator has been proposed as a novel circuit simulator [40] [38]. Given the scattering parameter description of the lumped elements, interconnect junctions, and single transmission lines, and using two efficient reduction rules on the original distributed and lumped network can be reduced to a network containing one multi-port component together with the sources and the loads of interest [40]. In addition, the trade-off between accuracy and efficiency of the simulation can be controlled by choosing the order of approximation. The utility of the Scattering-Parameter Macromodel Based Simulator, however, is very limited due to the small number of macromodels that are available, thus far, due primarily to the relatively short course of its existence. It is therefore important for us to pursue other macromodels that can deal with the crosstalk noise and the frequency-dependent issues. This thesis offers some newly developed macromodels that add extensions and enhancements to the existing macromodel based simulator. More specifically, this thesis includes four additional macromodels that will provide solutions to the previously mentioned problems: namely the crosstalk among multi-conductors and the frequency dependence of the parasitics.

As the system clock speed increases, the crosstalk becomes one of the major source of noise, in addition to the delay and ringing which can limit the performance of high-speed digital systems [7] [9] [12] [14] [19] [46] [26] [48] [57]. The crosstalk can often lead to excessive overshoots, undershoots and glitches. It can also cause false switchings on the non-active lines as well as undetected switchings on the active lines, not to mention that it can also increase power dissipation of the output drivers. The coupled noise (crosstalk) is inversely proportional to the inter-line spacing and is directly proportional to several parameters including those of the thickness of the dielectric material, the distance which the coupled lines are in parallel, the rate of change of the input waveforms, as well as the line impedances.

Due to the high switching speed of today's digital systems, the magnitudes of the harmonics of the transmitted signals above 1 GHz can often become the significant components in the power spectrum [29] [48] [2] [14] [28] [44]. At these high frequencies, the interconnects exhibit more frequency-dependent behaviour resulting from conductor losses (skin effect) and dielectric losses (dispersion) [44]. The electro-static field solution can no longer predict the correct parasitics for the interconnects that carry the high-speed digital signals. The macromodel of the transmission lines is required to accurately model both the frequency-dependent conductor (skin effect) and the dielectric losses.

The correct modeling of the transmission lines requires accurate modeling of both the frequency-dependent conductor (skin effect) and the dielectric losses. The measured S-parameter data can easily capture the behavior of these losses. The discontinuities in interconnects are difficult to describe with closed form equations. They are better characterized with the measured S-parameter data [21] [45] [29] [47] [34]. This calls for a more accurate and efficient circuit simulator which can accept devices characterized with measured S-parameter data.

Typical approaches for the transient analysis of the coupled lossy transmission lines are adapted in the new scattering-parameter macromodel based simulator. This thesis extends the scattering-parameter macromodel to include the congruence transformation to

analyze the crosstalk in coupled transmission-line systems. It also extends the scattering-parameter macromodel to include the frequency-dependent losses (such as skin effect) and coupling in the transient analysis of the coupled transmission-line systems. Scattering-parameter macromodel for interconnects characterized with measured data is also included. An indirect method is developed to find the moments from the measured S-parameter data of interconnects. The novel indirect method of computing the moments calls for a conversion of the S-parameter data into the parasitic functions before finding the Taylor series expansions. The exact moments of the S-parameter functions are subsequently computed from the moments of these four parasitic functions. Based on the exact moments found using the indirect approach, the macromodel of transmission lines characterized with frequency-dependent losses is constructed and the transient simulation is performed.

The task of designing the high-performance system does not stop with viewing the simulation waveforms; this research takes one step further to optimally design the interconnects based on the simulation waveforms. Based on the Integral of Time and Error criterion (ITAE), wire-sizing method, and Least-Square Estimation Optimization, the lossy interconnects on the multichip modules can be designed to meet all the user's electrical requirements. A simple and robust optimal design method is developed for the design of the lossy-transmission-line interconnects for multichip modules [52] [53] [54]. This method uses wire-sizing to meet the electrical damping criteria, to solve the problems encountered in propagating high-speed signals through unterminated lossy transmission lines on the substrates of MCMs. The optimal design method is based on a new improved scattering-parameter macromodel of transmission lines that keeps track of the time-of-flight term in the transfer functions. The optimal self-damped design concept is to relate the layout parameter (line width and line spacing) and the transfer function (damping ratio, and natural undamped frequency) to the signal propagation delay. This method produces a faster and a more stable signal propagation for single-source multi-receiver networks on MCMs without using termination resistors.

The optimal wire-sizing problem is a general nonlinear programming problem. Prin-

Principal Least Square Algorithms include the Durbin method, the Levinson/Wiener method, the Cholesky Decomposition method, the Burg method, the Lattice method, the Kalman method, and the Square Root Kalman method [27]. The nonlinear programming problem is not only hard to solve but also takes significant computation time. One alternative to solve the Min-Max optimization problem, which obtains a good solution in a relative short period of time, is to transform the original Min-Max optimization problem into a Least-Square Estimation (LSE) problem, which has a number of well defined methods for finding its optimal solution. It can be shown that the target function of the original Min-Max optimization problem and that of the transformed LSE problem linearly bound each other. Although these two problems have different objective functions, optimizing the transformed problem can produce a solution to the original Min-Max problem. Thus for non-linear Least Square Estimation problems, the Levenberg-Marquardt method is recommended [15] [36] [16].

The Levenberg-Marquardt Method is used in single variable optimization by transforming the original Min-Max problem into the Least-Square Estimation (LSE) problem. To use it in the nonlinear programming problem such that multiple objectives must be simultaneously minimized together, one must change the transformation. An new approach has been developed in this thesis. By transforming the original Min-Max problem into two Least-Square Estimation problems, we prove that a tighter bound exists which guarantee the optimization of the transformed multi-objective LSE problems is equivalent to the optimization of the original Min-Max problem.

This thesis is divided largely into two parts: The Part I deals with the discussion of the proposed macromodels. Part II is dedicated exclusively to the optimal design method. Including in Part I are macromodels for:

- The coupled transmission lines (Chapter 2)
- The single and coupled transmission lines characterized with frequency-dependent losses (Chapter 3)
- The elements characterized with measured S-parameter data (Chapter 4)

Including in Part II are the optimal design method and the proof of the tighter bound for the optimization transformation. The optimal design method is shown in Chapter 5 and the proof of the tighter bound is shown in Chapter 6. The theoretical work for all the macromodels, the optimal design method, and the proof of the tighter bound have been completed and are presented in the thesis.

## CHAPTER 2. Transient Analysis of Coupled Transmission Lines

Previous research work which uses scattering parameter for the analysis of the coupled transmission lines include those of the general convolution by Winklestein et al. [57], the full-wave analysis by Cooke et al. [14] and time domain simulation by Schutt-Aine et al. [48]. While Schutt-Aine et al. demonstrated a great accuracy improvements in simulating circuits that includes non-linear drivers and terminations, Cooke et al. illustrated an ability to simulate frequency dependent model propagation. Recently, a novel frequency domain simulator using scattering-parameter macromodels has been presented by Liao et al [39] [40]. Based on the scattering-parameter macromodel, Pade techniques or Exponentially Decayed Polynomial Function (EDPF) can be used to approximate transfer functions of the coupled interconnects. This approach avoids the costly matrix computation for converting the frequency domain scattering-parameter matrix representation into the one of time domain transfer/reflection matrix as in [48], and the time consuming full-wave analysis as in [14].

It has been shown that the congruence transformation can be used to decouple the modal wave propagation in a coupled transmission-line system [7] [9] [30] [46]. After decoupling the  $n$  coupled transmission lines using the congruence transformation, the task of computing the scattering matrix of the coupled transmission lines then becomes the task of computing the scattering matrix of the congruence transformers and that of the  $n$  decoupled single transmission lines. This congruence transformation has been extended to build the novel macromodel for the coupled lossy transmission lines. A very simple scattering-parameter description of the congruence transformer for coupled lossless transmission lines is subsequently derived. Incorporating the novel macromodel of the coupled transmission



lines in the scattering parameter based simulator provides a more efficient way of analyzing crosstalk for the transient simulation.

The contribution of this portion of the thesis is to build the S-parameter macromodel for the congruence transformer. In Section 2.1, the congruence transformation of the lossless and lossy coupled transmission lines is derived based on frequency-independent per-unit-length  $\mathbf{L}$  and  $\mathbf{C}$  matrices. The scattering-parameter macromodel of the coupled lossless transmission lines is derived from the congruence transformation. An added simplification of the scattering-parameter macromodel is derived through the similarity transformation property. The derivation will be further extended to the coupled lossy transmission lines. In Section 2.4, an example of one of the MCMC-94 Benchmarks will be presented. The results will be compared with the previously published results and the state-of-the-art simulators such as those of SWEC [41] and Coffee [11].

## 2.1 Congruence Transformation of the Coupled Transmission Lines

Wave propagation in multiconductors has been extensively studied by the microwave, electronic magnetic compatibility (EMC), and electrical engineers. Due to the coupling between transmission lines, different modes which have different propagation velocities exist simultaneously in the system. For an  $n$  conductor system shown in Figure 2.1 (a), there exists  $n$  fundamental modes of propagation.

With the assumption of quasi-TEM wave propagation, the distributions of voltages and currents in an  $n$  coupled lossy transmission-line system can be described by the generalized Telegraphist's equations [9]:

$$\frac{\partial \mathbf{v}(x, t)}{\partial x} = -\mathbf{L} \frac{\partial \mathbf{i}(x, t)}{\partial t} - \mathbf{R} \mathbf{i}(x, t) \quad (2.1)$$

$$\frac{\partial \mathbf{i}(x, t)}{\partial x} = -\mathbf{C} \frac{\partial \mathbf{v}(x, t)}{\partial t}, \quad (2.2)$$

where  $0 \leq x \leq l$ , and  $\mathbf{v}(x, t)$  and  $\mathbf{i}(x, t)$  are column vectors defining the voltages  $v_k(x, t)$  and currents  $i_k(x, t)$  distributed on the conductors  $k = 1, 2, 3, \dots, n$ . The  $\mathbf{L}$  and  $\mathbf{C}$  are the  $n$  by  $n$  symmetric matrices of the per-unit-length inductance and capacitance of the  $n$



where  $R_{kk}$  is the  $k$ -th diagonal element of the  $\mathbf{R}$  matrix,  $W$  is the right eigenvector matrix of the time constant matrix  $\mathbf{T} \equiv \mathbf{R}^{-1/2}\mathbf{L}\mathbf{R}^{-1/2} = W \text{diag}(\tau_k)W^{-1}$ ,  $\tau_k$ s are the eigenvalues of time constant matrix  $\mathbf{T}$ , and  $L_k$ s are some arbitrary constants.

Bayard first outlines the transformation,  $\mathbf{A}^t\mathbf{Z}\mathbf{A}$ , and calls it “translator” [3]. Hazony is the first one to name the transformation “congruence transformer” in his book [30]. Chang uses the congruence transformer to decouple both the lossless [7] and lossy coupled transmission lines [9]. Chang’s method for the analysis of coupled transmission lines relies on simultaneously diagonalizing all the matrices using a special conditioned matrix. Romeo and Santomauro present a different method of finding the right eigenvector matrices for coupled lossless transmission lines with tridiagonal  $\mathbf{L}$  and  $\mathbf{C}$  matrices [46]. In this part of the thesis, we find the right eigenvector matrices of a full  $\mathbf{L}$  matrix. The method proposed here for the lossless case is preferred to Chang’s method [7] because it leads to simpler equations for scattering parameter based macromodel representation of congruence transformer. In Chang’s method [7], where the similarity transformation property  $X^{-1} = X^t$  may not hold, the scattering parameter matrix representation in Equation (2.39) for the coupled lossless case cannot therefore be simplified to Equation (2.40).

After the coupled transmission-line system is decoupled into the congruence transformers and the decoupled single transmission lines, the macromodels can be built for all of them from their individual parameters.

## 2.2 Scattering Parameter Macromodel of a Congruence Transformer

By applying the congruence transformation, we can change the variable basis from  $v$  to  $u$  and from  $i$  to  $j$ . The terminal voltages and currents at opposite sides of the transformer are related by (see Figure 2.1 (b)) [7]:

$$v_k(x, t) = \sum_{m=1}^n X_{km} u_m(x, t) \quad (2.6)$$

$$j_k(x, t) = - \sum_{m=1}^n X_{mk} i_m(x, t), \quad (2.7)$$

The negative sign is used to indicate the direction of current  $j_k$  is flowing into the transformer. Rewriting Equation (2.6) and (2.7) in a matrix notation, one obtains:

$$V(x, t) = XU(x, t) \quad (2.8)$$

$$I(x, t) = -\left(X^t\right)^{-1} J(x, t), \quad (2.9)$$

where  $V$ ,  $U$ ,  $I$  and  $J$  are column vectors,  $V = [v_1(x, t), \dots, v_n(x, t)]^t$ ,  $U = [u_1(x, t), \dots, u_n(x, t)]^t$ ,  $I = [i_1(x, t), \dots, i_n(x, t)]^t$ , and  $J = [j_1(x, t), \dots, j_n(x, t)]^t$ .

Substituting Equation (2.8) and (2.9) into Equation (2.1), (2.2), (2.3) and (2.4), after rearrangements, one obtains the following Equations (2.10) to (2.13):

$$\frac{\partial \mathbf{u}(x, t)}{\partial x} = -X^{-1} \mathbf{L} \left(X^t\right)^{-1} \left[-\frac{\partial \mathbf{j}(x, t)}{\partial t}\right] - X^{-1} \mathbf{R} \left(X^t\right)^{-1} [-\mathbf{j}(x, t)] \quad (2.10)$$

$$-\frac{\partial \mathbf{j}(x, t)}{\partial x} = -X^t \mathbf{C} X \frac{\partial \mathbf{u}(x, t)}{\partial t} \quad (2.11)$$

$$\frac{\partial^2 \mathbf{u}(x, t)}{\partial x^2} = X^{-1} \mathbf{L} \mathbf{C} X \frac{\partial^2 \mathbf{u}(x, t)}{\partial t^2} + X^{-1} \mathbf{R} \mathbf{C} X \frac{\partial \mathbf{u}(x, t)}{\partial t} \quad (2.12)$$

$$\frac{\partial^2 \mathbf{j}(x, t)}{\partial x^2} = X^t \mathbf{C} \mathbf{L} \left(X^t\right)^{-1} \frac{\partial^2 \mathbf{j}(x, t)}{\partial t^2} + X^t \mathbf{C} \mathbf{R} \left(X^t\right)^{-1} \frac{\partial \mathbf{j}(x, t)}{\partial t}. \quad (2.13)$$

It can be shown that the coefficient matrices are all diagonal matrices as represented by Equations (2.14) to (2.20):

$$\tilde{\mathbf{R}} = X^{-1} \mathbf{R} (X^t)^{-1} = \text{diag}(R_k) = \text{diag}(L_k / \tau_k) \quad (2.14)$$

$$\tilde{\mathbf{L}} = X^{-1} \mathbf{L} (X^t)^{-1} = \text{diag}(L_k) \quad (2.15)$$

$$\tilde{\mathbf{C}} = X^t \mathbf{C} X = \text{diag}(C_k) = \text{diag}(1/\nu^2 L_k) \quad (2.16)$$

$$\tilde{\mathbf{L}} \tilde{\mathbf{C}} = X^{-1} \mathbf{L} \mathbf{C} X = X^{-1} \mathbf{L} \left(X^t\right)^{-1} X^t \mathbf{C} X = \text{diag}(L C_k) = \text{diag}(1/\nu^2) \quad (2.17)$$

$$\tilde{\mathbf{C}} \tilde{\mathbf{L}} = X^t \mathbf{C} \mathbf{L} \left(X^t\right)^{-1} = X^t \mathbf{C} X X^{-1} \mathbf{L} \left(X^t\right)^{-1} = \text{diag}(C L_k) = \text{diag}(1/\nu^2) \quad (2.18)$$

$$\tilde{\mathbf{R}} \tilde{\mathbf{C}} = X^{-1} \mathbf{R} \mathbf{C} X = X^{-1} \mathbf{R} \left(X^t\right)^{-1} X^t \mathbf{C} X = \text{diag}(R C_k) \quad (2.19)$$

$$\tilde{\mathbf{C}} \tilde{\mathbf{R}} = X^t \mathbf{C} \mathbf{R} \left(X^t\right)^{-1} = X^t \mathbf{C} X X^{-1} \mathbf{R} \left(X^t\right)^{-1} = \text{diag}(C R_k). \quad (2.20)$$

where  $k = 1 \dots n$  and  $\text{diag}(L_k)$  represents an  $n$  by  $n$  diagonal matrix  $L$  with all off-diagonal elements equal to zero. The  $R_k$ ,  $L_k$ ,  $C_k$ ,  $(LC)_k$ ,  $(CL)_k$ ,  $(RC)_k$ , and  $(CR)_k$  are the  $k$ -th eigenvalue of the matrices  $\mathbf{R}$ ,  $\mathbf{L}$ ,  $\mathbf{C}$ ,  $\mathbf{LC}$ ,  $\mathbf{CL}$ ,  $\mathbf{RC}$ , and  $\mathbf{CR}$ .

Redefined the Telegraphist Equations using Equations as shown in (2.21) to (2.24) are:

$$\frac{\partial \mathbf{u}(x, t)}{\partial x} \equiv -\tilde{\mathbf{L}} \left[ -\frac{\partial \mathbf{j}(x, t)}{\partial t} \right] - \tilde{\mathbf{R}} [-\mathbf{j}(x, t)] \quad (2.21)$$

$$-\frac{\partial \mathbf{j}(x, t)}{\partial x} \equiv -\tilde{\mathbf{C}} \frac{\partial \mathbf{u}(x, t)}{\partial t} \quad (2.22)$$

$$\frac{\partial^2 \mathbf{u}(x, t)}{\partial x^2} \equiv \tilde{\mathbf{L}} \tilde{\mathbf{C}} \frac{\partial^2 \mathbf{u}(x, t)}{\partial t^2} + \tilde{\mathbf{R}} \tilde{\mathbf{C}} \frac{\partial \mathbf{u}(x, t)}{\partial t} \quad (2.23)$$

$$\frac{\partial^2 \mathbf{j}(x, t)}{\partial x^2} \equiv \tilde{\mathbf{C}} \tilde{\mathbf{L}} \frac{\partial^2 \mathbf{j}(x, t)}{\partial t^2} + \tilde{\mathbf{C}} \tilde{\mathbf{R}} \frac{\partial \mathbf{j}(x, t)}{\partial t}. \quad (2.24)$$

The  $n$  coupled transmission lines shown in Figure 2.1 (a) can be decoupled using the congruence transformation method as shown in either Figure 2.1 (b) or (c) depending on whether they are lossless or lossy. Each decoupled single transmission line carries one mode of propagation. The incident waves are decoupled into different modes, and propagate through each decoupled transmission line in different mode velocity, then all the modes are combined at the other end to form the output and the reflected waves.

The task of finding the scattering parameter matrix of  $n$  coupled lossless transmission lines becomes one of finding the scattering parameter matrices of the two identical congruence transformers and the scattering parameter matrices of the  $n$  decoupled single transmission lines. The scattering parameter matrices are not combined because this is the task of the scattering parameter based macromodel simulator. The scattering parameter based macromodel simulator takes full advantage of being a frequency domain simulator and lumps the multiport components together using the Pade or EDPF approximation [39] [40].

Due to the choice of identical reference impedance  $Z_0$  at any port for the scattering parameter based macromodel simulator, the scattering parameter matrix  $S$  of any multiport component is equivalent to its voltage scattering parameter matrix  $S^V$ . Since only the terminal voltages and currents are of an interest, the following representations are introduced (see Figure 2.1):

$$v_{1k}(t) \equiv v_k(x = 0, t) \quad v_{2k}(t) \equiv v_k(x = l, t) \quad i_{1k}(t) \equiv i_k(x = 0, t) \quad i_{2k}(t) \equiv i_k(x = l, t),$$



$$V_p^+ + V_p^- = X(U_p^+ + U_p^-) \quad (2.29)$$

$$I_p^+ - I_p^- = -(X^t)^{-1}(J_p^+ - J_p^-) \quad (2.30)$$

Based on the definition of the incident voltage and current waves, the reflected voltage and current waves, and the reference impedance  $Z_0$ , one can write:

$$V_p^+ = Z_0 I_p^+ \quad (2.31)$$

$$V_p^- = Z_0 I_p^- \quad (2.32)$$

$$U_p^+ = Z_0 J_p^+ \quad (2.33)$$

$$U_p^- = Z_0 J_p^-. \quad (2.34)$$

The definitions of voltage scattering parameters matrix  $S^V$  and its submatrices  $S_{11}^V$ ,  $S_{12}^V$ ,  $S_{21}^V$ , and  $S_{22}^V$  are:

$$\begin{aligned} S^V &= \begin{bmatrix} S_{11}^V & S_{12}^V \\ S_{21}^V & S_{22}^V \end{bmatrix} \\ S_{11}^V &= \left. \frac{V_p^-}{V_p^+} \right|_{U_p^+ = \mathbf{0}} \\ S_{12}^V &= \left. \frac{V_p^-}{U_p^+} \right|_{V_p^+ = \mathbf{0}} \\ S_{21}^V &= \left. \frac{U_p^-}{V_p^+} \right|_{U_p^+ = \mathbf{0}} \\ S_{22}^V &= \left. \frac{U_p^-}{U_p^+} \right|_{V_p^+ = \mathbf{0}} \end{aligned}$$

By arithmetic manipulation of Equations (2.25), (2.26), (2.27), (2.28), (2.29), (2.30), (2.31), (2.32), (2.33), and (2.34), and by setting  $U_p^+$  to be an all-zero column vector, one can find  $S_{11}^V$  as:

$$S_{11}^V = -[X^{-1} + X^t]^{-1}[X^{-1} - X^t]. \quad (2.35)$$

Similarly one can find other submatrices of  $S^V$ :

$$S_{12}^V = 2[X^{-1} + X^t]^{-1} \quad (2.36)$$

$$S_{21}^V = 2[X + (X^t)^{-1}]^{-1} \quad (2.37)$$

$$S_{22}^V = -[X + (X^t)^{-1}]^{-1}[X - (X^t)^{-1}]. \quad (2.38)$$

The voltage scattering parameter matrix  $S^V$  is:

$$S = \begin{bmatrix} -[X^{-1} + X^t]^{-1}[X^{-1} - X^t] & 2[X^{-1} + X^t]^{-1} \\ 2[X + (X^t)^{-1}]^{-1} & -[X + (X^t)^{-1}]^{-1}[X - (X^t)^{-1}] \end{bmatrix}, \quad (2.39)$$

where the submatrix  $X$  is defined in Equation (2.5).

Since the similarity transformation property  $X^{-1} = X^t$  holds for all  $X$ , so it can simultaneously diagonalize the  $\mathbf{L}$ ,  $\mathbf{C}$ ,  $\mathbf{LC}$ , and  $\mathbf{CL}$  matrices. For the coupled lossless transmission line systems, the scattering parameter matrix  $S$  can be simplified to [55]:

$$S = \begin{bmatrix} 0 & X \\ X^t & 0 \end{bmatrix}. \quad (2.40)$$

### 2.3 Scattering Parameter Macromodel of a Decoupled Transmission Line

For each of the decoupled transmission lines shown in Figure 2.1 (b) and (c), its scattering parameter matrix is [20]:

$$S(s) = \frac{1}{2Z_0Z_c(s)\cosh(\gamma(s)) + (Z_c^2(s) + Z_0^2)\sinh(\gamma(s))} \begin{bmatrix} (Z_c^2(s) - Z_0^2)\sinh(\gamma(s)) & 2Z_0Z_c(s) \\ 2Z_0Z_c(s) & (Z_c^2(s) - Z_0^2)\sinh(\gamma(s)) \end{bmatrix}, \quad (2.41)$$

where  $Z_0$  is the reference impedance. For the lossy case, both  $Z_c(s)$  and  $\gamma(s)$  are computed from the eigenvalues  $R_k$ ,  $L_k$ , and  $C_k$  obtained from the diagonalization of the  $\mathbf{R}$ ,  $\mathbf{L}$ , and  $\mathbf{C}$  matrices respectively. The  $Z_c(s) = \sqrt{\frac{R_k + sL_k}{sC_k}}$  is the characteristic impedance of the  $k$ -th line where  $k = 1 \dots n$ , and  $\gamma(s) = \sqrt{(R_k + sL_k)(sC_k)} \cdot l$  is the propagation constant of the  $k$ -th line where  $l$  is the coupling length. For the lossless case, one can simply set  $R_k$  to zero and compute both  $Z_c$  and  $\gamma$  from the eigenvalues  $L_k$  and  $C_k$  of the  $\mathbf{L}$  and  $\mathbf{C}$  matrices respectively. The  $Z_c(s) = \sqrt{\frac{L_k}{C_k}}$  is the characteristic impedance of the  $k$ -th line and  $k = 1 \dots n$ , and  $\gamma(s) = s\sqrt{L_kC_k} \cdot l$  is the propagation constant of the  $k$ -th line.



## 2.4 Experimental Result of the Transient Analysis of Coupled Transmission Lines

The example presented here is one of the MCMC-94 Benchmarks (1994 IEEE Multi-Chip Module Conference Interconnect Simulation Benchmarks). All of the far-end waveforms are simulated with time-of-flight captured explicitly [37]. The circuit and geometry parameters of this example as shown in Figure 2.3 are taken from Cooke's paper [14]. The driving signal is 100-MHz, 50% duty-cycle pulse with  $0.1ns$  rise/fall time. The  $\mathbf{L}$  and  $\mathbf{C}$  of the configuration are:

$$L = \begin{bmatrix} 5.033 \text{ nH/cm} & 1.734 \text{ nH/cm} & 0.818 \text{ nH/cm} \\ 1.734 \text{ nH/cm} & 4.972 \text{ nH/cm} & 1.734 \text{ nH/cm} \\ 0.818 \text{ nH/cm} & 1.734 \text{ nH/cm} & 5.033 \text{ nH/cm} \end{bmatrix}$$

$$C = \begin{bmatrix} 0.667 \text{ pF/cm} & -0.163 \text{ pF/cm} & -0.0145 \text{ pF/cm} \\ -0.163 \text{ pF/cm} & 0.722 \text{ pF/cm} & -0.163 \text{ pF/cm} \\ -0.0145 \text{ pF/cm} & -0.163 \text{ pF/cm} & 0.667 \text{ pF/cm} \end{bmatrix}$$

Two enlarged portions of the simulation waveforms are shown in Figure 2.4 (a) and (b). The simulation waveforms of this example are shown in Figure 2.5 (a), (b), (c), and (d). The ASTAP (IBM circuit simulator) and Cooke's simulation waveforms are digitized from the results published in Cooke's paper [14] while the SWEC results are taken from MCMC-94 benchmark results. The SWEC uses an analytic method to find the derivatives of the admittance in order to compute the moments and because of its complexity, it can only find very low order moments [41]. In all of the plots, the results obtained from SWEC and our macromodel simulators agree very well with the published ASTAP results. However, it appears that the macromodel simulation waveforms match those of the ASTAP simulator better than those derived from the SWEC. In all of the waveform plots, Cooke's results deviate from the ASTAP results the most. The total running time for this example is 7.24 seconds on a SUN SPARC station 1+. The running times of other simulators are not listed because they are executed on different machines. For small examples such as the



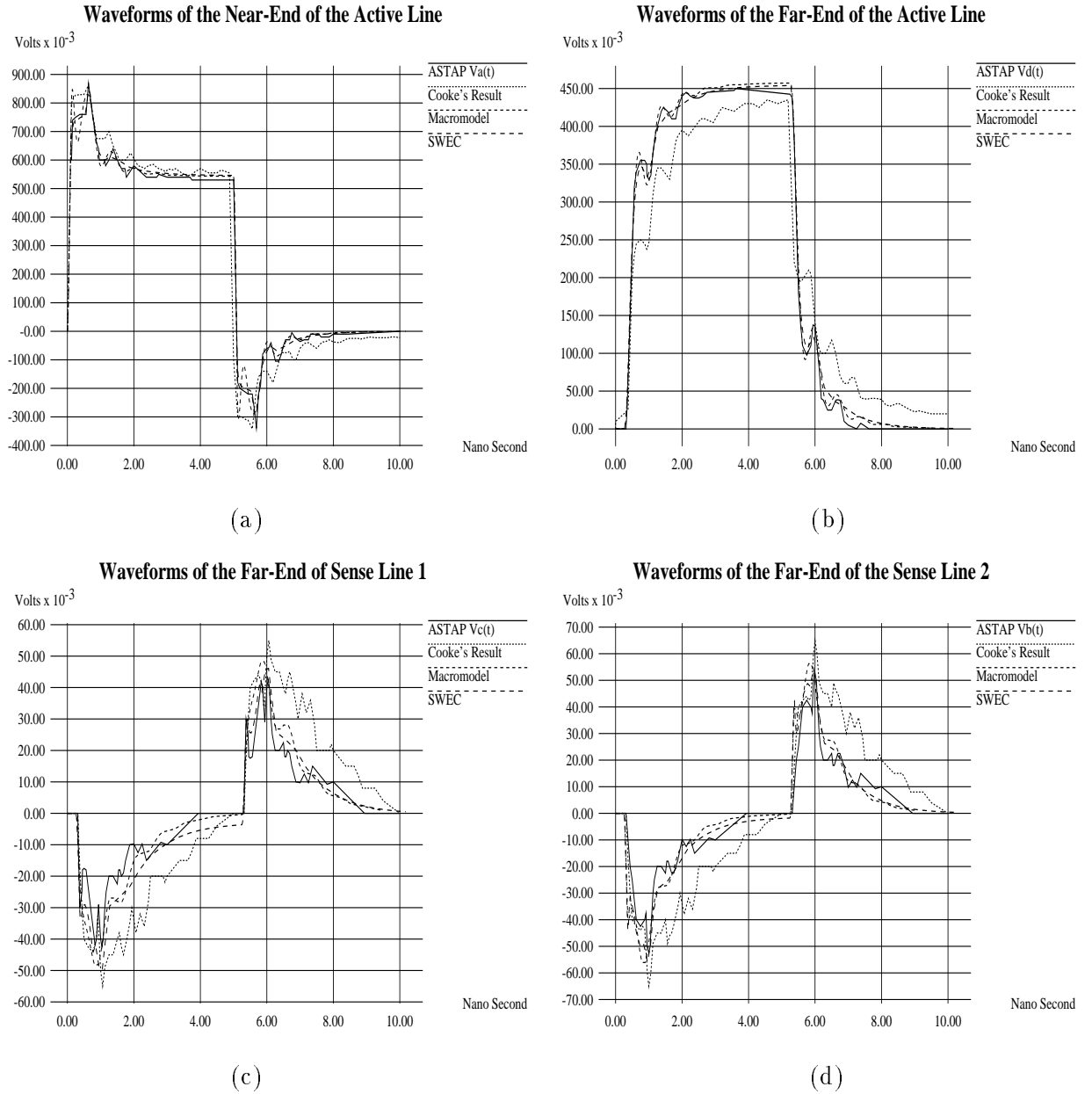


Figure 2.5: **Simulation Waveforms of the Two Groups of Three Coupled Lossless Transmission Lines:** The topology is shown in Figure 2.3. The output waveforms of the near end of the active line are shown in (a) together with its corresponding Cooke's, ASTAP and SWEC simulation results. The output waveforms of the far end of all three lines are shown in (b), (c), and (d) together with their corresponding Cooke's ASTAP and SWEC simulation results.

## CHAPTER 3. Transient Analysis of Coupled Transmission Lines Characterized with the Frequency-Dependent Losses

The objectives of this part of the thesis is to provide a method of finding the Taylor series expansions of the S-parameter functions from the measured parasitic data,  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ . A curve-fitting is first applied to find the moments of these four parasitic functions, which are subsequently used to compute the exact moments of the S-parameter functions. Based on the exact moments found by this approach, the macromodel of transmission lines characterized with the frequency-dependent losses is constructed and the transient simulation is performed.

Comparing to other accurate time domain simulators, the S-parameter macromodel based simulator is significantly more efficient, at the least thirty times faster in handling large circuits. However, it can only provide moderate accuracy because it utilizes lower order Pade approximations. This trade-off between accuracy and efficiency has to be made in order to play “what-if” scenarios for a performance-driven layout synthesis in which thousands of simulations must be executed to obtain timing and amplitude information. Classical approaches which use empirical equations might not work in this case because the over-simplification assumptions. Only a simulator with sufficient accuracy, and that does not require excessive time for simulate can fit this requirement. The S-parameter macromodel simulator is well suited for this kind of application.

### 3.1 Representing the Frequency-Dependent Coupled Transmission Lines Using the S-Parameter Macromodel

Previous research which took frequency-dependent losses into consideration include the work of Gruodis et al. [29], Schutt-Aine et al. [48], Chang et al. [8], Beyene [4], Baumgartner [2], Cooke et al. [14], Gordon et al. [28], and Nguyen [44]. Gruodis et al. measured the admittance matrix  $\mathbf{Y}_{2n}$  and impedance matrix  $\mathbf{Z}_{2n}$  of the transmission lines, computed the  $\mathbf{Y}_0$  matrix and propagation constant  $\Gamma$ , and then simulated the circuit's transient behavior using the state variable transfer function method [29]. Schutt-Aine et al. utilized the scattering parameter matrix method [48]. Chang et al. chose the “method of characteristic” with a network synthesis [8]. Beyene combined the bi-level waveform relaxation with scattering parameters [4]. Baumgartner used a state variable transfer function with an exponential approximation [2]. Cooke et al. chose the scattering parameter frequency domain simulation with the Fast Fourier Transformation (FFT) method [14]. Gordon et al. used the impulse response convolution method [28], and Nguyen used the state variable transfer function with the rational function approximation improvement [44]. In some of these previous researches including those by Schutt-Aine et al. [48], Beyene [4], and Cooke et al. [14], they chose to use the scattering parameter for the analysis of the transmission lines with the frequency-dependent losses. The scattering parameter matrix method adopted by Schutt-Aine et al. [48] used those of  $S_{11}(s)$  and  $S_{21}(s)$  to find the transmission  $T(t)$  and reflection  $\Gamma(t)$  matrix for the time domain convolution. The drawback of this Schutt-Aine method is the large number of matrix operations that were required. The bi-level waveform relaxation method adopted by Beyene [4] utilized the FFT and the inverse FFT (IFFT) to iterate between the frequency domain simulation and the time domain simulation. The deficiency of Beyene's method is the need for more than one thousand data points in order to do the evaluation of the FFT (IFFT) operation with the same degree of accuracy as other approaches. The scattering parameter frequency domain simulation with the FFT method by Cooke et al. [14] uses all the scattering parameters without reduction. The drawbacks of Cooke et al. method are its poor efficiency and its lack of accuracy when compared

to the published ASTAP results. In summary, all of the above methods lack the required efficiency because they do not employ lower order approximations and the macromodel reductions that were found in the scattering parameter macromodel based simulator.

Building the novel macromodels for both the frequency-dependent single and coupled transmission lines facilitates an accurate and a more efficient transient simulation of the interconnections that are characterized by lossy transmission lines with skin effects. The contribution of this part of the thesis is to determine the moments of the S-parameters of the decoupling congruence transformer and the decoupled transmission lines characterized with frequency-dependent losses from the curve-fitting coefficients of the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  data sets. Section 3.2 will derive the representations of the S-parameters macromodel for the frequency-dependent single transmission line and Section 3.3 will derive the representations of the S-parameters macromodel for the frequency-dependent coupled transmission line.

### 3.2 S-Parameter Macromodel of the Frequency-Dependent Single Transmission Line

With the assumption of quasi-TEM wave propagation, the distributions of voltages and currents in a single lossy transmission line can be described by the generalized Telegraphist's equations [9] shown in Equation (2.1) and (2.2). The frequency-dependent parasitic functions:  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  are used to characterize transmission line with frequency-dependent losses. The scattering parameter matrix for a transmission line is shown in Equation (3.1).

$$S(s) = \frac{1}{2Z_0Z_c(s)\cosh(\gamma(s)) + (Z_c^2(s) + Z_0^2)\sinh(\gamma(s))} \begin{bmatrix} (Z_c^2(s) - Z_0^2)\sinh(\gamma(s)) & 2Z_0Z_c(s) \\ 2Z_0Z_c(s) & (Z_c^2(s) - Z_0^2)\sinh(\gamma(s)) \end{bmatrix}, \quad (3.1)$$

The  $Z_c(s)$  is the characteristic impedance and  $\gamma(s)$  is the propagation constant. Both  $Z_c(s)$  and  $\gamma(s)$  are computed from the frequency-dependent values of  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$

based on the following equations:

$$\begin{aligned}\gamma(s) &= \sqrt{(R(s) + sL(s))(G(s) + sC(s))} \cdot l \\ Z_c(s) &= \sqrt{\frac{R(s) + sL(s)}{G(s) + sC(s)}}.\end{aligned}$$

For an uniform conductor, the S-parameter matrix is symmetrical [21]. Rearranging the representations of  $S_{11}(s)$  and  $S_{21}(s)$  in Equation (2.41), one has:

$$S_{11}(s) = \frac{[Z_c^2(s) - Z_0^2] \cdot \sinh(\gamma(s))}{2Z_0Z_c(s) \cosh(\gamma(s)) + (Z_c^2(s) + Z_0^2) \sinh(\gamma(s))}, \quad (3.2)$$

$$S_{21}(s) = \frac{2Z_0}{2Z_0Z_c(s) \cosh(\gamma(s)) + (Z_c^2(s) + Z_0^2) \sinh(\gamma(s))}. \quad (3.3)$$

In order to find the lower order approximations of  $S_{11}(s)$  and  $S_{21}(s)$ , the representations of the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  must first be found. A curve-fitting method is then adapted to find the coefficients of the polynomials which model the parasitics. At first, the linear fit is run on each data set to obtain two coefficients as the initial assignment for the successive curve-fitting steps. Then the problem of finding the coefficients of the polynomials are transformed into one of the least-square error estimation. The Levenberg-Marquardt method [42] [16] is used to solve this least-square error estimation problem. This method is chosen because it combines the best features of both the Taylor series expansion and the gradient methods. It can find the best solution as well as gradient methods even if the solution is outside the circle of convergence; and the rate of convergence is as fast as Taylor series methods.

The two coefficients found in the linear fit are passed on to the Levenberg-Marquardt method as an initial guess. The Levenberg-Marquardt method then iterates to find the best fit coefficients for the curves of  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ . This method stops when the results converges or the number of iteration exceeds a preset limit. For all of the experiments, this curve-fitting method shows better results than the one-pass least-square fit or the singular-value decomposition fit methods.

If one defines:

$$T(s) \equiv \gamma^2(s) \equiv \sum_{i=0}^q t_i \cdot s^i + o(s^q) \quad (3.4)$$

$$\begin{aligned} A(s) &\equiv Z_c(s)\gamma(s) - Z_0^2 \cdot \frac{\gamma(s)}{Z_c(s)} \\ &\equiv \sum_{i=0}^q a_i \cdot s^i + o(s^q) \end{aligned} \quad (3.5)$$

$$\begin{aligned} B(s) &\equiv Z_c(s)\gamma(s) + Z_0^2 \cdot \frac{\gamma(s)}{Z_c(s)} \\ &\equiv \sum_{i=0}^q b_i \cdot s^i + o(s^q), \end{aligned} \quad (3.6)$$

whereas  $\gamma^2(s) = (R(s) + sL(s))(G(s) + sC(s))$ ,  $\frac{\gamma(s)}{Z_c(s)} = G(s) + sC(s)$ , and  $Z_c(s)\gamma(s) = R(s) + sL(s)$ , there is no square root involved in the evaluation of the approximations of the  $T(s)$ ,  $A(s)$ , and  $B(s)$  complex functions. The approximation of the  $R(s)$ ,  $L(s)$ ,  $G(s)$ , and  $C(s)$  real functions are known through curve-fitting. The coefficients of the  $T(s)$ ,  $A(s)$ , and  $B(s)$  complex functions can be found through simple polynomial operations.

If one further defines:

$$\begin{aligned} T(s) &\equiv t_0 + \sum_{i=1}^q t_i \cdot s^i + o(s^q) \\ &\equiv t_0 + T_q(s) + o(s^q). \end{aligned} \quad (3.7)$$

where the constant term is separated from the rest of function. The separation will make the finding of the coefficients in Equation (3.12) and (3.13) much easier. The expansions of  $2Z_0 \cosh(\gamma(s))$  and  $\frac{\sinh(\gamma(s))}{\gamma(s)}$  are shown in Equation (3.8) and (3.9):

$$\begin{aligned} U(s) &\equiv 2Z_0 \cosh(\gamma(s)) \equiv 2Z_0 \sum_{i=0}^{\infty} \frac{\gamma^{2i}(s)}{2i!} \equiv 2Z_0 \sum_{i=0}^{\infty} \frac{T^i(s)}{2i!} \\ &\equiv 2Z_0 \sum_{i=0}^q u_i \cdot s^i + o(s^q) \end{aligned} \quad (3.8)$$

$$\begin{aligned} V(s) &\equiv \frac{\sinh(\gamma(s))}{\gamma(s)} \equiv \sum_{i=0}^{\infty} \frac{\gamma^{2i+1}(s)}{(2i+1)!} \equiv \sum_{i=0}^{\infty} \frac{T^i(s)}{(2i+1)!} \\ &\equiv \sum_{i=0}^q v_i \cdot s^i + o(s^q). \end{aligned} \quad (3.9)$$

Based on Equations (3.5), (3.6), (3.8), and (3.9), Equations (3.2) and (3.3) can be rewritten as:



$$S_{11}(s) = \frac{A(s) \cdot V(s)}{U(s) + B(s) \cdot V(s)} \quad (3.10)$$

$$S_{21}(s) = \frac{2Z_0}{U(s) + B(s) \cdot V(s)}. \quad (3.11)$$

It can be shown that:

$$U(s) = 2Z_0 \sum_{k=0}^q \beta_k \cdot T_q^k(s) + o(s^q) \quad (3.12)$$

$$V(s) = \sum_{k=0}^q \alpha_k \cdot T_q^k(s) + o(s^q), \quad (3.13)$$

where

$$\beta_k = \sum_{i=0}^{\infty} \frac{C_i^k}{2i!} \cdot t_0^{i-k} \quad (3.14)$$

$$\alpha_k = \sum_{i=0}^{\infty} \frac{C_i^k}{(2i+1)!} \cdot t_0^{i-k}, \quad (3.15)$$

and  $t_0$  is the separated constant term in Equation (3.7). Although the summation of both  $\alpha_k$  and  $\beta_k$  are an infinite series, in reality, the inverse of the factorial is a fast converging series and can be truncated at a certain point without introducing appreciable error. The coefficients of  $U(s)$  and  $V(s)$  are:

$$u_i = \sum_{k=1}^i \beta_k \sum_{j_1+j_2+\dots+j_k=i} t_{j_1} \cdot t_{j_2} \cdots t_{j_k} \quad (3.16)$$

$$v_i = \sum_{k=1}^i \alpha_k \sum_{j_1+j_2+\dots+j_k=i} t_{j_1} \cdot t_{j_2} \cdots t_{j_k}. \quad (3.17)$$

Since the coefficients of the polynomial  $T(s)$  are known from Equation (3.4), the coefficients of  $U(s)$  and  $V(s)$  can then be computed.

If one defines:

$$S_{11}(s) = \frac{C(s)}{D(s)} \quad (3.18)$$

$$S_{21}(s) = \frac{C'(s)}{D(s)}. \quad (3.19)$$

Comparing Equations (3.10), (3.11), (3.18), and (3.19), one finds:

$$C(s) \equiv \sum_{i=0}^q c_i \cdot s^i + o(s^q) = A(s) \cdot V(s) \quad (3.20)$$

$$\begin{aligned}
D(s) &\equiv \sum_{i=0}^q d_i \cdot s^i + o(s^q) \\
&= U(s) + B(s) \cdot V(s)
\end{aligned} \tag{3.21}$$

$$C'(s) = 2Z_0. \tag{3.22}$$

It can be shown that the coefficients of  $C(s)$  and  $D(s)$  are:

$$c_k = \sum_{i+j=k} a_i v_j \tag{3.23}$$

$$d_k = u_k + \sum_{i+j=k} b_i v_j. \tag{3.24}$$

If the moments of  $S_{11}(s)$  and  $S_{21}(s)$  are  $m_i$  and  $n_i$ , one can write:

$$S_{11}(s) = \sum_{i=0}^q m_i \cdot s^i + o(s^q) \tag{3.25}$$

$$S_{21}(s) = \sum_{i=0}^q n_i \cdot s^i + o(s^q). \tag{3.26}$$

From Equation (3.18) and (3.26), one can derive:

$$\sum_{i+j=k} m_i d_j = c_k. \tag{3.27}$$

If one denotes:

$$D_q = \begin{bmatrix} d_0 & 0 & 0 & \dots & 0 \\ d_1 & d_0 & 0 & \dots & 0 \\ d_2 & d_1 & d_0 & \dots & 0 \\ \dots & & & & \\ d_q & d_{q-1} & d_{q-2} & \dots & d_0 \end{bmatrix} \tag{3.28}$$

$$M_q = \begin{bmatrix} m_0 & m_1 & m_2 & \cdot & \cdot & \cdot & m_q \end{bmatrix}^T \tag{3.29}$$

$$C_q = \begin{bmatrix} c_0 & c_1 & c_2 & \cdot & \cdot & \cdot & c_q \end{bmatrix}^T, \tag{3.30}$$

where  $T$  represents the transpose of the vector, one can rewrite  $D_q M_q = C_q$ , i.e.  $M_q = D_q^{-1} C_q$ . Thus the moments  $m_i$  of the  $S_{11}(s)$  function can be found through simple backward substitutions. Similarly, if one denotes:

$$N_q = \begin{bmatrix} n_0 & n_1 & n_2 & \cdot & \cdot & \cdot & n_q \end{bmatrix}^T \quad (3.31)$$

$$C'_q = \begin{bmatrix} 2Z_0 & 0 & 0 & \cdot & \cdot & \cdot & 0 \end{bmatrix}^T, \quad (3.32)$$

where  $T$  again represents the transpose of the vector, one can rewrite  $D_q N_q = C'_q$ , i.e.  $N_q = D_q^{-1} C'_q$ , and the moments  $n_i$  of the  $S_{21}(s)$  function can again be found through simple backward substitutions.

### 3.3 S-Parameter Macromodel of the Frequency-Dependent Coupled Transmission Line

Previous research which used scattering parameter for the analysis of the coupled transmission lines include those of the general convolution by Winklestein et al. [57], the full-wave analysis by Cooke et al. [14] and the time domain transient simulation by Schutt-Aine et al. [48]. While Schutt-Aine et al. demonstrated a greater accuracy improvement in simulating the circuits that included non-linear drivers and terminations, their method involved costly matrix computation for converting the frequency domain scattering parameter matrix representation into the time domain transmission/reflection matrix representation [48]. Cooke et al. illustrated an ability to simulate frequency dependent modal propagation, but a time consuming full-wave analysis was often required [14]. Recently, a novel frequency domain simulator using scattering parameter based macromodels has been presented by Liao et al [39] [40]. Based on the scattering parameter based macromodel, Pade techniques or Exponentially Decayed Polynomial Function (EDPF) can be used to approximate transfer functions of the coupled interconnects. In the following section, we overcome some of these shortcomings by deriving the S-parameter macromodel for the frequency-dependent coupled transmission lines.

Based upon the assumption of quasi-TEM wave propagation, the distributions of voltages and currents in a  $n$  coupled lossy transmission-line system can be described by the generalized Telegraphist's equations [9]:

$$\frac{\partial \mathbf{v}(x, t)}{\partial x} = -\mathbf{L}(\mathbf{f}) \frac{\partial \mathbf{i}(x, t)}{\partial t} - \mathbf{R}(\mathbf{f}) \mathbf{i}(x, t) \quad (3.33)$$

$$\frac{\partial \mathbf{i}(x, t)}{\partial x} = -\mathbf{C}(\mathbf{f}) \frac{\partial \mathbf{v}(x, t)}{\partial t} - \mathbf{G}(\mathbf{f}) \mathbf{v}(x, t), \quad (3.34)$$

where  $0 \leq x \leq l$ .  $\mathbf{v}(x, t)$  and  $\mathbf{i}(x, t)$  are column vectors defining the voltages distributions  $v_k(x, t)$  and currents distributions  $i_k(x, t)$  on the conductors  $k = 1, 2, 3, \dots, n$ . The  $\mathbf{L}(\mathbf{f})$  and  $\mathbf{C}(\mathbf{f})$  are the  $n$  by  $n$  symmetric matrices of the frequency-dependent per-unit-length inductance and capacitance of the  $n$  conductor system. The  $\mathbf{R}(\mathbf{f})$  and  $\mathbf{G}(\mathbf{f})$  is the  $n$  by  $n$  symmetric matrices of the frequency-dependent per-unit-length resistance and conductance of the  $n$  conductor system [48].

It is very important to accurately model the frequency dependence of the parasitics for the case of the coupled transmission lines. There are off-diagonal elements in the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  matrices describing the mutual coupling effects which do not exist in the case of the single transmission line. The mutual inductance  $\mathbf{L}_{ij}(f)$ , where  $i \neq j$ , increases as frequency increases due to more coupling between lines at higher frequency. The mutual capacitance  $\mathbf{C}_{ij}(f)$ , where  $i \neq j$ , stays constant as the  $\mathbf{C}_{ii}(f)$ . Both the  $\mathbf{R}_{ii}(f)$  and  $\mathbf{R}_{ij}(f)$ , where  $i \neq j$ , increase as the frequency rises.

There are two major methods to find the time domain transient response waveforms of a coupled transmission line system. One method is to find the impulse response of the linear coupled transmission line system and then use either the convolution or the waveform relaxation to find the time domain waveforms. However, this method suffers from both the large memory requirement and the long computation time that are required. The other method is a modal wave propagation decoupling method which is preferred over the first method because it models the physical phenomenon of  $n$  fundamental mode of the wave propagation that exists in the  $n$  multi-conductor transmission line system. By decoupling the modal waves, the simulator is only required to memorize a period of the waveforms equal to the time-of-flight of each decoupled transmission line, which is much shorter when compared to the duration of the impulse response. Furthermore, with the help of the S-parameter macromodel, the recursive convolution can be applied with a significantly shorter computation time. After successfully decoupling of the coupled transmission lines system, the computation of the scattering parameter macromodel of the entire system becomes

the computation of the decoupling networks and those of the decoupled transmission lines with frequency-dependent losses. The macromodel of the latter is already available and is presented in Section 3.2. Because the coupled transmission lines are also characterized with frequency-dependent losses, the decoupling of the modal wave propagation of the coupled transmission lines has to take these losses into consideration.

To incorporate the macromodel of the frequency-dependent decoupling networks into the S-parameter macromodel simulator presents a very difficult challenge. The decoupling networks are known to be represented by a matrix with frequency-dependent elements [4]. This process requires the finding of a frequency-dependent transformation matrix in order to decouple the system. It is a complex process and requires the eigenvalues at each frequency point prior to diagonalization. The resulting matrix elements are characterized by tabulated scattering parameter data which present the same difficulty in finding the moments as in the case of the measured data macromodel.

Taking the Laplace transform of the Equation (3.33) and (3.34), they can be rewritten as:

$$\frac{\partial \mathbf{V}(x, s)}{\partial x} = -\mathbf{Z}\mathbf{I}(x, s) \quad (3.35)$$

$$\frac{\partial \mathbf{I}(x, s)}{\partial x} = -\mathbf{Y}\mathbf{V}(x, s), \quad (3.36)$$

where  $\mathbf{Z} = \mathbf{R} + jw\mathbf{L}$ , and  $\mathbf{Y} = \mathbf{G} + jw\mathbf{C}$ . Throughout this thesis, the following assumptions reported by Blazeck et al. are used [5]. The assumptions are: the modes of propagation must be TEM or quasi-TEM, and the lines are of uniform cross-section throughout their length; that is,  $\mathbf{R}$ ,  $\mathbf{L}$ ,  $\mathbf{C}$ , and  $\mathbf{G}$  are assumed to be constant with respect to the spatial variable  $x$ .

Solving Equation (3.35) and (3.36), one has:

$$\frac{\partial^2 \mathbf{V}(x, s)}{\partial x^2} = \mathbf{Z}\mathbf{Y}\mathbf{V}(x, s) = \mathbf{\Gamma}^2 \mathbf{V}(x, s) \quad (3.37)$$

$$\frac{\partial^2 \mathbf{I}(x, s)}{\partial x^2} = \mathbf{Y}\mathbf{Z}\mathbf{I}(x, s) = (\mathbf{\Gamma}^2)^T \mathbf{I}(x, s), \quad (3.38)$$

where  $T$  indicates transpose and  $\mathbf{\Gamma}$  is defined as  $\mathbf{\Gamma}^2 = \mathbf{Z}\mathbf{Y}$  [5]. Define

$$\mathbf{\Gamma} = \mathbf{X}\mathbf{\Lambda}\mathbf{X}^{-1}, \quad (3.39)$$

where  $\mathbf{X}$  is the eigenvectors of  $\mathbf{\Gamma}$ ; therefore, they are also the eigenvectors of  $\mathbf{\Gamma}^2$ , and  $\mathbf{\Lambda}$  is the diagonal matrix of the eigenvalues of  $\mathbf{\Gamma}$ . It can be shown that [56] [18]

$$\frac{\partial \mathbf{V}_{\mathbf{m}}(x, s)}{\partial x} = -\mathbf{\Lambda} \mathbf{Z}_{\mathbf{m}} \mathbf{I}_{\mathbf{m}}(x, s) \quad (3.40)$$

$$\frac{\partial \mathbf{I}_{\mathbf{m}}(x, s)}{\partial x} = -\mathbf{\Lambda} \mathbf{Y}_{\mathbf{m}} \mathbf{V}_{\mathbf{m}}(x, s), \quad (3.41)$$

where  $\mathbf{V}(x, s) = \mathbf{X} \mathbf{V}_{\mathbf{m}}(x, s)$ ,  $\mathbf{I}(x, s) = (\mathbf{X}^T)^{-1} \mathbf{I}_{\mathbf{m}}(x, s)$ . The modal impedance matrix  $\mathbf{Z}_{\mathbf{m}}$  and modal admittance matrix  $\mathbf{Y}_{\mathbf{m}}$  are related to the eigenvector matrix  $\mathbf{X}$  and the impedance matrix  $\mathbf{Z}$  by

$$\mathbf{Z}_{\mathbf{m}} = (\mathbf{Y}_{\mathbf{m}})^{-1} = \mathbf{\Lambda}^{-1} \mathbf{X} \mathbf{Z} (\mathbf{X}^T)^{-1}. \quad (3.42)$$

With the eigenvector matrix  $\mathbf{X}$ , the original coupled transmission lines can be decoupled into two congruence transformers and a set of  $n$  decoupled transmission lines [9]. The task of finding the macromodel of the frequency-dependent coupled transmission lines becomes that of finding the macromodel representations of the congruence transformers and the frequency-dependent single transmission lines [56]. The macromodel of the single transmission line that is characterized with frequency-dependent losses has already been developed in Section 3.2. The remaining task is to find the macromodel representation of a frequency-dependent congruence transformer.

Bayard first outlined the transformation,  $\mathbf{A}^t \mathbf{Z} \mathbf{A}$ , and called it “translator” [3]. Hazony is the first one to name the transformation “congruence transformer” in his book [30]. Chang used the congruence transformer to decouple both the lossless [7] and lossy coupled transmission lines [9]. Chang’s method for the analysis of coupled transmission lines relies on simultaneously diagonalizing all the matrices using a special conditioned matrix.

It is known that the modal eigenvectors of two symmetrical coupled transmission lines are frequency independent constant vectors even if the lines are characterized by frequency-dependent parasitics [28]. It can be shown that the decoupling networks can be constructed from the constant eigenvectors:

$$X = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}. \quad (3.43)$$

These two vectors correspond to the odd and even mode of propagation that exists in the symmetrical coupled transmission lines.

For asymmetrical coupled dual transmission lines and for coupled transmission lines with more than two conductors which are characterized with frequency-dependent parasitics, the model structure of the lines becomes frequency-dependent [28], and finding the moments of the model structure poses an even greater challenge.

It can be shown that for the coupled lossy transmission line systems, the congruence transformer matrix is the eigenvector matrix of the complex matrix  $\mathbf{\Gamma}$  [5]. The existence of such eigenvector matrix  $\mathbf{X}$ , that simultaneously diagonalize the complex matrices  $\mathbf{\Gamma}$ ,  $\mathbf{\Gamma}^2$ ,  $\mathbf{ZY}$ , and  $\mathbf{YZ}$ , is assumed throughout this thesis. This assumption is also adopted by Gordon et al. [28], Blazeck et al. [5] and Schutt-Aine et al. [48].

Gordon et al. suggested that the frequency-dependent congruence transformer can be found by performing the congruence decoupling at each frequency point, and by checking the orthogonality of all the eigenvectors for all the frequency points [28]. If the eigenvectors were not orthogonal to each other, column swapping must be performed so that for all the congruence transformation matrices at all the frequency points are orthogonal to any other one. During this process, the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  matrices have been diagonalized to be:  $\text{diag}(\mathbf{R}(f))$ ,  $\text{diag}(\mathbf{L}(f))$ ,  $\text{diag}(\mathbf{C}(f))$  and  $\text{diag}(\mathbf{G}(f))$  eigenvalue matrices. The  $i$ -th diagonal eigenvalues of each frequency points constitute the parasitic of the decouple frequency-dependent single transmission line. Thus these diagonal eigenvalue matrices can be used to form the macromodels using a method outlined in the Section 3.2.

Once the tabulated S-parameter data for the congruence transformation have been found, a curve-fitting using Levenberg-Marquardt method [42] [16] is used to find the coefficients for the construction of the congruence transformer:  $X(f)$ .

Having defined the scattering parameter matrix  $S$  of the congruence transformer  $X$  in Equation 2.39 [56], one can extend the representation of the congruence transformer to

include frequency-dependent effects, namely  $X(f)$ , and rewrite:

$$S(s) = \begin{bmatrix} -[X^{-1}(f) + X^t(f)]^{-1}[X^{-1}(f) - X^t(f)] \\ 2[X^{-1}(f) + X^t(f)]^{-1} \\ 2[X(f) + (X^t(f))^{-1}]^{-1} \\ -[X(f) + (X^t(f))^{-1}]^{-1}[X(f) - (X^t(f))^{-1}] \end{bmatrix}, \quad (3.44)$$

where the sub-matrix  $X(f)$  is found using curve-fitting of the tabulated congruence transformation data.

The macromodels of the two congruence transformers and the  $n$  decoupled transmission lines are passed onto the S-parameter macromodel based simulator to perform the transient analysis.

### 3.4 Experimental Results for Frequency-Dependent Transmission Lines

The data in the first two examples as well as the fourth example presented here are obtained from Dr. J. C. Liao of Intel Corporation. The data in the third example is obtained from the user manual of Mittra's "mtltda" simulator. Figure 3.1 (b) only shows the simulation result which takes frequency-dependent losses into consideration. Figure 3.2 shows the different simulation results between *taking* and *not taking* the frequency-dependent losses into consideration. The discrepancy in simulation waveforms confirms that one needs to include frequency-dependent losses in circuit simulation. Figure 3.4 and Figure 3.5 show the comparison between the S-parameter macromodel based simulator and a time-domain circuit simulator. For all of the simulation results of the S-parameter macromodels, all of the far-end waveforms are simulated with the time-of-flight captured explicitly with the method stated in Section 5.5. Although the S-parameter macromodel based simulator does not have the accuracy demonstrated by the time-domain simulator, it provides more than thirty times speedup in Example 4. This kind of efficiency lands it in the application of performance-driven layout synthesis.



### 3.4.1 Example 1: A Single Transmission Line

This example is a single transmission line characterized with frequency-dependent losses. The frequency dependence of the per-unit-length inductance and resistance are given in Table 3.1. The per-unit-length capacitance is  $1.460\text{pF/cm}$  and the per-unit-length conductance is assumed to be zero. The driving signal is 100-MHz, 50% duty-cycle pulse with  $0.5\text{ns}$  rise/fall time. The circuit schematic is shown in 3.1 (a) with the component values. The simulation waveforms of this example are shown in Figure 3.1 (b).

Frequency	L (nH/cm)	R (ohm/cm)
10 kHz	4.070	5.000
100 MHz	4.069	5.000
250 MHz	4.064	5.000
500 MHz	4.050	5.150
750 MHz	4.032	5.310
1 GHz	4.012	5.520
2 GHz	3.904	6.750
4 GHz	3.789	8.960
6 GHz	3.724	10.85
8 GHz	3.645	12.35

Table 3.1: The Frequency-Dependent Per-Unit-Length Inductance and Resistance

### 3.4.2 Example 2: Two Coupled Transmission Lines

This is an example with two coupled transmission lines characterized with frequency-dependent losses. The frequency-dependent per-unit-length inductance and resistance are given in Table 3.2. The per-unit-length capacitance matrix is a constant matrix which does not vary with frequency:

$$\begin{bmatrix} 1.637\text{pF/cm} & -0.177\text{pF/cm} \\ -0.177\text{pF/cm} & 1.637\text{pF/cm} \end{bmatrix} \quad (3.45)$$



The per-unit-length conductance is assumed to be zero. The driving signal is 100-MHz, 50% duty-cycle pulse with  $0.5ns$  rise/fall time. The circuit schematic is shown in 3.2 (a) with the component values. The simulation waveforms of this example are shown in Figure 3.2 (b), (c), (d) and (e).

For all of the figures, there are three output waveforms of the same coupled transmission lines. These are modeled:

- using the lossy macromodel.
- using the frequency-independent macromodel.
- using the frequency-dependent macromodel.

The per-unit-length parasitics of the lossy macromodel are taken from the DC values of the frequency-dependent model. The frequency-independent macromodel is created using the DC values of the frequency-dependent model at all of the frequency points. The results show that the output waveforms of the frequency-independent macromodel match that of the lossy macromodel exactly as expected. The results also indicate that the output waveforms of the frequency-dependent macromodel differ from that of the lossy and frequency-independent model because of the frequency-dependent nature of the per-unit-length parasitic. This demonstrates the importance of taking frequency-dependent losses into consideration when doing circuit simulation.

### 3.4.3 Example 3: Two Cascade Sections of Coupled Transmission Lines

This is an example of two cascade sections of two coupled transmission lines, one section is characterized with frequency-dependent losses, and the other is characterized only with lossless model. The frequency-dependent per-unit-length inductance and resistance of the first section are given in Table 3.3. The frequency-dependent per-unit-length capacitance matrix of the first section is given in Table 3.4. The per-unit-length conductance is assumed to be zero.



Frequency	$L_{11} \ L_{22}$ (nH/cm)	$L_{12} \ L_{21}$ (nH/cm)	$R_{11} \ R_{22}$ (ohm/cm)	$R_{12} \ R_{21}$ (ohm/cm)
DC	5.150	0.995	0.6406	0.1283
20 MHz	4.777	0.943	0.6829	0.1325
50 MHz	4.511	0.818	0.7793	0.1736
100 MHz	4.354	0.818	1.0147	0.2359
200 MHz	4.022	0.688	1.2572	0.2609
500 MHz	3.601	0.588	2.1743	0.4284
1 GHz	3.409	0.580	3.6886	0.6957

Table 3.3: **The Frequency-Dependent Per-Unit-Length Inductance, Mutual Inductance, and Resistance**

Frequency	$C_{11} \ C_{22}$ (pF/cm)	$C_{12} \ C_{21}$ (pF/cm)
DC	0.862	-0.140
20 MHz	0.752	-0.159
50 MHz	0.756	-0.140
100 MHz	0.754	-0.153
200 MHz	0.786	-0.147
500 MHz	0.768	-0.142
1 GHz	0.766	-0.146

Table 3.4: **The Frequency-Dependent Per-Unit-Length Capacitance Matrix**

The second section is characterized as a lossless coupled transmission lines. The per-unit-length inductance and capacitance matrices are as follows:

$$\begin{aligned}
 \mathbf{L} &= \begin{bmatrix} 5.105nH/cm & -0.995nH/cm \\ -0.995nH/cm & 5.105nH/cm \end{bmatrix} \\
 \mathbf{C} &= \begin{bmatrix} 0.862pF/cm & -0.140pF/cm \\ -0.140pF/cm & 0.862pF/cm \end{bmatrix}
 \end{aligned} \tag{3.46}$$

The driving signal is 100-MHz, 50% duty-cycle pulse with 0.5ns rise/fall time. The circuit schematic is shown in 3.3 with the component values. The simulation waveforms of this example are shown in Figure 3.4 (a) (b), (c), (d), (e) and (f).



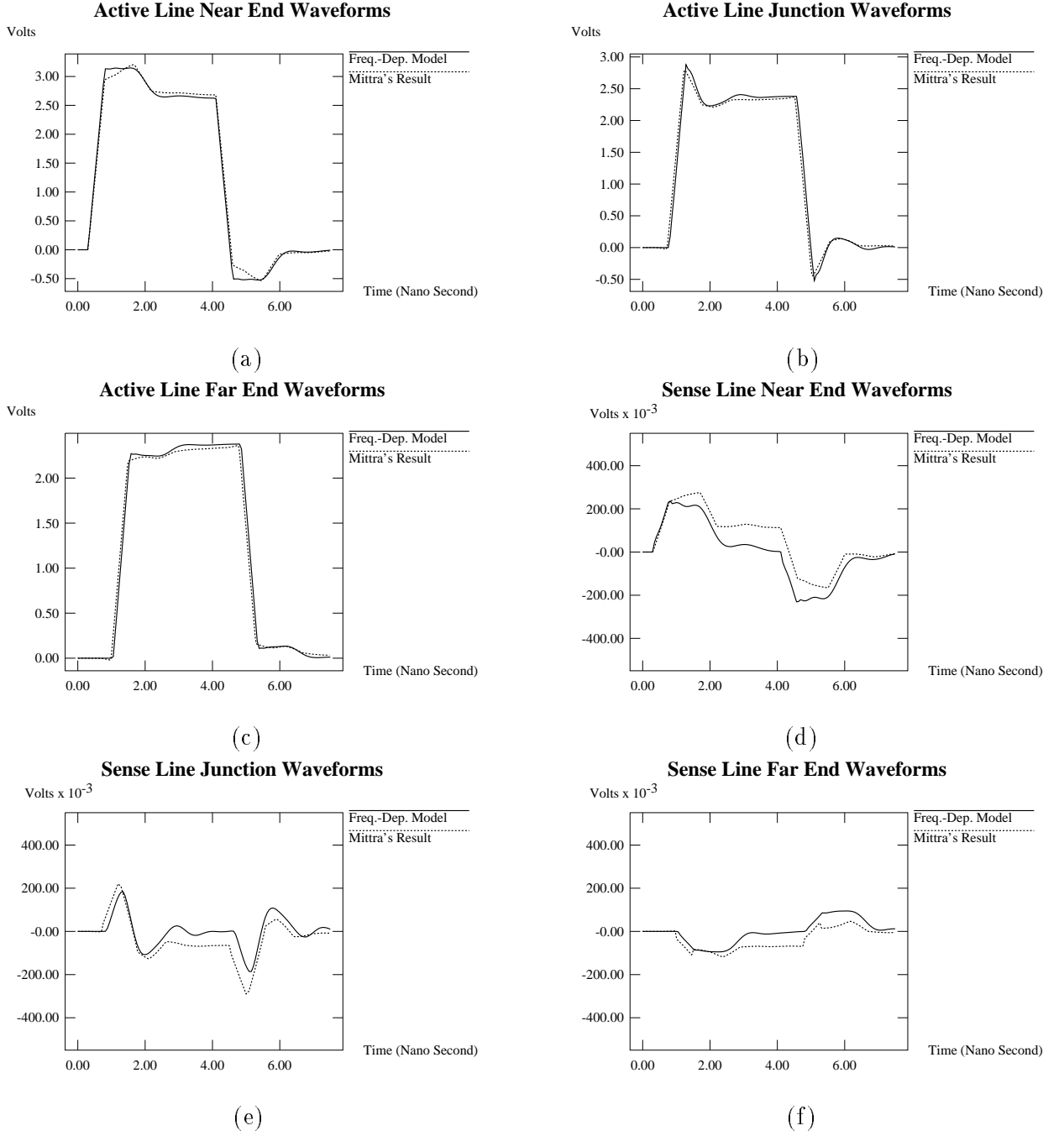


Figure 3.4: **Simulation Waveforms of the Coupled Transmission Line Characterized with Frequency-Dependent Losses:** The waveforms are obtained from the S-parameter macromodel based simulator and Mittra's "mtltda" simulator. The output waveforms of the near end of the active line is shown in (a), waveforms of the junction of the active line is shown in (b), waveforms of the far end of the active line is shown in (c), waveforms of the near end of the sense line is shown in (d), waveforms of the junction of the sense line is shown in (e), waveforms of the far end of the sense line is shown in (f).

in order to conserve space. The per-unit-length conductance is assumed to be zero. The driving signal is 200-MHz, 50% duty-cycle pulse with  $0.5ns$  rise/fall time. The circuit schematic is shown in 3.5 (a) with the component values. The simulation waveforms of this example are shown in Figure 3.5 (b), (c), (d) and (e).

Similar to Example 3, for all of the figures, there are two output waveforms, they are:

- the output waveform of the frequency-dependent macromodel.
- the output waveform obtained from Mittra's "mtltda" simulator.

In this particular example, increasing the order of Pade approximation does not imply an increase in accuracy; one must adopt other method such as Complex-Frequency-Hopping (CFH) or Pade-via-Lanczos (PVL) to solve this accuracy problem.

The S-parameter macromodel based simulator takes 9.23 seconds on the SUN Sparc1+ workstation, whereas the time-domain simulators takes 288 seconds on the same machine. This difference in simulation time is due to the fact that the second data point is taken at  $1MHz$  instead at  $20MHz$ . In order to use the time-domain simulator, the actual data taken at  $10KHz$  and  $100KHz$  are intentionally removed. Mittra's "mtltda" simulator generates better results, however, it suffers from the common drawback of long simulation time like other time-domain simulators.





## CHAPTER 4. Transient Analysis of Interconnects Characterized with Measured S-Parameter Data

Integrating the measured models into a moment-matching simulator is a problem because derivatives of the frequency dependent measured data which are required to generate the moments are not available explicitly. Generating the required moments using numerical differentiation can often lead to large computation errors [47]. Since the information necessary for the construction of the scattering parameter macromodel does not exist explicitly, a new method for retrieving the information needed must be developed with special attention paid to avoid numerical instabilities.

Gruodis et al. has previously reported the measurements of the admittance matrix  $\mathbf{Y}_{2n}$  and the impedance matrix  $\mathbf{Z}_{2n}$  of the transmission lines, the computation of the  $\mathbf{Y}_0$  matrix propagation constant  $\Gamma$ , and the simulation using the state variable transfer function method [29]. The deficiency associated with Gruodis' method was the fact that the admittance matrix  $\mathbf{Y}_{2n}$  and the impedance matrix  $\mathbf{Z}_{2n}$  were difficult to measure at a frequency above several MHz (as is specifically stated in Gruodis et al. paper [29].) Unlike the impedance matrix  $\mathbf{Z}_{2n}$  and the admittance matrix  $\mathbf{Y}_{2n}$  which may not exist for the cases of the serial circuits or shunt circuits respectively. The S-parameter matrix always exists for any physical circuits [47] [50]. So the choice is to measure the S-parameter data for the transient simulation. Sanaie et al. used the balance-reduction method [47] and Silveria et al. utilized the curve-fitting of the transfer function section-by-section [50]. Kuznetsov et al. used the direct rational function curve-fitting of the wave propagation function [34]. All

of them try to extract information from the curve-fitting of the S-parameter data directly. Thus Silveria et al. suggested performing curve-fitting section-by-section with each section no larger than a decade [50]. Although accurate, this section-by-section method generates more information than necessary. This leads to Silveria's proposal to use a balance-reduction method to reduce the order. Chang et al. has implemented a direct convolution method into HP Spice [51]. This method requires the impulse response found from inverse Fast Fourier Transform (IFFT) of the measured S-parameter data. Chang et al. further improves their method using direct fitting of the measured data with a rational function [10]. Hu et al. has incorporated the recursive convolution method into SWECC [32]. This method also calls for the impulse response found through IFFT. The drawback of convoluting the impulse response is the assumption made about the relationship between the real part and the imaginary part of all the S-parameters. The implementation usually requires the data of the imaginary part to be thrown away, whereas the Gruodis method of converting the measured S-parameter data into the parasitic functions makes use of all the data on hand.

Although the direct curve-fitting methods could be very accurate interpolation tools, to find the moments of the measured S-parameter data still requires an extrapolation method which at least extrapolate to  $s = 0$ . Extrapolation methods must preserve the periodic nature of the measured S-parameter data not only within the spectrum of the measurements but also in the spectra above and below.

By the definition of the Taylor series expansion, the moments of a Taylor series expansion are the coefficients of the polynomial of  $s$ . These coefficients are computed from the derivatives of different orders at a given expansion point. The number of terms in a Taylor series is called the expansion order. Another major drawback of the direct curve-fitting methods is the moments that are produced by these methods are often expansion order dependent. This is in violation of the definition given by the Taylor series expansion which states that the moments should be expansion order independent.

The purpose of this part of the thesis is to propose an indirect method of finding the moments from the measured S-parameter data. This novel indirect method of computing

the moments calls for a conversion of the S-parameter data into the parasitic functions before finding the Taylor series expansions. After the S-parameters have been converted into the parasitic functions,  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ , one can compute the S-parameter from these four data sets. A least-square curve-fitting of the computed S-parameter with respect to the measured S-parameter data, is used in order to find the moments of these four parasitic functions. The exact moments of the S-parameter functions are subsequently computed from the moments of these four parasitic functions. Based on the exact moments found using the indirect approach, the macromodel of transmission lines characterized with frequency-dependent losses is constructed and the transient simulation is performed. Currently, this indirect approach handles the measured S-Parameter data of the interconnects only. It cannot handle the general black box characterized by the measured S-parameter data.

The motivation of taking this indirect approach is shown in Section 4.1. The general steps of this novel indirect approach will be presented in Section 4.2, whereas in Section 4.3, two methods to convert the measured S-parameter data into the parasitic functions will be shown. Section 4.4 will present the algorithm for the finding of the moments of the parasitic functions through the indirect curve-fitting method.

## 4.1 Motivations

The most compelling reasons to use this indirect approach to find the moments of the measured S-parameters data are:

- finding the moments of the measured S-Parameter data requires extrapolation.
- the moments in a Taylor series expansion should be order independent.

The measured S-parameter data curves demonstrate the periodic nature of the S-parameter data. This periodic nature should be preserved not only within the sampling spectrum but also in the spectra above and below. This characteristic can only be obtained by a well behavior extrapolation method.

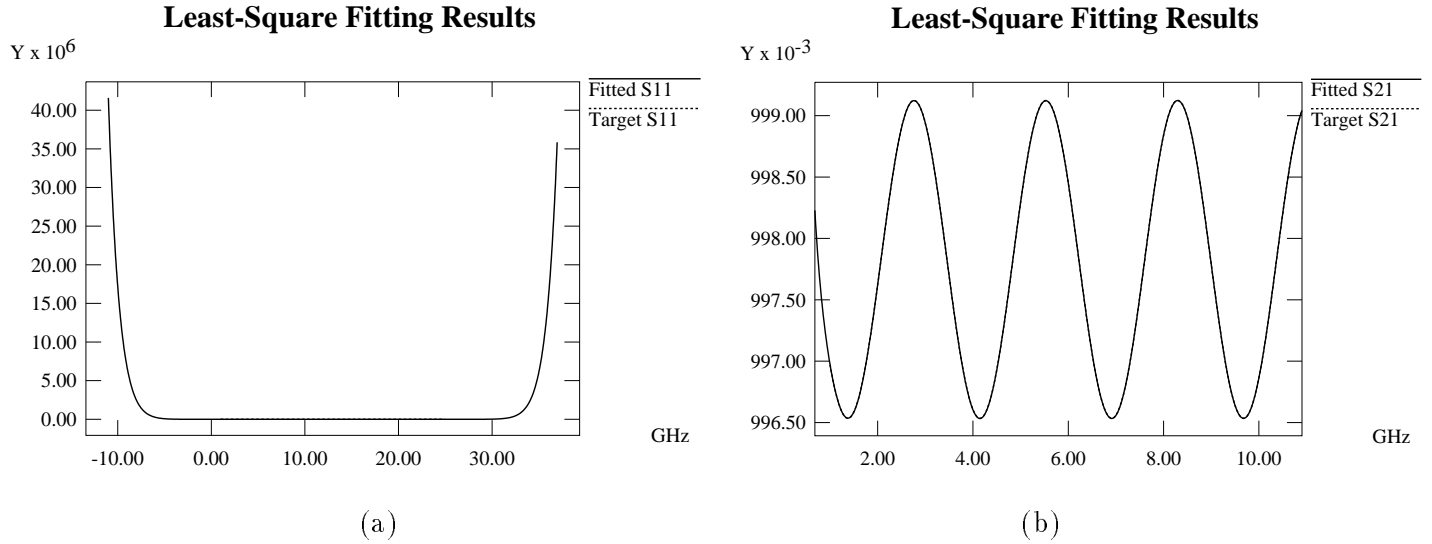


Figure 4.1: **A Particular Curve-Fitting Example:** The curves in (a) is plotted against extended frequency range beyond the sampled frequency spectrum. The curves in (b) is plotted against only the sampled frequency spectrum which spans from 1.0 GHz to 11.5 GHz.

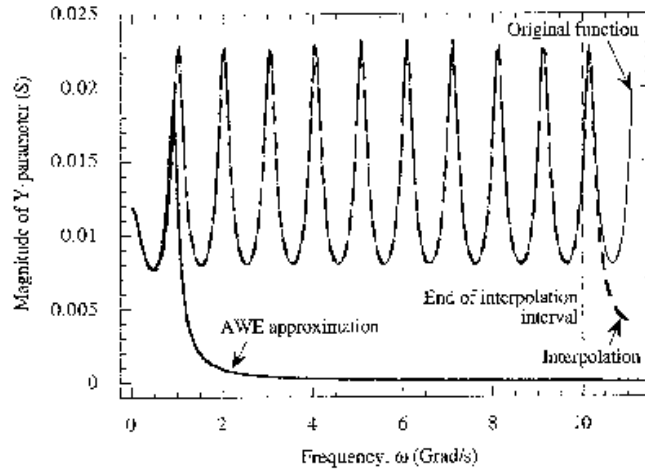


Figure 4.2: **AWESim and Rational function Curve-Fitting Comparison:** AWESim is proven inadequate to be used in curve-fitting and interpolation. Direct fitting to rational function is good only for use as interpolation tool.

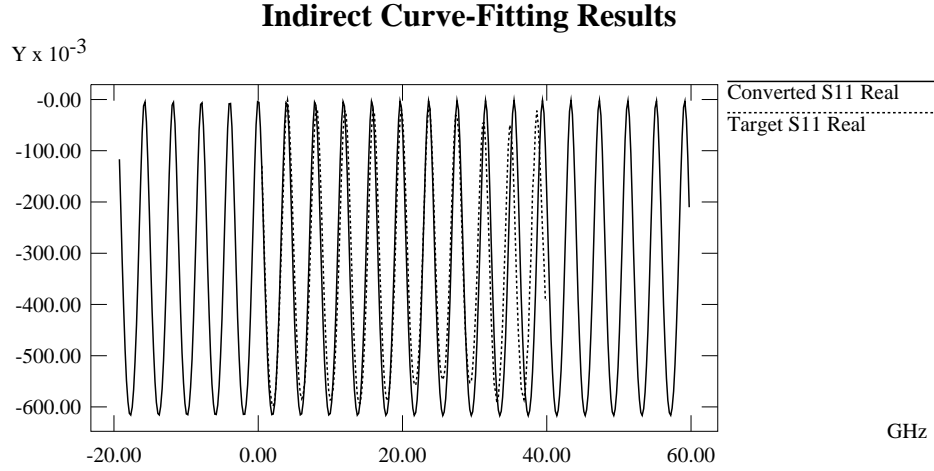


Figure 4.3: **Indirect Curve-Fitting Example:** The indirect method produces bounded and periodical computed S-parameter data both inside and outside of the sampling spectrum which spans from 0.5 GHz to 40.5 GHz. It is good both as an interpolation and extrapolation tool.

The direct curve-fitting methods are good interpolation tools which could outperform the indirect approach presented here in the interpolation task. However, these direct curve-fitting methods fail poorly in the extrapolation task and thus cannot be used to find the accurate moments of the measured S-parameter data.

For example, the Root-Mean-Square (RMS) error for the Least-Square curve-fitting within the sampling spectrum could be very small, but can be unbounded outside of the sampling spectrum. Figure 4.1 (a) shows a particular example where the RMS error of the Least-Square curve-fitting equals to 0.0245 and the resulting function is unbounded outside of the sampled spectrum. Figure 4.1 (b) shows how well the least-square curve-fitting performs over the sampled spectrum. The least-square curve-fitting method is known for its inability to handle the measured data with many periods of oscillations. Miguel Silveria's section-by-section method is an improved least-square curve-fitting method which can be used to solve the above problem [50]. However, his method is still an interpolation tool and hence cannot be used to perform extrapolation to find the moments of the measured S-parameter data. Figure 4.2 shows the AWESim curve-fitting characteristic and the direct rational function curve-fitting results [35]. AWESim is not suitable for the curve-fitting

to functions with many periods of oscillations. The direct rational function curve-fitting proposed by Kuznetsov et al. is good for use as an interpolation tool, but cannot be used in extrapolating. Clearly, its accuracy cannot be relied on beyond the interpolation interval. A state-of-the-art direct rational function curve-fitting method which could outperform the indirect approach which is presented in here both in the interpolation and extrapolation [10]. Dr. Lee Barford of HP uses this direct curve-fitting method to find the poles and the residues of the measured S-parameter data for their recursive convolution simulation. This direct rational function curve-fitting method is a commercial product as part of the Matlab software package published by The MathWorks, Inc. However, this direct rational function curve-fitting method still produces moments that are expansion order dependent. Figure 4.3 shows that the indirect method produces the computed the S-parameters data which are not only bounded but also preserve the periodic nature above and below the spectrum of the measured frequencies.

The lower order moment terms in a Taylor series expansion around a given point should be expansion order independent. It does not hold true for all of the direct curve-fitting methods. Computing the S-parameter moments from parasitic functions does guarantees that the moments found are independent of the order of the approximation.

The “jitters” in the measured S-parameter data usually comes from the instrument limitations and human errors, and requires many measurements to average them out. Due to the smooth nature of the S-parameter computed from the analytic equations, there is no jitter shown in the computed S-parameter. Fitting the computed S-parameter data to the measured S-parameter data will not be affected by the small measurement error jitter abundant in the examples tested.

Sometimes the measured S-parameter data can rise above 1.0 for a passive system. Dr. Lee Barford of Hewlett-Packard Laboratories states:“ It is quite common to collect measured S-parameters that are non-passive. It occurs at any frequency where the device being measured has very low loss. This is the case for any transmission line or interconnect below its cutoff frequency. Measurement noise may make the device appear to be very

slightly active instead of very slightly passive.” [1] With the computed S-parameters which are evaluated from the parasitic functions, they never violate the power conservation rule and hence are always stable for a passive system.

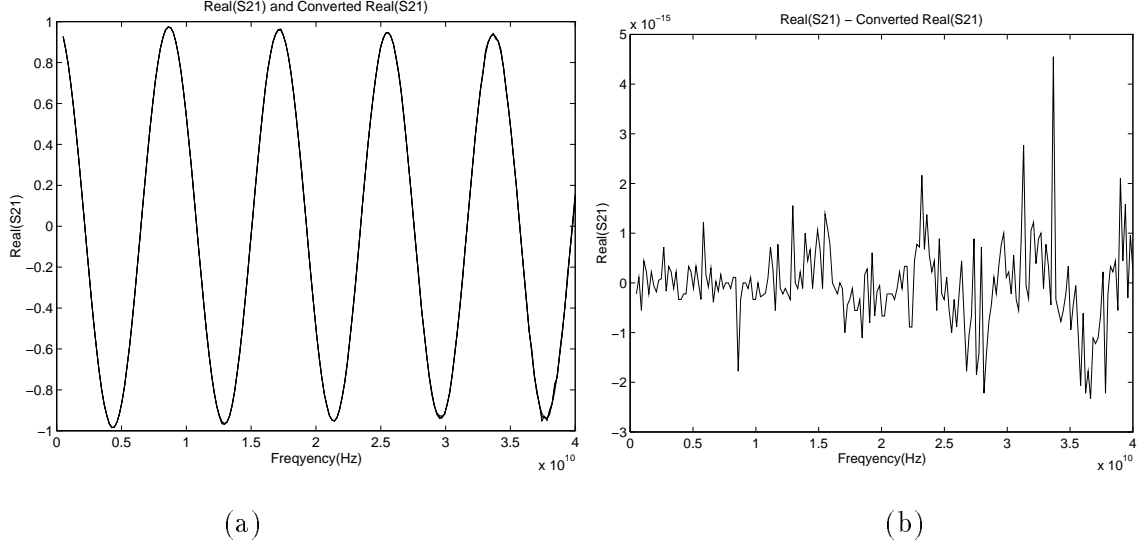


Figure 4.4: **The Introduced Conversion Error:** The original measured S-parameter data and the converted S-parameter are plotted in (a). The difference between them are plotted in (b).

The conversion processes for both the two-port measured data and  $2N$ -port measured data shown in Section 4.3 are almost lossless. The regenerated S-parameters based on the converted **RLCG** compared with the original S-parameter data show that this conversion process introduces very little error. Figure 4.4 demonstrates the conversion error is less than  $10^{-14}$ .

As a by-product, this indirect method can also be used to predict measurement errors. One can use the converted parasitic to compute the S-parameters. When compared to the measured S-parameters, one can immediately identify those sampling points with large measurement errors (Situations such as the S-parameter value is greater than 1.0). The general practice is to discard the S-parameter data of those points to avoid their influence in the final fitting results. However, if there are a large number of points with error, it is advisable that the measurement be repeated.



## 4.2 Proposed Method

The steps required for the transient analysis of the interconnect characterized with the measured S-parameter data macromodel include:

- Converting measured S-parameter data into parasitic functions:  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ .
- Finding moments of the four parasitic functions through least-square curve-fitting of the computed S-parameter w.r.t. the measured S-parameter data.
- Computing moments of the S-parameters from the moments of the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  data sets.
- Constructing macromodels for the congruence transformers for the  $2N$ -port coupled interconnects.
- Incorporating the computed S-parameter macromodels into the S-parameter macromodel based simulator to perform the transient analysis.

## 4.3 Converting Measured S-Parameter Data into Parasitic Functions

For the completeness of this thesis, the methods for converting two-port and  $2N$ -port S-parameter data into the Parasitic functions are presented here.

Previous research that converted the measured two-port S-parameter data into the **RLCG** data sets include those of Eisenstadt et al. [21] and Owazr [45]. Their primary focus was on the interconnect characterization and not on the transient simulation. Also, only a two-port device, and not the  $2N$ -port device, was described in their papers.

The following Equations are taken from the Eisenstadt et al. paper to demonstrate the steps taken to convert the S-parameters into  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  [21]. Starting with the S-parameter matrix of a lossy transmission line which is presented in Equation (4.1).

$$S(s) = \frac{1}{2Z_0 Z_c(s) \cosh(\gamma(s)) + (Z_c^2(s) + Z_0^2) \sinh(\gamma(s))}$$

$$\begin{bmatrix} (Z_c^2(s) - Z_0^2) \sinh(\gamma(s)) & 2Z_0 Z_c(s) \\ 2Z_0 Z_c(s) & (Z_c^2(s) - Z_0^2) \sinh(\gamma(s)) \end{bmatrix}, \quad (4.1)$$

The  $Z_c(s)$  is the characteristic impedance and  $\gamma(s)$  is the propagation constant. Both  $Z_c(s)$  and  $\gamma(s)$  are computed from the frequency-dependent values of  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  based on the following Equations:

$$\begin{aligned} \gamma(s) &= \sqrt{(R(s) + sL(s))(G(s) + sC(s))} \cdot l \\ Z_c(s) &= \sqrt{\frac{R(s) + sL(s)}{G(s) + sC(s)}}. \end{aligned}$$

It can be shown that:

$$e^{-\gamma l} = \left\{ \frac{1 - S_{11}^2 + S_{21}^2}{2S_{21}} \pm K \right\}^{-1}, \quad (4.2)$$

where

$$K = \left\{ \frac{(1 - S_{11}^2 + S_{21}^2)^2 - (2S_{11})^2}{(2S_{21})^2} \right\}^{\frac{1}{2}} \quad (4.3)$$

$$Z_c^2 = Z_0^2 \frac{(1 + S_{11})^2 - S_{21}^2}{(1 - S_{11})^2 - S_{21}^2}. \quad (4.4)$$

During the extraction of  $\gamma$ , one must choose the root with the positive attenuation factor  $\alpha$  and correct the negative propagation  $\beta$  into a positive value. After extracting the  $\gamma$  and  $Z$  from  $e^{-\gamma l}$  and  $Z_c^2$ , the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  can be found from the following Equations:

$$\mathbf{R}(f) = \text{Re}\{\gamma Z_c\} \quad (4.5)$$

$$\mathbf{L}(f) = \text{Im}\{\gamma Z_c\} / \omega \quad (4.6)$$

$$\mathbf{G}(f) = \text{Re}\{\gamma / Z_c\} \quad (4.7)$$

$$\mathbf{C}(f) = \text{Im}\{\gamma / Z_c\} / \omega. \quad (4.8)$$

The conversion for  $2N$ -port measured S-parameter data requires a different approach which must handle  $\mathbf{RLCG}$  matrices. Gruodis et al. has previously reported the conversion of the measured admittance and the measured impedance data into  $\mathbf{RLCG}$  data sets

[29]. Applying the standard equations found in microwave textbooks [20], one can find the admittance matrix  $\mathbf{Y}_{2n}$  and impedance matrix  $\mathbf{Z}_{2n}$  from the measured S-parameter data matrix  $\mathbf{S}_{2n}$ , and then apply Gruodis et al. method to find the parasitic function matrices:  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ .

Assume there only exists the TEM or Quasi-TEM mode of wave propagation, for the  $(n+1)$  conductor coupled transmission lines, there are  $2N$  ports. The differential equations in the frequency domain are [29]:

$$\frac{\partial}{\partial x} \begin{bmatrix} \mathbf{V}_S \\ \mathbf{V}_R \end{bmatrix} = -\mathbf{Z} \begin{bmatrix} \mathbf{I}_S \\ \mathbf{I}_R \end{bmatrix} \quad (4.9)$$

$$\frac{\partial}{\partial x} \begin{bmatrix} \mathbf{I}_S \\ \mathbf{I}_R \end{bmatrix} = -\mathbf{Y} \begin{bmatrix} \mathbf{V}_S \\ \mathbf{V}_R \end{bmatrix}, \quad (4.10)$$

where  $\mathbf{V}_S$  and  $\mathbf{V}_R$  are the  $n$  by 1 column vector for line voltage w.r.t. the reference conductor on the sending and the receiving side,  $\mathbf{I}_S$  and  $\mathbf{I}_R$  are the  $n$  by 1 column vector for line current on both sides,  $\mathbf{Z} = \mathbf{R} + jw\mathbf{L}$ , and  $\mathbf{Y} = \mathbf{G} + jw\mathbf{C}$ .  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  are  $n$  by  $n$  symmetric matrices in which all elements are functions of frequency but independent of  $x$ .

Following the same assumptions as proposed in Gruodis et al.'s paper, that there exists a complex square root matrix  $(\mathbf{ZY})^{1/2}$ , the solutions of the above differential equations become [29]:

$$\begin{bmatrix} \mathbf{I}_S \\ \mathbf{I}_R \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_0 \coth \Gamma d & -\mathbf{Y}_0 \csc h \Gamma d \\ -\mathbf{Y}_0 \csc h \Gamma d & \mathbf{Y}_0 \coth \Gamma d \end{bmatrix} \begin{bmatrix} \mathbf{V}_S \\ \mathbf{V}_R \end{bmatrix}. \quad (4.11)$$

where

$$\mathbf{\Gamma} = (\mathbf{ZY})^{1/2} = \mathbf{P}\mathbf{\gamma}\mathbf{P}^{-1}, \quad (4.12)$$

$$\mathbf{Y}_0 = \mathbf{Z}^{-1}\mathbf{\Gamma} = \mathbf{Y}\mathbf{\Gamma}^{-1}. \quad (4.13)$$

$\mathbf{P}$  is the eigenvector matrix of  $\mathbf{\Gamma}$  as well as the eigenvector matrix of the  $\mathbf{ZY}$  product, and  $\mathbf{\gamma}$  is the diagonal eigenvalue matrix of  $\mathbf{\Gamma}$ .

Given the measured  $2N$  port S-parameter data  $\mathbf{S}_{2n}^{measd}$ , the symmetry of the S-parameter matrix can be assured by taking the arithmetic average as follows:

$$\mathbf{S}_{2n} = \frac{1}{2} \left[ (\mathbf{S}_{2n}^{measd}) + (\mathbf{S}_{2n}^{measd})^T \right]. \quad (4.14)$$

From standard microwave textbooks [20], one can find:

$$\mathbf{Y}_{2n} = Z_0^{-1}([\mathbf{I}] - \mathbf{S}_{2n})([\mathbf{I}] + \mathbf{S}_{2n})^{-1} \quad (4.15)$$

$$\mathbf{Z}_{2n} = Z_0([\mathbf{I}] + \mathbf{S}_{2n})([\mathbf{I}] - \mathbf{S}_{2n})^{-1}, \quad (4.16)$$

where  $Z_0$  is the reference impedance.

After securing the  $\mathbf{Y}_{2n}$  and  $\mathbf{Z}_{2n}$  matrices, one can use the data analysis method 2 in Gruodis et al. paper to give [29]:

$$\begin{aligned} -\mathbf{Y}_{21}^{-1}\mathbf{Y}_{11} &= \cosh \Gamma d \\ \Gamma d &= \mathbf{P}[\cosh^{-1}\mathbf{\Lambda}_{-\mathbf{Y}_{21}^{-1}\mathbf{Y}_{11}}]\mathbf{P}^{-1}, \end{aligned} \quad (4.17)$$

where  $\mathbf{P}$  is the eigenvector matrix of  $-\mathbf{Y}_{21}^{-1}\mathbf{Y}_{11}$  as well as the eigenvector matrix of  $\Gamma d$ . From Equations (4.17) and  $\mathbf{Y}_{21} = -\mathbf{Y}_0 \cosh \Gamma d$ , one has:

$$\begin{aligned} \mathbf{Y}_0 &= -\mathbf{Y}_{21} \sinh \Gamma d \\ &= -\mathbf{Y}_{21} \mathbf{P}[\sinh(\cosh^{-1}\mathbf{\Lambda}_{-\mathbf{Y}_{21}^{-1}\mathbf{Y}_{11}})]\mathbf{P}^{-1}. \end{aligned} \quad (4.18)$$

Having derived  $\Gamma d$  and  $\mathbf{Y}_0$ , one can use Equation (4.13) for the following:

$$\mathbf{Z} = \mathbf{R} + j\omega\mathbf{L} = \Gamma d \cdot \mathbf{Y}_0^{-1} \quad (4.19)$$

$$\mathbf{Y} = \mathbf{G} + j\omega\mathbf{C} = \mathbf{Y}_0 \cdot \Gamma d. \quad (4.20)$$

The  $\mathbf{R}_{ij}$ ,  $\mathbf{L}_{ij}$ ,  $\mathbf{C}_{ij}$  and  $\mathbf{G}_{ij}$  can be found from the following Equations:

$$\mathbf{R}_{ij}(f) = \text{Re}\{\mathbf{Z}_{ij}\} \quad (4.21)$$

$$\mathbf{L}_{ij}(f) = \text{Im}\{\mathbf{Z}_{ij}\}/\omega \quad (4.22)$$

$$\mathbf{G}_{ij}(f) = \text{Re}\{\mathbf{Y}_{ij}\} \quad (4.23)$$

$$\mathbf{C}_{ij}(f) = \text{Im}\{\mathbf{Y}_{ij}\}/\omega. \quad (4.24)$$

#### 4.4 Finding the Moments of the Four Parasitic Functions

The key contribution of this part of the thesis is to find the moments of the parasitic functions indirectly. The indirect approach proposed here is to find the moments of the measured S-parameter data indirectly through converting the strongly frequency-dependent S-parameter data into parasitic functions:  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  data sets. When taking advantage of the fact that the  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  are weak functions of frequency and the  $\mathbf{R}(f)$  is a strong function of frequency [2], the moments of these four parasitic functions are much simpler to find.

The input is the measured S-parameter data tabulated w.r.t. the sampling frequencies. The output is the moments of the parasitic functions.  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ . These moments are the coefficients of the Taylor series expansion around  $s = 0$  for the parasitic functions: First, one converts the measured S-parameter data into frequency-dependent parasitic data sets. Then a least-square curve-fitting is performed to find the coefficients of the polynomials which represent the parasitic data sets. These coefficients are used as the initial values for the following optimization procedure. The S-parameter data sets are computed from the moments of the parasitic functions based upon analytic equations. The Levenberg-Marquardt optimization method is used to find the moments of the parasitic functions through comparing the computed S-parameter and the original measured S-parameter data [16]. The partial derivatives of the least-square error between these two S-parameter data sets w.r.t. each of the moments of the parasitic functions are obtained using a perturbation method [43]. Assume the initial coefficients are  $C^{(0)}$ , and there are total  $m$  coefficients. The next coefficients  $C^{(k+1)}$  can be computed from the previous coefficients  $C^{(k)}$ , which according to Levenberg-Marquardt, is optimized according to the following formula [42] [16]:

$$C^{(k+1)} = C^{(k)} - (J^T J + \lambda \Lambda)^{-1} J^T \left[ \frac{\partial LSE}{\partial c_j} \right], \quad (4.25)$$

where  $k$  is the number of iteration,  $C^{(k)}$  the column vector of the  $k$ -th iteration, and  $LSE$  the least-square error.  $J$  is the sensitivity matrix,  $J^T$  is the transposition vector of  $J$  where

the  $j$ -th element  $J^T(j) = J(j)$ ,  $\Lambda$  is a value which are equal to the product of  $J^T J$ , and  $\lambda$  is the *Lagrange Multiplier* properly selected to speed up the convergence of the optimization process [42] [16].  $J^T \left[ \frac{\partial LSE}{\partial c_j} \right]$  represents the gradient around the current coefficients  $C^{(k)}$ . To obtain the sensitivity matrix  $J$ , the  $j$ -th element is defined as:

$$J(j) = \frac{\partial LSE}{\partial c_j}, \quad 1 \leq j \leq m. \quad (4.26)$$

The partial derivatives are computed using a central difference method. The optimization continues until the least-square error can no longer be improved, or the iteration number exceeds a preset limit. The convergence to the optimal values of Levenberg-Marquardt method is proved in [42].

#### 4.5 Creation of the S-Parameter Macromodel

Given the moments of the four parasitic functions  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$ , one can compute the moments of the S-parameters using the methods outlined in Section 3.1. Once found, the same macromodels are integrated into the S-parameter macromodel based simulator to perform the transient analysis.

#### 4.6 Experimental Results for Measured Scattering Parameter Data

The measured S-parameter data files are courtesy of HP Santa Rosa division, Hewlett Packard Company. The measured S-parameter data in all of the examples are not given in order to conserve space. The driving signal is 1-GHz, 50% duty-cycle pulse with 0.1ns rise/fall time. All of the drivers are modeled with a piecewise-linear input voltage source in series with the parallel combination of a 25Ω resistor and a 4.3pF capacitor. The receivers in all of the examples are modeled using the parallel combination of a 100Ω resistor and a 1.0pF capacitor. For all of the simulation results of the S-parameter macromodels, all of the far-end waveforms are simulated with the time-of-flight captured explicitly with the method stated in Section 5.5. The testing circuits are the same for all of the examples. The circuit schematic is shown in Figure 4.5. For all of the figures of the simulation results,



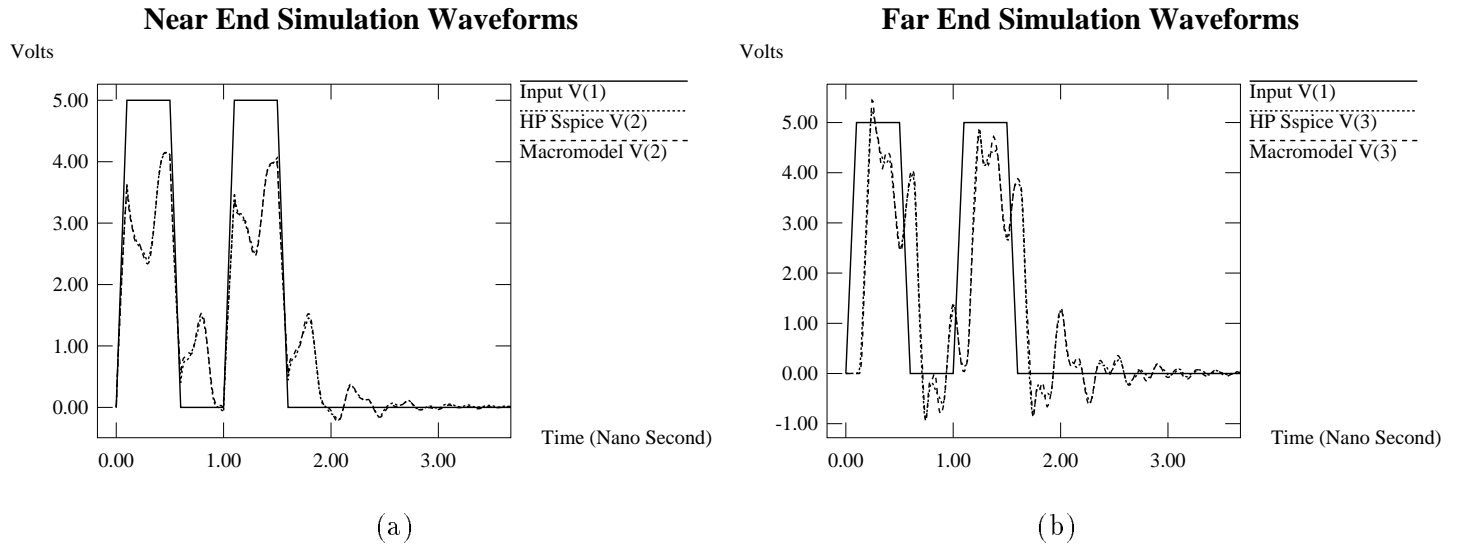


Figure 4.6: **Simulation Waveforms of the Step Structure Characterized with Measured S-Parameter Data:** The output waveforms of the near end of the two-port is shown in (a), waveforms of the far end of the two-port is shown in (b).

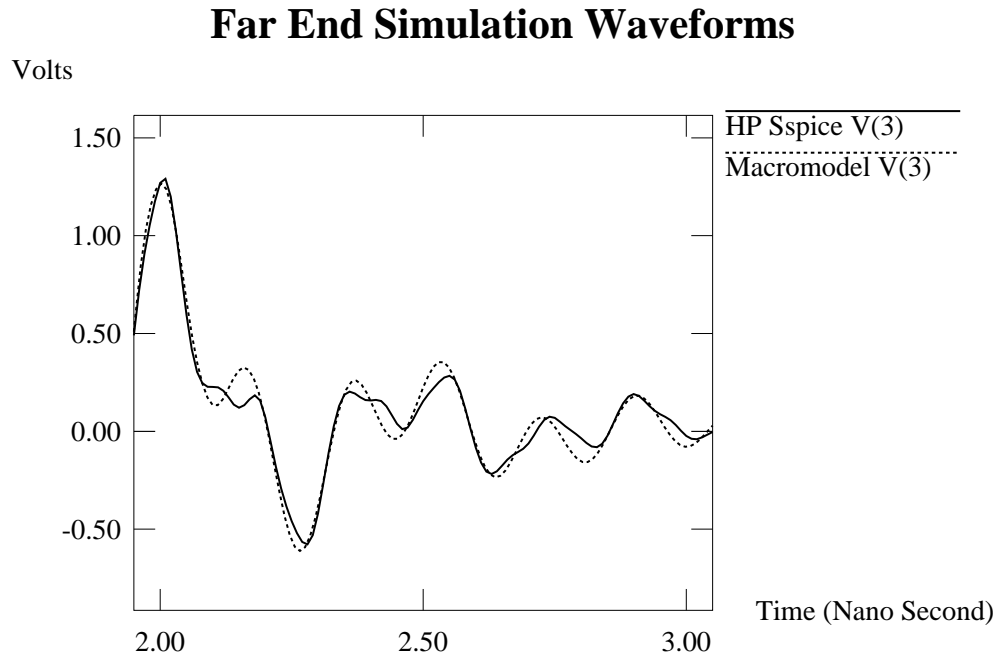


Figure 4.7: **Enlarged Far End Simulation Waveforms:** The two closely follow each other are the simulation waveforms of the HP Spice and the S-parameter based Macromodel simulator.



### 4.6.2 Example 2: An Uniform Microstrip Structure

This is an example with an uniform transmission line characterized with measured S-parameter data. It is simulated using the same test circuit and input driving signal. The near end and far end simulation waveforms of the HP Spice simulator and the S-parameter macromodel based simulator are shown in Figure 4.8 (a) and Figure 4.8 (b) respectively.

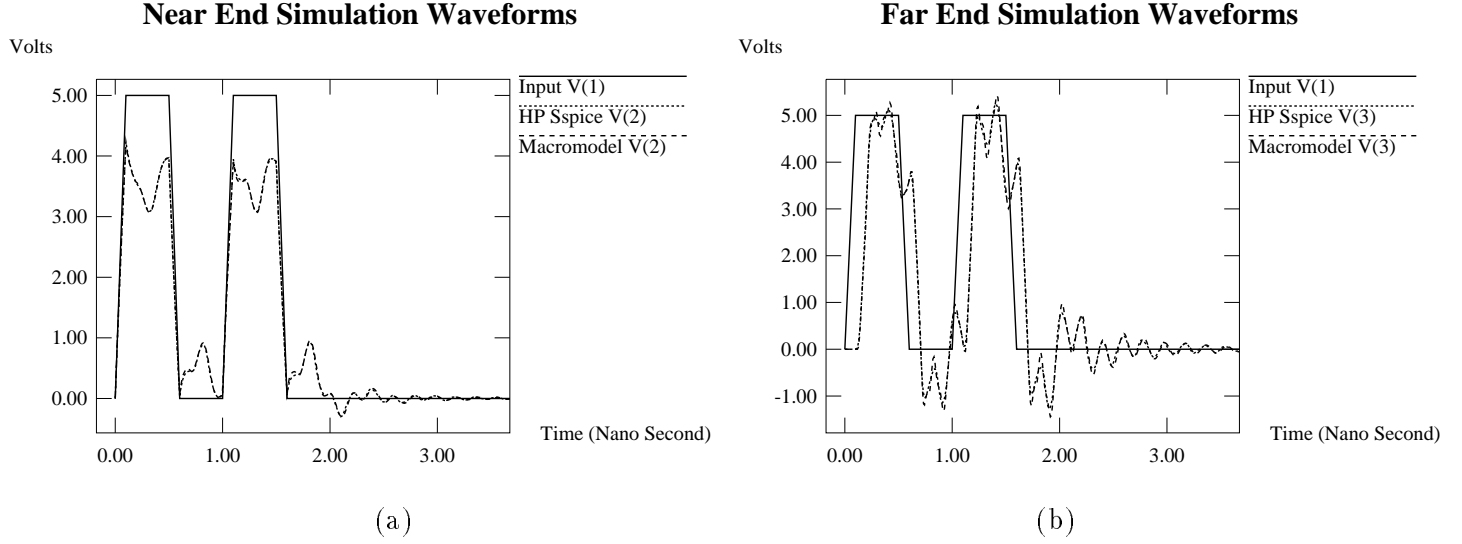


Figure 4.8: **Simulation Waveforms of the Uniform Microstrip Characterized with Measured S-Parameter Data:** The output waveforms of the near end of the two-port is shown in (a), waveforms of the far end of the two-port is shown in (b).

### 4.6.3 Example 3: A Bend Structure

This is an example with two uniform transmission line sections joined by a  $90^\circ$  bend, and the whole structure is characterized with measured S-parameter data. It is simulated using the same test circuit and input driving signal. The near end and far end simulation waveforms of the HP Spice simulator and the S-parameter macromodel based simulator are shown in Figure 4.9 (a) and Figure 4.9 (b) respectively.

All of the results show that the output waveforms of the measured S-parameter macromodel match well with those produced by HP Spice simulator as was expected. The small

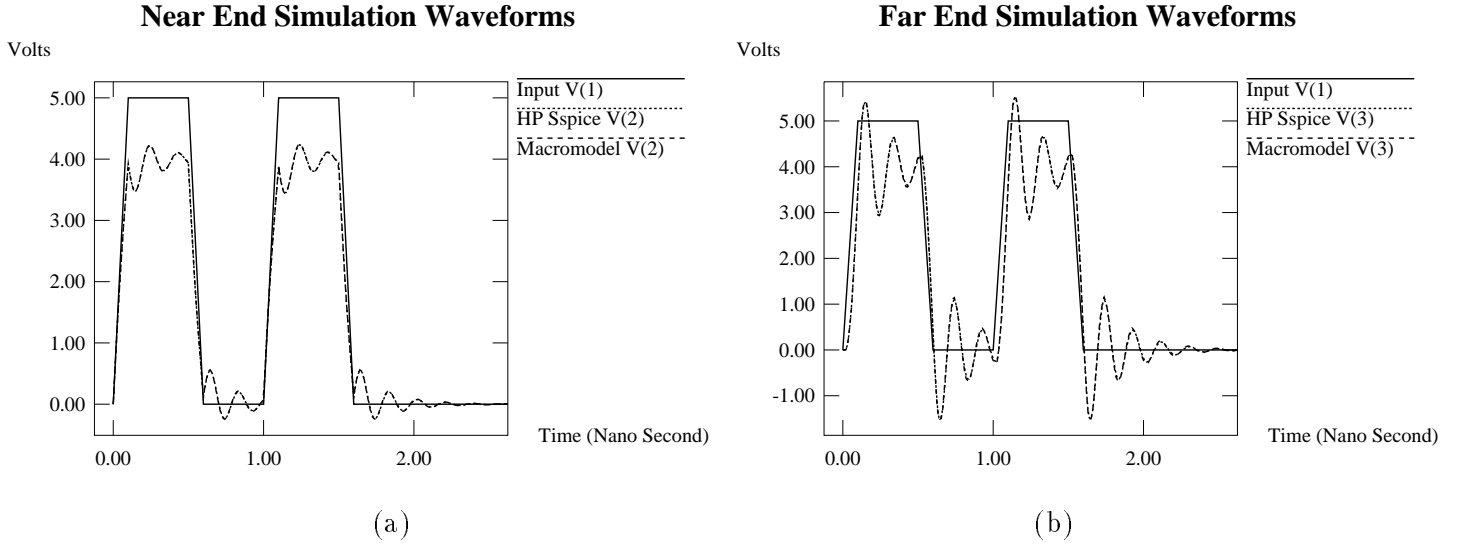


Figure 4.9: **Simulation Waveforms of the Bend Structure Characterized with Measured S-Parameter Data:** The output waveforms of the near end of the two-port is shown in (a), waveforms of the far end of the two-port is shown in (b).

differences that were shown on both the near end and far end waveforms are due to the lower number of order chosen and the Pade approximation error. In all of the examples, increasing the order of Pade approximation does not imply an improvement in accuracy. One must turn to other methods such as the Pade-via-Lanczos (PVL) to solve this accuracy problem.

## CHAPTER 5. Optimal Design of Self-Damped Interconnects for Multichip Modules

The metal line width on MCM are much wider than those of modern VLSI chip. As a result, the line can exhibit a non-negligible inductance. When compared with the metal lines on PCB, these metal lines on MCM exhibit a much larger resistance per-unit-length due to a smaller cross-section. Thus, the metal lines on MCM must be treated as lossy transmission lines.

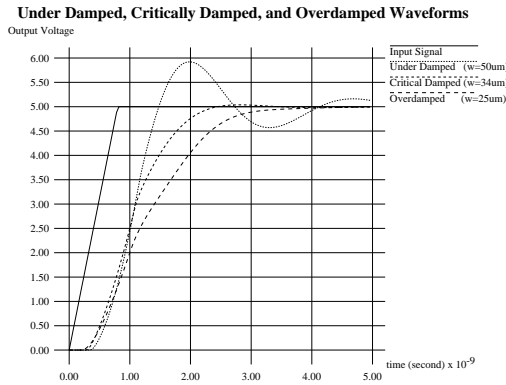


Figure 5.1: **Simulation Waveforms of Different Line Widths:** The different damping conditions are achieved through sizing the width of the interconnect to:  $50\mu m$ ,  $34\mu m$ , and  $25\mu m$  respectively.

The interconnection lines on multichip modules exhibit reflections and resonances due to its transmission line characteristics and are therefore usually terminated by two methods: one is to terminate the lines with clamping diodes and the other is to terminate with resistors. In the former method, when a line on a multichip module is terminated by a pair of voltage-clamping diodes to limit the positive and negative signal swings, the diode when turned on by signal voltage overshoot can drive a large current, which can often

cause an increase in power consumption and an increase in power-distribution disturbance. On the other hand, when a line on a multichip module is terminated with a resistor of appropriate value to minimize signal reflections and resonances, the voltage divider formed by the termination resistance and the line characteristic impedance can cause a substantial and unacceptable attenuation of the propagated signals. In addition, each resistor can dissipate a quiescent power when the line is at non-zero voltage. In modules containing thousands of such lines, this power dissipation can become excessive. The problem can be further compounded by the small packaging size of MCM with a limited heat removal capability.

In order to propagate high speed signals, the lines on thin-film substrates are often unterminated. Since voltage doubling occurs at the end of an open line due to reflection, a controlled amount of attenuation can be tolerated. Instead of terminating with diodes or resistors, the long lines on the thin-film multichip modules can be constructed to critically damp the signal in order to avoid resonances. Doing without the termination eliminates the heat generated by the terminator and solves the heat removal problem. Unterminated, the long lines in the substrate are structured to exhibit a total resistance that is related to the source resistance of the active devices which drives the lines [31] [6]. For performance reasons, *slightly* under-damped designs give shorter signal propagation delays with a tolerable amount of overshoot [6] [23] [31]. These lines are thus called *optimal self-damped lossy transmission lines* [25].

## 5.1 Wire-Sizing

In order to have a MCM design that meets all the electrical performance requirements, one needs to control the delay, the noise, and the impedance through geometry parameters. It is easiest to fabricate interconnection lines with fixed metal and dielectric thicknesses and vary only the line width. This results in a specific dependency of line width on length for self-damped lines. The *wire-sizing* method is often used to vary the width of each branch of the network to meet certain electrical criteria. This method is used to design the

self-damped transmission lines on MCMs.

Different line widths result in different output waveforms. Figure 5.1 shows the under-damped, critically damped, and over-damped output waveforms of a ramp-step input waveform achieved through wire-sizing. The contribution of this part of the thesis is to provide optimal performance totally through a wire-sizing method.

## 5.2 Performance Criteria

The performance criteria for the optimal design of interconnects are much like criteria of feedback control systems design. The performance of the interconnects and feedback control systems are designed for a faster response time with a minimum amount of steady state Root-Mean-Square error. A significant collection of literature deals with the performance criteria subject in the feedback control systems design [22] [49] [17]. The Integrate of Time and Error (ITAE) criterion has the optimal transient response and no steady state error [49]. It is chosen as the target for the approximated transfer function.

## 5.3 Levenberg-Marquardt Method

Most algorithms for solving the Least-Square Estimation problems involving nonlinear parameters have been centered around either the Taylor series method or the Gradient method [16]. Both of these methods have advantages depending on the applications. In the Taylor series method, the objective function is expanded as a Taylor series and the corrections to the nonlinear parameters are calculated at each iteration based upon the assumption of local linearity. This method converges very fast within the circle of convergence of the Taylor series. The drawback of the algorithms based on the Taylor series method is that they cannot find the global minimum if it is outside of the circle of convergence of the expanded Taylor series, and they suffer from the possible divergence of the successive iterations. The Gradient method, on the other hand, can find the global minimum even it is far away from the initial guess. However, this method suffers from slow

convergence to the optimal solution after the first several rapid advance in the first few iterations.

The Levenberg-Marquardt Method tries to perform an optimum interpolation between the Taylor series method and the Gradient method so that the size and the direction of the steps can be determined simultaneously [36] [16]. The properties proved in Levenberg-Marquardt's paper show that their method combines the best features of the previous two methods and generally avoids their limitations [42]. The Levenberg-Marquardt Method benefits from the maximum neighborhood in which the truncated Taylor series gives an adequate linearized representation of the nonlinear objective function. It can find global minimum outside of this maximum neighborhood. It also converges to the final solution quickly with the help of the Lagrange Multiplier. The Lagrange Multiplier is dynamically adjusted at each iteration. The key feature of Levenberg-Marquardt's method is to search the feasible values of the Lagrange Multiplier such that the objective function is always decreased at each iteration until the convergency to the global minimum is reached. The choice of the Lagrange Multiplier also has some hill-climbing property similar to simulated annealing so as to avoid a local minimum. The Levenberg-Marquardt Method is chosen as the optimization method throughout this thesis.

## 5.4 Previous Work

Papers published by Brews [6] and Frye [25] show design methods for the design of critically damped *point-to-point* interconnections on VLSI and MCM. Another paper published by Cong et al. [13] uses wire-sizing to achieve optimal design for the tree network based upon the distributed-RC model [13]. Wang et al. also published a report which refines Frye's approach to handle the design of critically damped interconnect tree networks on MCM [52].

Later Zhou et al. presented a distributed-RLC model and a second order approximation in their performance-driven-layout paper [58], which extends the wire-sizing algorithm to cover transmission lines on MCM. It is reasonable to assume that the second order

approximation as adopted by Zhou et al. was to keep the optimization manageable. In the same paper, Zhou also reports the use of a second order approximation without the extraction of the time-of-flight to formulate the performance-driven-layout. However, there are three problems arise with the performance-driven-layout as described in Zhou et al. paper.

- The formulation of the optimization problem with critical damping constraints is changed to an optimization without any constraint in implementation.
- The distributed-RLC model is used in the formulation but a distributed-RC model is actually used in the optimization.
- The approach used by Zhou et al. is a two-step approach; the first step is to perform the wire-sizing to achieve shorest delay, and the second step is to critically damp the output signals.

The Zhou’s critical damping design method consists of adding resistors in series at the receiving terminals. The addition of a serial resistor at the receiver is like adding the matching serial termination resistor at the driver, both take extra space to accommodate the resistors.

## 5.5 S-Parameter Macromodel and the Time-of-Flight Extraction

In order to precisely analyze the lossy transmission lines on MCM substrates, the scattering parameter (S-parameter) based macromodel [39] [40] is used to find the approximated transfer function  $H(s)$ . The S-parameter based macromodel simulator can handle both lumped circuit elements and lossy/lossless transmission lines including loops. However, the delay associated with transmission line networks consists of the exponentially charging time and a pure propagation delay representing the finite propagating speed of electromagnetic signals in the dielectric medium. This propagation delay, so called “time-of-flight delay” and denoted by  $\tau$ , is impossible to model perfectly by a finite order of approximation. So, the time-of-flight  $\tau$ , more precisely the factor  $e^{-s\tau}$ , must be extracted from the transfer function of the circuit. The new and important improvement in the S-parameter based

macromodel simulator is the extraction of the exact time-of-flight term in transfer functions [37]. By extracting the time-of-flight of scattering parameters for basic components, an effective network reduction is developed to compute the lower order macromodel of an interconnect system. The output responses, due to the extraction of the time-of-flight, can be greatly improved.

## 5.6 Transfer Function Approximation

The transfer function is defined as:

$$H(s) = \frac{V_{output}(s)}{V_{input}(s)} = e^{-s\tau} \times \hat{H}(s)$$

where  $s$  is the complex variable of the Laplace transformation,  $V_{output}(s)$  and  $V_{input}(s)$  are the output and input waveforms,  $\tau$  is the time-of-flight term, and  $\hat{H}(s)$  is the remaining part of the transfer function after the time-of-flight is extracted. If the  $h(t)$  and  $\hat{h}(t)$  are the inverse transformation of  $H(s)$  and  $\hat{H}(s)$  respectively, then the relationship between  $h(t)$  and  $\hat{h}(t)$  is:

$$h(t) = \hat{h}(t - \tau).$$

The above equation describes a plain shift of the output waveforms  $\hat{h}(t)$  in the time domain for the amount of the time-of-flight  $\tau$ .

The performance criteria of the optimal design must be determined quantitatively. In addition, attention must be paid specifically to some of the qualitative issues such as stability and static accuracy. Significant collection of literature exists that treats the subject of performance criteria in the feedback control systems design [22] [49] [17]. From a thorough investigation, the Integral of Time and Error (ITAE) seems to be the best performance criterion for optimizing the transient response and the steady state Root-Mean-Square error [49]. Transfer functions that meet ITAE criterion have a minimum overshoot and yet faster than critical damping response delay. These are the two key performance issues in designing the optimal self-damped lossy transmission lines on multichip modules. Based on the ITAE criterion, a set of equations for the different orders of the closed-loop transfer functions gives the optimal performance. These transfer functions listed in Equations (5.1) to (5.6) can be





- the undamped case where the roots reside on the imaginary axis.
- the optimal damped case based on ITAE criterion where the roots reside on the circle with 45 degree elevation.
- the critically damped case where both roots overlapped on the point where the circle intersects the negative real axis.
- the overdamped case where both roots reside on the negative real axis but moving in opposite directions.

When a second order approximation is applied to the charging part of the transfer function  $\hat{H}(s)$ , it can be rewritten as:

$$\hat{H}(s) \approx \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \quad (5.7)$$

where  $\omega_n$  is the natural undamped frequency and  $\zeta$  the damping ratio of the transfer function  $\hat{H}(s)$  [33]. Some representations of the pole-pair on the left-half plane are given by [33]:

$$\left\{ \begin{array}{ll} \text{Case 1 : } s_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2} & \text{for } 0 < \zeta < 1; \\ \text{Case 2 : } s_{1,2} = -\omega_n & \text{for } \zeta = 1; \\ \text{Case 3 : } s_{1,2} = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2-1} & \text{for } 1 < \zeta; \end{array} \right.$$

The damping ratio  $\zeta$  and the natural undamped frequency  $\omega_n$  can be expressed in terms of two poles  $s_1$  and  $s_2$ :

$$\left\{ \begin{array}{l} \omega_n = \sqrt{s_1 \cdot s_2} \\ \zeta = \frac{s_1 + s_2}{-2\sqrt{s_1 \cdot s_2}}. \end{array} \right.$$

For a given input waveform in the time domain, the three quantities of the second order approximated transfer function: the time-of-flight  $\tau$ , the natural undamped frequency  $\omega_n$  and the damping ratio  $\zeta$  *uniquely* determine the time domain response waveform at the receiver. Similarly, the propagation delay for a signal to reach 50% of its final value at the receiver is also *uniquely* determined by these three quantities:  $\tau$ ,  $\omega_n$ , and  $\zeta$ . The damping ratio  $\zeta$  and the propagation delay will be used together in the optimal wire sizing process. According to the ITAE criterion, the best transient performance happens when  $\zeta = \frac{1}{\sqrt{2}}$  for

a second order system [49] [22] [33] [6]. This value will be used as the target of optimization method in all the experiments.

For higher order systems, the coefficients of the approximated transfer functions are used. For a given input waveform in the time domain, these coefficients *uniquely* determine the time domain response waveform at the receiver. So instead of using the damping ratio  $\zeta$ , the coefficients of the approximated transfer functions and the propagation delay will be used together in the optimal wire sizing process. To simplify the presentation in this thesis, the second order approximation of the transfer function is used throughout the subsequent sections.

### 5.7 Formulation of the Optimal Self-Damped Design

The optimal design problem can be formulated as follows: Given a network ( $N$ ) consists of a set of edge  $E$  with fixed lengths, the width of each edge belongs to a set of feasible widths  $C = \{c_1, c_2, \dots, c_{max}\}$  where  $(c_i < c_{i+1}, 1 \leq i \leq (max - 1))$ . The optimal design is to find an optimal width assignment:  $W^* \subset C$  and a mapping  $M : B \rightarrow W^*$  such that every path from root to leaves has the shortest delay and the same damping criteria if possible. The objective is to optimize the performance which includes minimizing the signal propagation delay for each path from the source to a receiver with either no overshoot or a controlled amount of overshoot. The width of edge  $i$ ,  $w_i$ , is bounded by the minimum and maximum feasible width,  $min(w_i) \leq w_i \leq max(w_i)$ , where  $min(w_i), max(w_i) \in C$ . The incremental difference,  $\Delta c$ , where  $\Delta c = (c_{i+1} - c_i), 1 \leq i \leq (max - 1)$ , between feasible widths is depicted by the process technology.

The following notations are defined for the formulation of the optimal self-damped design.

$n$  : the number of the edges in the network.

$m$  : the number of the receivers (the output nodes).

$w_i$  : the width of edge  $i$ .

$l_i$  : the given length of edge  $i$ .

$\tau_j$  : the propagation delay for a signal to reach 50% of its final output value transmitted along path from the source to a receiver  $j$ .

$\zeta_j$  : the damping ratio of the lowest conditional frequency of the transfer function from the source to a receiver  $j$ .

$\zeta_{target}$  : the electric damping criteria set forth by the user.

The optimal self-damping design for a general network can be formulated as a general nonlinear programming problem as follows, Let

$$F(w_1, w_2, \dots, w_n) = \text{Maximum}(\tau_j), 1 \leq j \leq m. \quad (5.8)$$

$$G(w_1, w_2, \dots, w_n) = \text{Maximum}(|\zeta_j - \zeta_{target}|), 1 \leq j \leq m. \quad (5.9)$$

### Objective

$$\text{Minimize } F(w_1, w_2, \dots, w_n) \quad (5.10)$$

### Constraints

$$G(w_1, w_2, \dots, w_n) \leq \epsilon, w_i \in C, 1 \leq i \leq n, 1 \leq j \leq m, \text{ and } \epsilon \text{ is a preset constant.} \quad (5.11)$$

$$\zeta_j(w_1, w_2, \dots, w_n) \geq \frac{1}{\sqrt{2}}, w_i \in C, 1 \leq i \leq n, 1 \leq j \leq m. \quad (5.12)$$

$$g_i(w_1, w_2, \dots, w_n) = \{w_i - \min(w_i)\} \geq 0, w_i \in C, 1 \leq i \leq n. \quad (5.13)$$

$$h_i(w_1, w_2, \dots, w_n) = \{\max(w_i) - w_i\} \geq 0, w_i \in C, 1 \leq i \leq n. \quad (5.14)$$

The objective of an optimal self-damped design is to size the wire width of each edge in order to minimize both the maximum delay stated in Equation (5.8) and the maximum damping ratio error stated in Equation (5.9), while maintaining that all output damping ratios satisfy the constraint in Equation (5.11) and (5.12), and all edge widths satisfy the constraints in Equation (5.13) and (5.14).

The way to find the optimal answer is to use the *perturbation method* [43], which works by perturbing each of the design parameters  $\lambda_i$  a small amount in each direction, and finds the right direction and distance to change  $\lambda_i$ . In this thesis, the width of each edge  $w_i$  is the chosen design parameter to be changed and a simulation is run at each perturbation to find

the propagation delay and the damping ratio. The perturbation results of the propagation delays and the damping ratios are then used to compute the gradient matrix. Details of the optimization method is described in Section 5.9.

## 5.8 Proof of the Correctness of the Least-Squares Estimation Transformation

The optimal wire-sizing problem formulated in Equation (5.8) is a general nonlinear programming problem. Both the damping ratio error minimization and delay minimization problem can be transformed into one of a Least-Square Estimation. There are two parts in the problem of the original optimization. The first part is the minimization of the maximum delay and the second part is the minimization of the maximum damping ratio error.

The following is the proof of the correctness of the transformation for the first part. Let  $\tau_j$  be the 50% signal propagation delay from the source to the receiver  $j$  ( $1 \leq j \leq m$ ), where  $m$  is the number of receivers. Let column vector  $\Theta = \{\tau_1, \tau_2, \dots, \tau_m\}^T$  represents the delay vector, where  $T$  denotes matrix transposition. The summation of all squares of delay errors is:

$$\Phi(w_1, w_2, \dots, w_n) = \Theta^T \Theta = \sum_{j=1}^m \tau_j^2 \quad (5.15)$$

Define the root-mean-square (rms) error of the delay as:

$$\varphi = \sqrt{\frac{\Phi}{m}} = \sqrt{\frac{\sum_{j=1}^m \tau_j^2}{m}} \quad (5.16)$$

From Equation (5.8), the maximum delay is  $F(w_1, w_2, \dots, w_n) = \text{Maximum}(\tau_j)$ .

**Theorem 1:** *Given a single-source multi-receiver network, the root-mean-square error defined in Equation (5.16) and maximum delay defined in Equation (5.8) are linearly bound to each other.*

**Proof:** Assume the largest delay is  $\tau_{max}$ . For all  $\tau_j$ ,  $\tau_j \leq \tau_{max}$ , ( $1 \leq j \leq m$ ), where  $m$  is the number of receivers and it is a constant for a given network. From Equation

(5.16), we have  $\varphi = \sqrt{\sum_{j=1}^m \tau_j^2 / m} \leq \sqrt{\sum_{j=1}^m \tau_{max}^2 / m} = \tau_{max} = F(w_1, w_2, \dots, w_n)$ . On the other hand, we have  $F(w_1, w_2, \dots, w_n) = \text{Maximum}(\tau_j) = \tau_{max} = \sqrt{\tau_{max}^2} \leq \sqrt{\sum_{j=1}^m \tau_j^2} = \sqrt{m \cdot \sum_{j=1}^m \tau_j^2 / m} = \sqrt{m} \cdot \varphi$ . So  $\varphi$  and  $F(w_1, w_2, \dots, w_n)$  are linearly bound to each other.

The following is the proof of the correctness of the transformation for the second part. Let column vector  $\Omega = \{|\zeta_1 - \zeta_{target}|, |\zeta_2 - \zeta_{target}|, \dots, |\zeta_m - \zeta_{target}|\}^T$  represents the damping ratio error vector. The summation of all squares of damping ratio errors is:

$$\Psi(w_1, w_2, \dots, w_n) = \Omega^T \Omega = \sum_{j=1}^m (\zeta_j - \zeta_{target})^2 \quad (5.17)$$

Define the root-mean-square (rms) error of the damping ratio as:

$$\psi = \sqrt{\frac{\Psi}{m}} = \sqrt{\sum_{j=1}^m \frac{(\zeta_j - \zeta_{target})^2}{m}} \quad (5.18)$$

From Equation (5.9), the maximum damping ratio error is:

$$G(w_1, w_2, \dots, w_n) = \text{Maximum}(|\zeta_j - \zeta_{target}|).$$

**Theorem 2:** *Given a single-source multi-receiver network, the root-mean-square damping ratio error defined in Equation (5.18) and maximum damping ratio error defined in Equation (5.9) are linearly bound to each other.*

**Proof:** Assume the largest damping ratio is  $\zeta_{max}$  and it deviates from  $\zeta_{target}$  the most, the maximum damping ratio error is  $|\zeta_{max} - \zeta_{target}|$ . (The proof can be applied to the case where the smallest damping ratio  $\zeta_{min}$  deviates from  $\zeta_{target}$  the most. We only have to replace all  $\zeta_{max}$  with  $\zeta_{min}$  in the proof.) For all  $\zeta_j$ ,  $|\zeta_j - \zeta_{target}| \leq |\zeta_{max} - \zeta_{target}|$ , ( $1 \leq j \leq m$ ), where  $m$  is the number of receivers and it is a constant for a given network. From Equation (5.18), we have  $\psi = \sqrt{\sum_{j=1}^m (\zeta_j - \zeta_{target})^2 / m} \leq \sqrt{\sum_{j=1}^m (\zeta_{max} - \zeta_{target})^2 / m} = |\zeta_{max} - \zeta_{target}| = \text{Maximum}(|\zeta_j - \zeta_{target}|) = G(w_1, w_2, \dots, w_n)$ . On the other hand, we have  $G(w_1, w_2, \dots, w_n) = \text{Maximum}(|\zeta_j - \zeta_{target}|) = |\zeta_{max} - \zeta_{target}| = \sqrt{(\zeta_{max} - \zeta_{target})^2} \leq \sqrt{\sum_{j=1}^m (\zeta_j - \zeta_{target})^2} = \sqrt{m \cdot \sum_{j=1}^m (\zeta_j - \zeta_{target})^2 / m} = \sqrt{m} \cdot \psi$ . So  $\psi$  and  $G(w_1, w_2, \dots, w_n)$  are linearly bound to each other.

From the above two theorems, minimizing the two Least-Square Estimation problem in Equations 5.16 and 5.18 bounds the maximum path delay  $F(w_1, w_2, \dots, w_n)$  defined in Equation (5.8) and maximum damping ratio error  $G(w_1, w_2, \dots, w_n)$  defined in Equation (5.9) of the original nonlinear programming problem.

## 5.9 Implementation of the Least-Square Estimation Optimization

The Levenberg-Marquardt method is used to solve the least-square estimation problem [42] [16]. Theorems 1 and 2 show the consistency between the minimizing of the original problem and the minimizing of the transformed root-mean-square estimation problem. Starting with an arbitrary initial solution of width assignment  $W^{(0)} = \{w_1^{(0)}, w_2^{(0)}, \dots, w_n^{(0)}\}^T$ , according to Levenberg-Marquardt, the width assignment  $W$  is optimized according to the following formula:

$$W^{(k+1)} = W^{(k)} - (J^T J + \lambda \Lambda)^{-1} J^T \begin{bmatrix} \Theta|_{W^{(k)}} \\ \Omega|_{W^{(k)}} \end{bmatrix} \quad (5.19)$$

where  $k$  is the number of iteration,  $\Theta|_{W^{(k)}}$  the column vector of delays from the source to all the receivers at the  $k$ -th iteration, and  $\Omega|_{W^{(k)}}$  the column vector of damping ratio errors from the source to all the receivers at the  $k$ -th iteration.  $J$  is the  $2m \times n$  sensitivity matrix,  $J^T$  is the transposition matrix of  $J$  where the  $(i, j)$ th element  $J^T(i, j) = J(j, i)$ ,  $\Lambda$  is a diagonal matrix in which the values of its diagonal elements are the same as the diagonal elements of  $J^T J$ , and  $\lambda$  is the *Lagrange Multiplier* properly selected to speed up the convergence of the optimization process [42]. Round-off occurs when computing  $w_i^{(k)}$  so that Equation (5.13) and (5.14) are always satisfied.  $J^T \begin{bmatrix} \Theta|_{W^{(k)}} \\ \Omega|_{W^{(k)}} \end{bmatrix}$  represents the gradient around the current width assignment  $W^{(k)}$ . To obtain the sensitivity matrix  $J$ , the  $(i, j)$ th element is defined as:

$$J(i, j) = \begin{cases} \frac{\partial \Theta[i]}{\partial w_j}, & \text{if } 1 \leq i \leq m \\ \frac{\partial \Omega[i-m]}{\partial w_j}, & \text{if } m+1 \leq i \leq 2m \end{cases}. \quad (5.20)$$

The partial derivatives are computed using a central difference method. The optimization continues until the maximum damping ratio error is less than a prescribed value, the maximum damping ratio error cannot be further improved, the maximum delay cannot be further improved, or the iteration number exceeds a preset limit. The convergence to the optimal values of Levenberg-Marquardt method is proved in [42].

## 5.10 Experimental Results for the Optimal Self-Damped Design

The examples tested are constructed with High Performance MCM process technologies published by Frye [24]. The important parameters of the MCM process are listed in Table 5.1. In the case of the uniform width, all of the widths are equal to  $25\mu m$  for all of the examples tested. All of the drivers are modeled with a step input voltage source in series with the parallel combination of a  $12\Omega$  resistor and a  $4.3pF$  capacitor. All of the receivers are modeled using a  $2.5pF$  capacitor. The damping ratio target is chosen to be  $\frac{1}{\sqrt{2}}$  for shorter propagation delay with a controlled amount of overshoot [6] [33] [22] [49] [17]. For all of the simulation results of the S-parameter macromodels, all of the far-end waveforms are simulated with the time-of-flight captured explicitly with the method stated in Section 5.5.

### 5.10.1 Example 1 A Tree Network

Example 5.10.1 is the tree network shown in Zhou's MCMC paper [58]. Figure 5.3 (a) shows the topology, lengths, and widths of the optimal design for all of the edges. The optimal design performance is compared with the uniform width case. The per-unit-length  $R$ ,  $C$ , and  $L$  of the uniform width design are listed in Table 5.1. The simulation waveforms of the optimal design and uniform width design are shown in Figure 5.3 (b). The maximum path delays, and the percentage of improvement are listed in Table 5.2.



HIGH PERFORMANCE MCM-D	
Thickness of Dielectric ( $\mu m$ )	5
$\epsilon_{rel}$	3.2
Thickness of Metal ( $\mu m$ )	2.5
R ( $\Omega/\mu m$ ) for typical edge width	2.4
L ( $nH/\mu m$ ) for typical edge width	2.9
C ( $pF/\mu m$ ) for typical edge width	1.39
lower bound metal line width ( $\mu m$ )	10
typical metal line width ( $\mu m$ )	25
upper bound metal line width ( $\mu m$ )	50

Table 5.1: **The High Performance MCM Technologies Process Parameters**

	Maximum Path Delay (Uniform) (nS)	Maximum Path Delay (Optimal) (nS)	Percent Improvement (%)
Example 5.10.1	0.9874	0.7182	27.26
Example 5.10.2	0.9354	0.8245	11.85
Example 5.10.3	1.2308	0.8533	30.67

Table 5.2: **Comparison between the Uniform Width Design and the Optimal Design**

### 5.10.2 Example 2 A Network with a Loop

Example 5.10.2 is a network with a loop. Figure 5.4 shows the topology, lengths, and widths of the optimal design for all of the edges. The simulation waveforms of the receiver 4 in the loop network for the optimal design and uniform width design are shown in Figure 5.4 (b). The maximum path delays and the percentage of improvement are listed in Table 5.2.



### 5.10.3 Example 3 A Clock Tree Network

Example 5.10.3 is a clock tree network shown in Zhu's ICCAD paper [59]. Figure 5.5 (a) shows the topology, lengths, and widths of the optimal design for all of the edges. The simulation waveforms of the optimal design and uniform width design are shown in Figure 5.5 (b). The maximum path delays and the percentage of improvement are listed in Table 5.2.

The experimental results show that this optimization method reduces the maximum of all of the signal propagation delays considerably when compared to the uniform width designs. This proves that the generally adopted practice of the uniform width design method is overly constrained. The optimal width assignment produces a fast and stable signal propagation. There is as much as 30% improvement that has been found by this optimal design method.



## CHAPTER 6. Transformation of Min-Max Optimization to Least-Square Estimation and Its Application to Interconnect Design Optimization

For most of the Computer-Aided Design software, optimization is required as part of the heuristic algorithm. The most common optimization problem encountered is the minimization of the maximum of all the observable outputs, which is often referred to as the Min-Max optimization problem. Because the Min-Max optimization is generally a nonlinear programming problem, it is not only hard to solve but also takes a long computation time. An alternative method to solve the Min-Max optimization problem and to obtain a good solution in a relative short period of time is to transform the original Min-Max optimization problem into one of a Least-Square Estimation (LSE), which has a number of well defined methods for solving it. It can also be shown that the target function of the original Min-Max optimization problem and that of the transformed LSE problem are linearly bound to each other. Although these two problems have different objective functions, optimizing the transformed problem can produce a solution to the original Min-Max problem. To optimize the solution to the LSE problem, one tries to minimize a *target function* which is the Root-Mean-Square (RMS) value of a given function. In this method, one can obtain a *solution function* which is a set of values of a given function. By the definition of the LSE problem, when evaluating the Root-Mean-Square value of a solution function, this RMS value should be at its minimum. The solution to the original Min-Max problem is

found from the maximum value of the solution function of the LSE problem. However, this does not mean that solving the LSE problem can lead to the exact solution to the original Min-Max problem, it merely states that solving the LSE problem gives one possible solution to the original Min-Max problem. The term: “linearly bound each other” means that obtaining the solution function of the LSE problem, one can translate the maximum value of the solution function at hand to a solution of the original Min-Max problem; it does not say anything about the quality of this solution compared with those obtained from other methods which are used to solve the same original Min-Max problem.

As illustrated in Section 6.1, for any single observable output Min-Max problem, the maximum value of the solution function of the LSE problem is the same as the solution of the original Min-Max problem. However, for multiple observable outputs Min-Max problem, the Root-Mean-Square value of the solution function of the LSE problem linearly bounds the solution of the original Min-Max problem with a wide range. Although the solution to the original Min-Max problem can still be found, however, one cannot be certain about the quality of the solution of the original Min-Max problem because of this wide range. Worst of all, for Min-Max problems with large number of observable outputs, the solution of the original Min-Max problem is bounded by a huge range, between one and several tens multiplying the solution of the LSE problem. For example, if a Min-Max problem has 1000 observable output variables, then the solution of the Min-Max problem can lie between one and 31.62 multiples of the Root-Mean-Square value of the transformed LSE problem, since the square root of 1000 is 31.62. The drawback of this huge range is that it gives rise to such a huge uncertainty in transferring the maximum value of the solution function of the LSE problem back to the Min-Max solution. This renders the idea of the transformation and solving it through LSE optimization much less useful.

Previous research that utilizes the transformation include those of Zhu et al. [59] and Wang et al. [54]. Both of these results lack the ability to precisely translate the maximum value of the solution function of the LSE problem back to the original Min-Max problem.

In this part of thesis, a novel transformation is formulated, followed immediately by a

proof that this transformation indeed gives the tightest bound for the ideal case. Because of this tightest bound, the maximum value of the solution function of the LSE problem can be transformed back as the solution of the original Min-Max problem even for the case of the multiple observable output variables.

As demonstrated in Example 6.4.3 in Section 6.4, the best case happens when the difference between all the values of a given function can be minimized. It makes this novel approach especially effective in solving the problem of minimizing the delay of the equal-path-length clock tree.

Section 6.1 presents the original Min-Max optimization problem, the transformation into one LSE problem, and the proof of their target functions linearly bound each other. The drawback of this type of transformation for the case of the multiple observable output variables is also discussed in detail. Section 6.2 presents the novel transformation into two LSE problems and the proof of the tighter linear bound between their target functions. The advantage of this novel transformation and an ideal case of the solution translation is also discussed in detail in Section 6.2. The details about the implementation of the optimization of the LSE problems is shown in Section 6.3, whereas Section 6.4 demonstrates the usefulness of this novel transformation with three examples.

## **6.1 Transform the Min-Max Problem to One Least-Square Estimation Problem**

Computer-Aided Design software usually requires the optimization to be a part of the heuristic algorithm. One of the optimization calls for the minimization of the maximum of all the observable outputs. The transformation defined in this section closely followed that of Wang et al. [54]. However, the notation used here is much more general, which can be apply to the Min-Max optimization problem of any positive function.

### **6.1.1 Physical Example**





If one defines the target function of the Min-Max problem to be  $F(W, J)$ , one can write:

$$\min_W [F(W, J)] \triangleq \min_W \left[ \max_J [G(W, J)] \right]. \quad (6.1)$$

From a general optimization textbook [36], an easier alternative of solving this Min-Max problem is to transform it into a Least-Square Estimation (LSE) problem. In order to solve this LSE problem, one has to minimize the RMS of the given function  $G(W, J)$ . Solving the LSE problem gives a solution function whose maximum can be translated into a solution of the original Min-Max problem.

### 6.1.3 Transformation Formulation

The detailed formulation of the transformation is as follows: let  $G(W, j)$  be the observed  $j$ -th output, and let the column vector

$$\Theta(W, J) \triangleq \{G(W, 1), G(W, 2), \dots, G(W, j), \dots, G(W, m)\}^T, \quad (6.2)$$

represents the estimation vector, where  $T$  denotes the transposition operation,  $j \in J$  and  $1 \leq j \leq m$ , and  $m$  is the number of the outputs. The summation of all squares of the estimations  $\Phi(W, J)$  is:

$$\Phi(w_1, w_2, \dots, w_n, 1, \dots, m) \triangleq \Theta(W, J)^T \Theta(W, J) = \sum_{j=1}^m [G(W, j)]^2 \quad (6.3)$$

If one defines the Root-Mean-Square (RMS) of the estimation as:

$$\varphi(w_1, w_2, \dots, w_n, 1, \dots, m) \triangleq \sqrt{\frac{\Phi(W, J)}{m}} = \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}}, \quad (6.4)$$

then new optimization problem becomes the minimization of the RMS of the given function  $G(W, J)$ . One can write the new target function of the optimization as:

$$\min_W [\varphi(W, J)] \triangleq \min_W \left[ \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} \right]. \quad (6.5)$$

The new target function (Equation (6.5)) does not correspond to the original target function (Equation (6.1)). However, it can be shown that the minimization result of the

new target function linearly bounds the original optimization solution. As can be seen later, if the linear bound is tight, the maximum value of the optimization result of the new solution function can be translated to be the solution of the original Min-Max problem.

The following is the proof of the target functions of the transformation linearly bound each other.

**Theorem 3:** *Given a function  $G(W, J)$ , the minimum of the Root-Mean-Square (RMS) as defined in Equation (6.5) and the minimum of the maximum as defined in Equation (6.1) linearly bound each other.*

**Proof:**

Given that the largest of all the  $G(W, J)$  is equal to  $F(W, J)$ ,  $F(W, J) \triangleq \max_J G(W, J)$ . For all  $G(W, j)$ ,  $G(W, j) \leq F(W, J)$ , ( $1 \leq j \leq m$ ), one can obtain Equation (6.6 from Equation (6.5)

$$\varphi(W, J) \triangleq \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} \leq \sqrt{\sum_{j=1}^m \frac{[\max_J G(W, J)]^2}{m}} = \max_J G(W, J) = F(W, J). \quad (6.6)$$

On the other hand, we have

$$\begin{aligned} F(W, J) \triangleq \max_J [G(W, J)] &= \sqrt{\max_J [G(W, J)]^2} \leq \sqrt{\sum_{j=1}^m [G(W, j)]^2} \\ &= \sqrt{m \cdot \sum_{j=1}^m \frac{[G(W, j)]^2}{m}} = \sqrt{m} \cdot \varphi(W, J). \end{aligned} \quad (6.7)$$

From Equation (6.6) and (6.7), one has

$$\varphi(W, J) \leq F(W, J) \leq \sqrt{m} \cdot \varphi(W, J). \quad (6.8)$$

This concludes the proof that  $\varphi(W, J)$  and  $F(W, j)$  linearly bound each other.

#### 6.1.4 Physical Meaning

Given a *positive* function  $G(W, J)$  which one wants to optimize, where  $W = \{w_i | i = 1 \dots n\}$  and  $J = \{j | j = 1 \dots m\}$ , one can define a mapping  $M$  from  $J = \{j | j = 1 \dots m\}$  to  $J' = \{j' | j' = \frac{1}{m} \dots 1\}$ , and plot the function  $G(W, J)$  as in Figure 6.2, where each  $G(W, j)$



the four values w.r.t. the function  $G(W, J)$  in Figure 6.3, one can see the relation between these four values. The areas under the rectangles of these four values are:

$$Area_{max} = \int_0^1 \left[ \max_J [G(W, J)] \right] dj = \max_J [G(W, J)], \quad (6.10)$$

$$Area_{min} = \int_0^1 \left[ \min_J [G(W, J)] \right] dj = \min_J [G(W, J)], \quad (6.11)$$

$$Area_{avg} = \int_0^1 [avg_J [G(W, J)]] dj = avg_J [G(W, J)], \quad (6.12)$$

$$Area_{RMS} = \int_0^1 \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} dj = RMS_J [G(W, J)]. \quad (6.13)$$

From Figure 6.3, one can tell that  $Area_{RMS} \leq Area_{max}$ . From the linear bound, one has  $Area_{max} \leq \sqrt{m} \cdot Area_{RMS}$  which is represented by the biggest rectangle in Figure 6.3. It is important to note that Figure 6.3 is not drawn to the scale, so the actual size of the area which equals to  $\sqrt{m} \cdot Area_{RMS}$  could be much larger. The term “linearly bound each other” means that the  $Area_{max} = F$  lies in between the  $Area_{RMS} = \varphi$  and the  $\sqrt{m} \cdot Area_{RMS} = \sqrt{m} \cdot \varphi$ . By minimizing the area of  $Area_{RMS}$ , one also brings down the area of  $\sqrt{m} \cdot Area_{RMS}$ , since the area of  $Area_{max}$  lies in between them, so it is also minimized. However, for the case of multiple observable output variable case as shown in the same figure, the range between the  $Area_{RMS}$  and  $\sqrt{m} \cdot Area_{RMS}$  can be quite large. This gives rise to the uncertainty about the minimization of the original target function when one attempts to optimize the LSE problem.

### 6.1.5 Practical Considerations

When applied to a case of the single observable output,  $m = 1$ , Equation (6.8) becomes  $\varphi \leq F \leq \sqrt{1}\varphi$ , which means that the two target functions are equal, that is, minimizing one is to minimizing the other. Thus, the solution to the LSE problem,  $\varphi$ , can be taken as the solution of the Min-Max Problem. However, for a case of the multiple observable output variable, Equation (6.8) is  $\varphi \leq F \leq \sqrt{m}\varphi$ , which means that the *target function* of the LSE problem,  $\varphi$ , linearly bounds the solution of the original Min-Max problem,  $F$ , with a range. This means that the two target functions are not the same, so minimizing one

does not guarantee the minimization of the other. This range is defined as from  $\varphi$  to  $\sqrt{m}\varphi$ . For example, if a Min-Max problem has 1000 observable outputs which is common for a global clock distribution net, then the solution of the original Min-Max problem is bounded between one and  $\sqrt{1000} = 31.62$  multiplies the *target function* of the LSE problem. This huge range give rise to one's hesitation about the quality of the solution to the Min-Max problem . This is the drawback when one tries to use the transformation to solve the Min-Max problem in the case of the multiple observable output. The huge linear bound range and the uncertainty in the solution translation renders the idea of the transformation and solving through LSE optimization much less useful.

In Section 6.2, a novel transformation approach is presented. This novel transformation can be shown to have a tighter linear bound than that of Equation (6.8) and, in the best case, can make the solution translation of the case of the multiple observable output identical to that of the single observable output.

## 6.2 Novel Transformation of the Min-Max Problem into Two Least-Square Estimation Problems

The transformation presented in Section 6.1 changes the target function of the Min-Max problem to the RMS of a different LSE problem. This transformation works only because minimization of the RMS of the different LSE problem is identical to the minimization of the original problem for a single observable output. In the case of the multiple observable output variable, the minimization of the LSE problem cannot be proved to be identical to the optimization of the Min-Max problem. This prevents the use of the transformation as a method of solving the original Min-Max problem.

In this Section, a novel transformation which gives even a tighter linear bound is presented. The reason for the development of this tighter bound is that it will work even in the case of the multiple observable output. This novel transformation transforms the original Min-Max problem into two LSE problem; one of them is the minimization of the RMS of a given function; and the other is the minimization of the RMS of the difference

between an arbitrary constant and a given function  $G(W, J)$ . The first LSE problem is the same as the one presented to in Section 6.1.3. Solving these two LSE problems together gives two *target function values* and one *solution function*. The maximum value of the solution function of the two LSE problem can be translated into a solution of the original Min-Max problem. The solution to the second LSE problem only serves as an assurance that when optimizing the first LSE problem, one is indeed minimizing the original Min-Max problem.

### 6.2.1 Problem Definition

The problem definition is identical to the one that presented in Section 6.1.2. Given a *positive* function  $G(W, J)$  where  $W = \{w_i | i = 1 \dots n\}$  and  $J = \{j | j = 1 \dots m\}$ , one can define the target function of the Min-Max problem to be  $F(W, J)$ , and write:

$$\min_W [F(W, J)] \triangleq \min_W \left[ \max_J [G(W, J)] \right]. \quad (6.14)$$

### 6.2.2 Transformation Formulation

The transformed optimization problem consists of two LSE problems. The first one is the minimization of the RMS of the given function  $G(W, J)$ . The second LSE problem is the minimization of the RMS of the difference between an arbitrary constant  $c$  and the given function  $G(W, J)$ . For the second LSE problem, one has to create a new Min-Max problem. Define a new function  $\Delta(W, J) \triangleq [\max_J [c - G(W, J)]]$ , where  $c$  is an arbitrary constant which remains the same during the entire LSE optimization process. The choice of  $c$  affects the outcome of the the solution translation a great deal and will be discussed in detail in Section 6.2.4. Define the target function of the new Min-Max problem to be  $\Delta(W, J)$ , one can write:

$$\min_W [\Delta(W, J)] \triangleq \min_W \left[ \max_J [c - G(W, J)] \right]. \quad (6.15)$$

Denote  $G(W, j)$  to be the discrete value of  $G(W, J)$  at  $j$ . Define the estimation vector for the first LSE problem to be  $\Theta(W, J) \triangleq \{G(W, 1), G(W, 2), \dots, G(W, j), \dots, G(W, m)\}^T$ ,

where  $T$  denotes the transposition operation,  $j \in J$  and  $1 \leq j \leq m$ , and  $m$  is the number of the observable outputs. Define the summation of all squares of the first estimations,  $\Phi(W, J)$ , to be:

$$\Phi(w_1, w_2, \dots, w_n, 1, \dots, m) \triangleq \Theta(W, J)^T \Theta(W, J) = \sum_{j=1}^m [G(W, j)]^2 \quad (6.16)$$

Define the Root-Mean-Square (RMS) of the first estimation as:

$$\varphi(w_1, w_2, \dots, w_n, 1, \dots, m) \triangleq \sqrt{\frac{\Phi(W, J)}{m}} = \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} \quad (6.17)$$

Similarly, denote  $\Delta(W, j)$  to be the discrete value of  $\Delta(W, J)$  at  $j$ . Define the estimation vector for the second LSE problem to be:

$$\Omega(W, J) \triangleq \{c - G(W, 1), c - G(W, 2), \dots, c - G(W, j), \dots, c - G(W, m)\}^T, \quad (6.18)$$

where  $T$  denotes the transposition operation. Define the summation of all squares of the second estimations,  $\Psi(W, J, c)$ , to be:

$$\Psi(w_1, w_2, \dots, w_n, 1, \dots, m, c) \triangleq \Omega(W, J)^T \Omega(W, J) = \sum_{j=1}^m [c - G(W, j)]^2 \quad (6.19)$$

Define the Root-Mean-Square (RMS) of the second estimation as:

$$\psi(w_1, w_2, \dots, w_n, 1, \dots, m, c) \triangleq \sqrt{\frac{\Psi(W, J, c)}{m}} = \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \quad (6.20)$$

The transformed optimization problem consists of two LSE problems. The first one is the minimization of the RMS of the given function  $G(W, J)$ . The second LSE problem is the minimization of the RMS of the difference between an arbitrary constant  $c$  and the given function  $G(W, J)$ . One can write the new target function of the first optimization as:

$$\min_W [\varphi(W, J)] \triangleq \min_W \left[ \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} \right]. \quad (6.21)$$

Likewise, one can write the new target function of the second optimization as:

$$\min_W [\psi(W, J, c)] \triangleq \min_W \left[ \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \right]. \quad (6.22)$$

The two new target functions (Equation (6.21)) and (Equation (6.22)) do not correspond to the original target function (Equation (6.14)). However, it can be shown that the linear combination of the minimization results of the two new target functions linearly bound the original optimization solution. As can be seen in Section 6.2.4, if the linear bound is the tightest for the idea case, the optimization result of the maximum value of the *solution function* of the two LSE optimization can be translated to be the solution of the original Min-Max problem.

The following is the proof of the target functions of the transformation linearly bound each other. From Theorem 3, one has:

$$\varphi(W, J) \leq F(W, J) \leq \sqrt{m} \cdot \varphi(W, J). \quad (6.23)$$

Similarly, for the new Min-Max problem and the new transformed LSE problem, one can prove their target functions linearly bound each other.

**Theorem 4:** *Given a function  $\Delta(W, J)$ , the minimum of the Root-Mean-Square (RMS) defined in Equation (6.20) and the minimum of the maximum defined in Equation (6.15) linearly bound each other.*

**Proof:**

Given that the largest of all the  $\Delta(W, J)$  is  $H(W, J)$ .  $H(W, J) \triangleq \max_J \Delta(W, J)$ . For all  $\Delta(W, j)$ ,  $\Delta(W, j) \leq H(W, J)$ , ( $1 \leq j \leq m$ ). From Equation (6.20), one has:

$$\begin{aligned} \psi(W, J, c) &\triangleq \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \leq \sqrt{\sum_{j=1}^m \frac{\max_J [c - G(W, J)]^2}{m}} \\ &= \max_J \Delta(W, J) = H(W, J). \end{aligned} \quad (6.24)$$

On the other hand,  $H(W, J)$  is defined as

$$\begin{aligned} H(W, J) &\triangleq \max_j [c - G(W, j)] = \sqrt{\max_j [c - G(W, j)]^2} \leq \sqrt{\sum_{j=1}^m [c - G(W, j)]^2} \\ &= \sqrt{m \cdot \sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} = \sqrt{m} \cdot \psi(W, J, c). \end{aligned} \quad (6.25)$$



From Equation (6.24) and (6.25), one has

$$\psi(W, J, c) \leq H(W, J) \leq \sqrt{m} \cdot \psi(W, J, c). \quad (6.26)$$

This concludes the proof that  $\psi(W, J, c)$  and  $H(W, J)$  linearly bound each other.

From the solution of the two LSE problem, one can transform them back to a solution of the original Min-Max problem.

**Theorem 5:** *Given a function  $G(W, J)$ , the minimum of the Root-Mean-Square (RMS) as defined in Equation (6.17), and the sum of the minimum of the RMS and  $\sqrt{m}$  times the minimum of the RMS as defined in Equation (6.20) linearly bound the target function of the Min-Max problem defined in Equation (6.14). i. e.*

$$\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c). \quad (6.27)$$

**Proof:**

From the Theorem 3, one has:

$$\varphi(W, J) \leq F(W, J). \quad (6.28)$$

It is shown in Appendix A that:

$$F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c). \quad (6.29)$$

From Equations (6.28) and (6.29), one has:

$$\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c).$$

### 6.2.3 Physical Meaning

Given the same *positive* function  $G(W, J)$  where  $W = \{w_i | i = 1 \dots n\}$  and  $J = \{j | j = 1 \dots m\}$ , the objective of the Min-Max problem is to minimize the maximum of all its values. The physical meaning of the novel transformation is not only to minimize the RMS of the given function  $G(W, J)$  but also to minimize the difference between the given





the solution value of the original Min-Max problem, then after the LSE optimization is done, compute a new constant  $c = \text{avg}_j [G(W, J)]$  and repeat the LSE optimization again. Several iterations of the LSE optimization process are required until the value of the function  $\psi(W, J, c)$  is sufficiently small and then the maximum value of the *solution function* of the two LSE optimization is taken as the solution to the Min-Max problem.

Generally speaking, the function  $\psi(W, J, c)$  is two to three orders of magnitudes smaller than  $\varphi(W, J)$ . This makes the linear bound  $\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c)$  much tighter than the linear bound  $\varphi(W, J) \leq F(W, J) \leq \sqrt{m} \cdot \varphi(W, J)$  as presented in Section 6.1. This tighter bound assures the equivalence in minimization of the two different kinds of optimization in the case of the multiple observable output variable. For the best case, when the constant  $c$  chosen is  $\text{avg}_j [G(W, J)]$ , the function  $\psi(W, J, c) = 0$ . Then  $\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c) = \varphi(W, J) + \sqrt{m} \cdot 0 = \varphi(W, J)$ , which means  $\varphi(W, J)$  is equal to  $F(W, J)$ . Because this ideal case makes the two target functions equal to one another, minimization of the LSE problem is indeed identical to the minimization of the original Min-Max problem. This is the tightest linear bound possible which makes the translation of the solution in the case of the multiple observable output variable exactly the same as that of the single observable output variable.

### 6.3 Implementation of the Least-Square Estimation Optimization

The Levenberg-Marquardt method is used to solve the two Least-Square Estimation problem similar to the implementation in Section 5.9. The only difference is the computation of  $\Theta|_{W^{(k)}}$  and  $\Omega|_{W^{(k)}}$ . The  $\Theta|_{W^{(k)}}$  is defined in Equation (6.2) and the  $\Omega|_{W^{(k)}}$  is defined in Equation (6.18).

### 6.4 Experimental Results for Better Optimal Self-Damped Design

The examples that were tested were constructed with High Performance MCM process technologies published by Frye [24]. The important parameters of the MCM process are listed in Table 6.1. In the case of the uniform width, all of the widths are equal to  $25\mu m$

for all of the examples tested. All of the drivers are modeled with a step input voltage source in series with the parallel combination of a  $12\Omega$  resistor and a  $4.3pF$  capacitor. All of the receivers in Examples 6.4.1 and 6.4.3 are modeled using a  $2.5pF$  capacitor. All of the receivers in Example 6.4.2 are modeled using a  $4.5pF$  capacitor. For all of the simulation results of the S-parameter macromodels, all of the far-end waveforms are simulated with the time-of-flight captured explicitly with the method stated in Section 5.5.

HIGH PERFORMANCE MCM-D	
Thickness of Dielectric ( $\mu m$ )	5
$\epsilon_{rel}$	3.2
Thickness of Metal ( $\mu m$ )	2.5
R ( $\Omega/\mu m$ ) for typical edge width	2.4
L ( $nH/\mu m$ ) for typical edge width	2.9
C ( $pF/\mu m$ ) for typical edge width	1.39
lower bound metal line width ( $\mu m$ )	10
typical metal line width ( $\mu m$ )	25
upper bound metal line width ( $\mu m$ )	50

Table 6.1: **The High Performance MCM Technologies Process Parameters**

	Maximum Path Delay (Uniform) (nS)	Lower Bound $\varphi$ (nS)	Maximum Path Delay (Optimal) (nS)	Upper Bound $\sqrt{m}\varphi$ (nS)	Percent Improvement (%)
Example 6.4.1	0.9874	0.5833	0.7182	1.4288	27.26
Example 6.4.2	0.9354	0.6266	0.8245	1.2532	11.85
Example 6.4.3	1.2308	0.8510	0.8533	2.0841	30.67

Table 6.2: **Comparison between the Uniform Width Design and the Optimal Design**





Example 6.4.2 is a network with a loop. Figure 6.8 shows the topology, lengths, and widths of the optimal design for all of the edges. The simulation waveforms of the receiver 4 with the maximum path delay for three different designs, namely the optimal design, old optimal design, and uniform width design are shown in Figure 6.8 (b). The maximum path delays, their respective bounds, and their percentage of improvements are listed in Tables 6.2 and 6.3. This novel optimal design has also been shown to have less overshoot than the old optimal design which tries to minimize overshoot by minimizing the maximum of all of the damping ratio errors together with the minimization of the maximum of all of the delays.

### 6.4.3 Example 3 A Clock Tree Network

Example 6.4.3 is a clock tree network shown in Zhu's ICCAD paper [59]. Figure 6.9 (a) shows the topology, lengths, and widths of the optimal design for all of the edges. All of the the simulation waveforms of the optimal design are shown in 6.9 (b), those of the old optimal design are shown in 6.9 (c), and those of the uniform width design are shown in Figure 6.9 (d). The maximum path delays, their respective bounds, and their percentage of improvements are listed in Tables 6.2 and 6.3.

The old optimization method presented in Chapter 5 demonstrated an improvement in the performance of this clock tree by reducing the maximum path delay and skew through minimization of the delays and the damping ratio errors. Although the skew was initially reduced to  $96pS$ , a further reduction to  $4.4pS$  can be achieved by using the novel optimization method which has the ability to minimize the RMS of all of the delays as well as the RMS of all of the differences between the delays and an arbitrary constant. This method not only guarantees the quality of the solution to the Min-Max delay problem but it also implies that the minimization of the skew has occurred.

The experimental results show that this optimization method reduces the maximum of the signal propagation delays considerably when compared to both the uniform width designs and the previous optimization designs presented in Chapter 5.





The results gathered here clearly demonstrate that this novel optimization method is most useful when the differences between all of the observable output variables are minimized. An excellent application of this optimization method, especially with this special characteristics, is in the minimization of the maximum path delay of the equal-path-length **H** clock tree.

## CHAPTER 7. Conclusions

The task of designing interconnect networks for today's high performance digital systems requires an accurate and a more efficient transient analysis of the coupled lossy transmission lines with frequency-dependent losses. Four novel macromodels of transmission lines and discontinuities have been developed for the purpose of a more efficient transient analysis. A novel method for the optimal self-damped design of interconnect networks on multichip modules (MCM) has also been presented.

We have presented here four new macromodels which extend and enhance the existing S-parameter based macromodels. The contributions in this thesis are:

- The construction of the macromodel for the congruence transformer.
- The computation of the moments of the S-parameters from the curve-fitted coefficients of the  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  data sets.
- The construction of the macromodel for the congruence transformer for the coupled transmission lines characterized with frequency-dependent parasitics.
- The indirect computation of the moments of the S-parameters from the measured data through first converting the measured S-parameter data into  $\mathbf{R}(f)$ ,  $\mathbf{L}(f)$ ,  $\mathbf{C}(f)$  and  $\mathbf{G}(f)$  data sets, and later finding the curve-fitted coefficients from the converted data sets.

These results are used in the transient analysis of:

- The coupled transmission lines (Chapter 2)
- The single and coupled transmission lines characterized with frequency-dependent losses (Chapter 3)

- The elements characterized with measured and tabulated S-parameter data (Chapter 4)

The experimental results show an excellent good agreement with other state-of-the-art simulators in addition to the clear efficiency advantages.

An optimal design technique for the design of the interconnects on MCMs has been presented. The self-damped design can be applied to any general network which may contain *loops*. The contributions in this thesis are the formulation of the self-damped design, the proofs of the tighter bounds between the original nonlinear programming problems and the Least-Square Estimation problems. The experimental results show significant improvements over the over-constrained uniform width design practice.

## CHAPTER 8. Future Work

In this research, only uniform transmission lines which support quasi-TEM wave propagation are investigated. The S-parameter macromodels of non-uniform transmission lines which support non-TEM wave propagation should be created in the future research.

For the creation of S-parameter macromodel of the measured data, only interconnections which have either two-port or  $2N$ -port configurations are considered. These result can be extended to handle general two-port elements such as open or gap. They should also include general  $N$ -port “black boxes” in future research.

Since the current S-parameter macromodel of the measured data can be used to extrapolate or interpolate the S-parameter value at any frequency point, it can be integrated seamlessly into the new frequency domain simulation method which evaluates S-parameters at every frequency of interest and then performs approximation at the very end.

The perturbation method used in finding the gradient matrix of the Least-Square Estimation optimization is a very time consuming one. There are two other methods, the incremental simulation and the sensitivity analysis which can be employed to further speed up the optimal design process. The use of S-parameter sensitivity analysis presents a great challenge because it only produces a voltage-time relationship for each parameter which does not bear a direct implication on the value of the partial derivatives. It requires a new way of thinking about the transformation and implementation of the Least-Square Estimation problem to solve the original Min-Max optimization problem. It also requires a new proof to show that the target function of the new LSE problem linearly bounds the target function of the original Min-Max problem. This topic is the most interesting and challenging among all of the topics in the future extensions of this thesis.

## Appendix A. Proof of the Tighter Bound

Given a *positive* function  $G(W, J)$  where  $W = \{w_i | i = 1 \dots n\}$  and  $J = \{j | j = 1 \dots m\}$ , one can define the target function of the Min-Max problem to be  $F(W, J)$ , and write:

$$\min_W [F(W, J)] \triangleq \min_W \left[ \max_J [G(W, J)] \right]. \quad (\text{A.1})$$

Denote  $G(W, j)$  to be the discrete value of  $G(W, J)$  at  $j$ . Define the estimation vector for the first LSE problem to be  $\Theta(W, J) \triangleq \{G(W, 1), G(W, 2), \dots, G(W, j), \dots, G(W, m)\}^T$ , where  $T$  denotes the transposition operation,  $j \in J$  and  $1 \leq j \leq m$ , and  $m$  is the number of the observable outputs. Define the summation of all squares of the first estimations,  $\Phi(W, J)$ , to be:

$$\Phi(w_1, w_2, \dots, w_n, 1, \dots, m) \triangleq \Theta(W, J)^T \Theta(W, J) = \sum_{j=1}^m [G(W, j)]^2 \quad (\text{A.2})$$

Define the Root-Mean-Square (RMS) of the first estimation as:

$$\varphi(w_1, w_2, \dots, w_n, 1, \dots, m) \triangleq \sqrt{\frac{\Phi(W, J)}{m}} = \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} \quad (\text{A.3})$$

Similarly, denote  $\Delta(W, j)$  to be the discrete value of  $\Delta(W, J)$  at  $j$ . Define the estimation vector for the second LSE problem to be:

$$\Omega(W, J) \triangleq \{c - G(W, 1), c - G(W, 2), \dots, c - G(W, j), \dots, c - G(W, m)\}^T, \quad (\text{A.4})$$

where  $T$  denotes the transposition operation. Define The summation of all squares of the second estimations,  $\Psi(W, J, c)$ , to be:

$$\Psi(w_1, w_2, \dots, w_n, 1, \dots, m, c) \triangleq \Omega(W, J)^T \Omega(W, J) = \sum_{j=1}^m [c - G(W, j)]^2 \quad (\text{A.5})$$

Define the Root-Mean-Square (RMS) of the second estimation as:

$$\psi(w_1, w_2, \dots, w_n, 1, \dots, m, c) \triangleq \sqrt{\frac{\Psi(W, J, c)}{m}} = \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \quad (\text{A.6})$$

One wants to prove that given a positive function  $G(W, J)$ , the minimum of the Root-Mean-Square (RMS) as defined in Equation (A.3), and the sum of the minimum of the RMS as defined in Equation (A.3) and  $\sqrt{m}$  multiplies the minimum of the RMS as defined in Equation (A.6) linearly bound the target function of the Min-Max problem defined in Equation (A.1). i. e.

$$\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c). \quad (\text{A.7})$$

From the Theorem 3, one has:

$$\varphi(W, J) \leq F(W, J). \quad (\text{A.8})$$

The second half of the equation is:

$$F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c). \quad (\text{A.9})$$

Substituting the definitions of  $\varphi$  and  $\psi$  into the above equation, one has:

$$\max_j [G(W, J)] \leq \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} + \sqrt{m} \cdot \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}}, \quad (\text{A.10})$$

or

$$\max_j [G(W, J)] \leq \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} + \sqrt{\sum_{j=1}^m [c - G(W, j)]^2}, \quad (\text{A.11})$$

The remainder of the proof of the theorem is to first establish the extreme value is in fact the global minimum, and later prove that this extreme value is equal to the left hand side of the Equation (A.11). If the global minimum is indeed equal to the left hand side of the Equation (A.11), then this Equation (A.11) holds for all value of  $c$ . Throughout the proof, because the case of the multiple outputs is analyzed here, it is assumed that there are at least *two* observable output variables.

It is apparent from Equation (A.7) that the choice of  $c$  determines how tight the bound will be. It can be shown that the  $c$  which makes  $\psi(W, J, c)$  assume the smallest value gives the tightest bound. In order to find such a constant  $c$ , one takes the partial derivative of the upper bound function w.r.t.  $c$ , one has:

$$\begin{aligned} \frac{\partial}{\partial c} [\varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c)] &= \frac{\partial}{\partial c} \varphi(W, J) + \sqrt{m} \cdot \frac{\partial}{\partial c} \psi(W, J, c) \\ &= 0 + \sqrt{m} \cdot \frac{\partial}{\partial c} \psi(W, J, c). \end{aligned} \quad (\text{A.12})$$

Setting the partial derivative equal to zero and solve, one can find the minimum of the right hand side of Equation (A.11). From the definition,

$$\psi(W, J, c) \triangleq \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}}. \quad (\text{A.13})$$

In order to find the extreme value of  $\psi(W, J, c)$  w.r.t.  $c$ , one has to set the first partial derivative of  $\psi(W, J, c)$  w.r.t.  $c$  equal to zero and solve. To find out whether it is a global minimum or global maximum, one needs to find out the sign of the second partial derivative of  $\psi(W, J, c)$  w.r.t.  $c$ . Take the first and second partial derivative of  $\psi(W, J, c)$  w.r.t.  $c$ , one has:

$$\frac{\partial \psi}{\partial c} = \frac{\sum_{j=1}^m \frac{2 \cdot [c - G(W, j)]}{m}}{\sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}}} = 0, \quad (\text{A.14})$$

$$\frac{\partial^2 \psi}{\partial c^2} = \frac{\sum_{j=1}^m \frac{2}{m}}{\left( \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \right)^3}. \quad (\text{A.15})$$

Solving for  $\frac{\partial \psi}{\partial c} = 0$ , one has:

$$\sum_{j=1}^m \frac{2 \cdot [c - G(W, j)]}{m} = 0. \quad (\text{A.16})$$

In order for the above equation to be true, the value of  $c$  must be

$$c = \text{avg}_j [G(W, J)]. \quad (\text{A.17})$$

To find out whether this extreme value of  $\psi(W, J, c)$  is a minimum or a maximum, one looks at the sign of  $\frac{\partial^2 \psi}{\partial c^2}$ . Simplify  $\frac{\partial^2 \psi}{\partial c^2}$ , one has:



$$\frac{\partial^2 \psi}{\partial c^2} = \frac{\sum_{j=1}^m \frac{2}{m}}{\left( \sqrt{\sum_{j=1}^m \frac{[c-G(W,j)]^2}{m}} \right)^3} = \frac{2}{\left( \sqrt{\sum_{j=1}^m \frac{[c-G(W,j)]^2}{m}} \right)^3} \quad (\text{A.18})$$

Because  $[c - G(W, j)]^2$  is always greater than or equal to zero, so  $\frac{\partial^2 \psi}{\partial c^2} > 0$ , which means  $\psi(W, J, c)$  is a concave upward function, and its value at  $c = \text{avg}_j [G(W, J)]$  is therefore a global minimum.

The following three cases list all the possible values of the choice of constant  $c$ :

- Case I:  $c \leq \text{avg}_j [G(W, J)]$
- Case II:  $c \geq \max_j [G(W, J)] \geq \text{avg}_j [G(W, J)]$
- Case III:  $\max_j [G(W, J)] \geq c > \text{avg}_j [G(W, J)]$

Because  $c = \text{avg}_j [G(W, J)]$  is a global minimum, one only need to find out the extreme value of the right hand side of Equation (A.11) in Case I.

Let

$$\text{avg}_j [G(W, J)] \triangleq \frac{\sum_{j=1}^m G(W, j)}{m}. \quad (\text{A.19})$$

Assume  $c \leq \text{avg}_j [G(W, J)]$ . It is known that Root-Mean-Square of a function is greater than or equal to the Mean of the function, that is:

$$\sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \geq \frac{\sum_{j=1}^m G(W, j)}{m} = \text{avg}_j [G(W, J)]. \quad (\text{A.20})$$

So

$$\sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} + \sqrt{\sum_{j=1}^m [c - G(W, j)]^2} \geq \frac{\sum_{j=1}^m G(W, j)}{m} + \sqrt{\sum_{j=1}^m [c - G(W, j)]^2}. \quad (\text{A.21})$$

Since  $c \leq \text{avg}_j [G(W, J)]$ , thus

$$\begin{aligned} \max_j [G(W, J)] - c &\geq \max_j [G(W, J)] - \text{avg}_j [G(W, J)], \\ \sqrt{\sum_{j=1}^m [c - G(W, j)]^2} &\geq \max_j [G(W, J)] - c, \\ \sqrt{\sum_{j=1}^m [c - G(W, j)]^2} &\geq \max_j [G(W, J)] - \text{avg}_j [G(W, J)]. \end{aligned}$$

Rearranging, one has:

$$\sqrt{\sum_{j=1}^m [c - G(W, j)]^2} + \text{avg}_j [G(W, J)] \geq \max_j [G(W, J)]. \quad (\text{A.22})$$

From Equation (A.21) and (A.22), if  $c \leq \text{avg}_j [G(W, J)]$ , one has:

$$\sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} + \sqrt{\sum_{j=1}^m [c - G(W, j)]^2} \geq \max_j [G(W, J)]. \quad (\text{A.23})$$

From Equations (A.23), one can write:

$$\begin{aligned} \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c) &= \sqrt{\sum_{j=1}^m \frac{[G(W, j)]^2}{m}} + \sqrt{m} \cdot \sqrt{\sum_{j=1}^m \frac{[c - G(W, j)]^2}{m}} \\ &\geq \max_j [G(W, J)] = F(W, J) \end{aligned} \quad (\text{A.24})$$

For a choice of  $c$  with  $c \leq \text{avg}_j [G(W, J)]$ , the minimum of  $\varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c)$  is equal to  $\max_j [G(W, J)] = F(W, J)$ , so

$$F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c). \quad (\text{A.25})$$

From Equations (A.8) and (A.25), one has:

$$\varphi(W, J) \leq F(W, J) \leq \varphi(W, J) + \sqrt{m} \cdot \psi(W, J, c).$$

## References

- [1] Lee Barford. On the measurement errors of the measured s-parameter data. Private communication, 1995.
- [2] C. E. Baumgartner. *Simulation methods for multi-conductor transmission lines in electronic applications*. PhD thesis, University of Arizona, Tucson, 1992.
- [3] M. Bayard. Synthesis of n-terminal pair networks. *Proc. Symposium on Modern Network Synthesis*, page 69, 1952.
- [4] Wendemagegnehu Beyene. Bi-level waveform relaxation analysis of package interconnects using scattering parameters. In *Proceeding of ISHM 92*, pages 156–161, 1992.
- [5] Thomas S. Blazek and Raj Mittra. Transient analysis of lossy multiconductor transmission lines in nonlinear circuits. *IEEE Trans. on Components, Hybrids and Manufacturing Technology*, CHMT-14(3):618–627, 1991.
- [6] John. R. Brews. Overshoot-controlled RLC interconnections. *IEEE Trans. on Electron Device*, ED-38:76–87, January 1991.
- [7] F. Y. Chang. Transient analysis of lossless coupled transmission lines in a nonhomogeneous dielectric medium. *IEEE Trans. on Microwave Theory and Techniques*, MTT-18, No. 9:616–626, Sep. 1970.
- [8] F. Y. Chang and Omar Wing. Transient simulation of lossy coupled transmission lines characterized with frequency-dependent parameters. *International Journal of High Speed Electronics*, 2(4):319–353, 1991.
- [9] F.Y. Chang. The generalized method of characteristic for waveform relaxation analysis of lossy coupled transmission lines. *IEEE Trans. on Microwave Theory and Techniques*, MTT-37, No. 12:2028–2038, 1989.
- [10] Norman H. J. Chang. Rational function curve-fitting for the transient analysis of measured s-parameter data. Private communication, 1995.
- [11] Eli Chiprout and M. Nakhla. Generalized moment-matching methods for transient analysis of interconnect networks. *9 th ACM/IEEE Design Automation Conference Proceedings*, pages 201–206, 1992.
- [12] Eli Chiprout and Michel Nakhla. Transient waveform estimation of high-speed mcm networks using complex frequency hopping. In *Proc. of IEEE Multi-Chip Module Conference MCMC-93*, pages 0–0, 1993.
- [13] Jason Cong, Kwok-Shing Leung, and Dian Zhou. Performance-driven interconnect design based on distributed RC delay model. Technical Report CSD-920043, University of Clifornia, Los Angeles, Oct. 1992.
- [14] Bradly J. Cooke, John Prince, and Andreas C. Cangellaris. S-parameter analysis of multiconductor, integrated circuit interconnect systems. *IEEE Trans. on Computer-Aided-Design*, 11, No. 3:353–360, 1992.
- [15] Cuthbert Daniel and Fred S. Wood. *Fitting Equations To Data*. Wiley-Interscience a Division of John Wiley & Sons, first edition, 1971.
- [16] Cuthbert Daniel and Fred S. Wood. *Fitting Equations To Data*. John Wiley & Sons, second edition, 1980.

- [17] John Joachim D'Azzo. *Feedback control system analysis and synthesis*. McGraw-Hill, New York, second edition, 1966.
- [18] Tom Dhaene and Daniel De Zutter. Selection of lumped element models for coupled lossy transmission lines. *IEEE Trans. on Computer-Aided Design*, 11(7):805,815, July 1992.
- [19] Antonije. R. Djordjevic, Tapan. K. Sarkar, and Roger. F. Harrington. Time-domain response of multiconductor transmission lines. *Proceedings of the IEEE*, 75(6):111–132, 1987.
- [20] J. Dobrowolski. *Introduction to Computer Methods for Microwave Circuit Analysis and Design*. Artech House, 1991.
- [21] William R. Eisenstadt and Yungseon Eo. S-parameter-based IC interconnect transmission line characterization. *IEEE Trans. on Components, Hybrids and Manufacturing Technology*, CHMT-15(4):483–490, August 1992.
- [22] O. I. Elgerd and W. C. Stephens. Effect of closed-loop transferfunction pole and zero locations on transient response of linear control systems. *AIEE*, pages 121–127, 1959.
- [23] R. C. Frye and K. L. Tai. Interconnection lines for wafer-scale-integrated assemblies. *U. S. Patent 4,703,288*, 1987.
- [24] Robert C. Frye. Physical scaling and interconnection delays in multichip module. *IEEE Trans. on CHMT*, 17(1):30–37, Feb. 1994.
- [25] Robert C. Frye and Howard Z. Chen. High speed interconnection using self-damped lossy transmission lines. In *Symposium on High Density Integration in Communication and Computer Systems*, pages 36–37, 1991.
- [26] David S. Gao, Andrew T. Yang, and Sung Mo Kang. Modeling and simulation of interconnect delays and crosstalks in high-speed integrated circuits. *IEEE Trans. on Circuits and Systems*, Vol. 37:1–9, January 1990.
- [27] Arthur A. Giordano and Frank M. Hsu. *Least Square Estimation with Applications to Digital Signal Processing*. John Wiley & Sons, 1985.
- [28] Colin Gordon, Thomas Blazek, and Raj Mittra. Time-domain simulation of multiconductor transmission lines with frequency-dependent losses. *IEEE Trans. on Computer-Aided Design*, 11(11):1372–1387, November 1992.
- [29] C. S. Gruodis and A. J. Chang. Coupled lossy transmission line characterization and simulation. *IBM J. Res. Develop*, 25(1):25–41, January 1981.
- [30] D. Hazony. *Elements of Network Synthesis*. Reinhold, New York, 1963.
- [31] C. W. Ho, D. A. Chance, C. H. Bajorek, and R. E. Acosta. The thin-film module as a high performance semiconductor package. *IBM Journal Res. Develop*, 26(286), 1982.
- [32] Rongxiang Julie Hu and Ernest S. Kuh. Integrating networks characterized by measured s-parameter data into swec. Technical Report Memorandum No. UCB/ERL M94/68, University of Clifornia, Berkeley, Aug. 1994.
- [33] Benjamin C. Kuo. *Automatic Control Systems*. Prentice-Hall, Inc., third edition, 1975.
- [34] Dmitri Kuznetsov and Jose Schutt-Aine. The optimal transient simulation of distributed lines, Part I: formulation of the optimal method. Submitted for publication in *IEEE Trans. of Circuits and Systems Fundamental Theory and Applications*.

- [35] Dmitri Kuznetsov and Jose Schutt-Aine. The optimal transient simulation of distributed lines, Part II: Difference approximation. Submitted for publication in IEEE Trans. of Circuits and Systems Fundamental Theory and Applications.
- [36] Charles L. Lawson and Richard J. Hanson. *Solving Least Square Problems*. Prentice-Hall Series in Automatic Computation. Prentice-Hall, Inc., 1974.
- [37] Haifang Liao and Wayne Dai. Extracting time-of-flight from scattering-parameter based macro-model. Technical Report 93-35, University of Clifornia, Santa Cruz, 1993.
- [38] Haifang Liao and Wayne Dai. Wave spreading evaluation of interconnect systems. In *Proc. of 3rd IEEE Multichip-module Conference*, pages 128–133, 1993.
- [39] Haifang Liao, Wayne Dai, Rui Wang, and F. Y. Chang. Scatering-parameter based macro model of distributed-lumped networks using expnontially decayed polynomial function. In *Proc. of 30th ACM/IEEE Design Automation Conference*, pages 726–731, 1993.
- [40] Haifang Liao, Rui Wang, Wayne Dai, and R. Chandra. Scatering-parameter based macro model of distributed-lumped networks using pade approximation. In *Proc. of International Symposium on Circuits and Systems*, 1993.
- [41] Shen Lin, Marek-Sadowska M., and E. S. Kuh. Swec: a stepwise equivalent conductance timing simulator for cmos vlsi circuits. *EDAC. Proceedings of the European Conference on Design Automation*, pages 142–148, Feburary 1991.
- [42] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of Society of Indust. Appl. Math.*, 11:431–441, 1963.
- [43] Edward M. Mikhail and F. Ackermann. *Observations and Least Squares*. IEP–A Dun-Donnelley Publisher, 1976.
- [44] Tuyen Nguyen. Efficient simulation of lossy and dispersive transmission lines. *31st th ACM/IEEE Design Automation Conference Proceedings*, pages 622–627, 1994.
- [45] Amir Owazr, M. Wickhmann, and M. Kasper. Line parameter characterization of laser deposited lines. In *Proc. of IEEE EPEP-94*, November 1994.
- [46] Fabio Romeo and Mauro Sanomauro. Time-domain simulation of  $n$  coupled transmission lines. *IEEE Trans. on Microwave Theory and Techniques*, MTT-35, No. 2:131–136, Feburary 1987.
- [47] R. Sanaie, Eli Chiprout, Michel S. Nakhla, and Qi-Jun Zhang. Integrating subnetworks characterized by measured data into moment-matching simulations. In *Proc. of IEEE Multi-Chip Module Conference MCMC-94*, pages 114–119, 1994.
- [48] Jose E. Schutt-Aine and Raj Mittra. Nonlinear transient analysis of coupled transmission lines. *IEEE Trans. on Microwave Theory and Techniques*, MTT-36, No. 7:959–966, 1989.
- [49] Stanley M. Shinnars. *Control system design*. Wiley, New York, 1964.
- [50] L. Miguel Silveria, Ibrahim M. Elfadel, and Jacob K. White. An efficient approach to transmission line simulation using measured of tabulated s-parameter data. *31st th ACM/IEEE Design Automation Conference Proceedings*, pages 634–639, 1994.
- [51] Boris Troyanovsky and Norman H. J. Chang. Hp Spice user manual. Private communication, 1994.

- [52] Jimmy Wang and Wayne Dai. Optimal design of self-damped lossy transmission lines in a tree network for multichip module. Technical Report UCSC-CRL-92-11, University of Clifornia, Santa Cruz, 1992.
- [53] Jimmy Wang and Wayne Dai. Optimal design of self-damped lossy transmission lines for multichip modules. Technical Report UCSC-CRL-93-43, University of Clifornia, Santa Cruz, 1993.
- [54] Jimmy Wang and Wayne Dai. Optimal design of self-damped lossy transmission lines for multichip modules. In *Proc. of IEEE ICCD-94*, October 1994.
- [55] Jimmy Wang and Wayne Dai. Transient analysis of coupled transmission lines using scattering parameter based macromodel. Technical Report UCSC-CRL-94-09, University of Clifornia, Santa Cruz, 1994.
- [56] Jimmy Wang and Wayne Dai. Transient analysis of coupled transmission lines using scattering parameter based macromodel. In *Proc. of IEEE EPEP-94*, November 1994.
- [57] D. Winklestein, M.B. Steer, and R. Pomerleau. Simulation of arbitrary transmission line networks with nonlinear terminations. *IEEE Trans. on Circuits and Systems*, 38(4):418–422, April 1991.
- [58] D. Zhou, F. Tsui, J. S. Cong, and D. S. Gao . A distributed-RLC model for MCM layout. In *Proc. of IEEE Multi-Chip Module Conference MCMC-93*, pages 191–197, 1993.
- [59] Qing Zhu and Wayne Dai. Optimal sizing of high speed clock networks based on distributed RC and lossy transmission line models. In *IEEE/ACM International Conference on Computer-Aided Design*, 1993.