© 2016 Tianjian Lu

## MULTIPHYSICS MODELING AND SIMULATION FOR LARGE-SCALE INTEGRATED CIRCUITS

BY

TIANJIAN LU

#### DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering in the Graduate College of the University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Jian-Ming Jin, Chair Professor José E. Schutt-Ainé Associate Professor Lynford L. Goddard Professor Philippe H. Geubelle

## ABSTRACT

This dissertation is a process of seeking solutions to two important and challenging problems related to the design of modern integrated circuits (ICs): the ever increasing couplings among the multiphysics and the large problem size arising from the escalating complexity of the designs. A multiphysicsbased computer-aided design methodology is proposed and realized to address multiple aspects of a design simultaneously, which include electromagnetics, heat transfer, fluid dynamics, and structure mechanics. The multiphysics simulation is based on the finite element method for its unmatched capabilities in handling complicate geometries and material properties. The capability of the multiphysics simulation is demonstrated through its applications in a variety of important problems, including the static and dynamic IR-drop analyses of power distribution networks, the thermal-ware highfrequency characterization of through-silicon-via structures, the full-wave electromagnetic analysis of high-power RF/microwave circuits, the modeling and analysis of three-dimensional ICs with integrated microchannel cooling, the characterization of micro- and nanoscale electrical-mechanical systems, and the modeling of decoupling capacitor derating in the power integrity simulations. To perform the large-scale analysis in a highly efficient manner, a domain decomposition scheme, parallel computing, and an adaptive time-stepping scheme are incorporated into the proposed multiphysics simulation. Significant reduction in computation time is achieved through the two numerical schemes and the parallel computing with multiple processors.

To all graduate students, who engineered an impact, no matter how small

## ACKNOWLEDGMENTS

My deepest gratitude goes to my advisor, Professor Jian-Ming Jin, who lead me into this field, guided me through the six years of discovery, and shaped me into who I am as a researcher today. It is difficult to overstate my appreciation. To me, he is much more than an advisor of profound knowledge and expertise. He educated me through his enthusiasm for research, working attitudes, disciplines, and persistence. By being his teaching assistant of his classic courses on electromagnetic theory and computational electromagnetics over the years, I am greatly inspired by the art of teaching. He could be a wonderful cook if not becoming a professor. The annual thanksgiving party at his house with the turkey he serves to his students is among the best parts of my graduate school experience. The values on work and life he instills in me only grows with my age.

My deepest gratitude extends to my thesis committee members, Professor José Schutt-Ainé, Professor Lynford Goddard, and Professor Philippe Geubelle. Professor José Schutt-Ainé exposed me to the field of signal/power integrity through his two wonderful courses, advanced signal integrity and microwave measurements. I also greatly benefited from serving as his teaching assistant during which time I found many interesting problems from industry and some of which turned out to be part of the dissertation. Professor Lynford Goddard held a weekly meeting with me for a collaborated project for almost two years. I start enjoying taking notes by observing him documenting details reported by every single student during the meeting on a black thick laboratory notebook. I really enjoyed the discussions with him and greatly benefited from his insights on physics, patience, and detail-oriented attitude. His valuable inputs to the two chapters of the dissertation on microchannel cooling and electrical-mechanical systems are greatly appreciated. I learned writing research proposals from Professor Philippe Geubelle by digesting his samples, which I strongly believe is a unique training especially in developing a big picture in this field. He encouraged me to incorporate structure mechanics into the multiphysics simulation after my preliminary exam. The few draft papers of formulas after our discussions in his office launched two chapters of the dissertation, one on the reliability analysis of interconnects and the other on micro- and nanoscale electrical-mechanical systems.

My deepest gratitude also goes to Mrs. Joanna Bao. To me, she is a constant source of support and encouragement. Her care has been significantly valuable to me especially during the difficult times when I was away from home. Though she does not contribute to technical details of the dissertation, she kept me sane on my path of making it possible. I am greatly indebted to her mother-like love and support over the years.

I am grateful to Professor Xudong Chen of the National University of Singapore. His guidance during my senior year project as an undergraduate student and his teaching on electromagnetics initiated my engagement in this field. I am indebted to his training in the early days and his mentorship over the years. I also want to thank Professor Weng Cho Chew for his advanced courses on electromagnetic theory. I am grateful that he shared his experience on several occasions with me. In many difficult situations I used his words to pace myself and calm down: "keep improving yourself through the work" and "success has multiple paths" are among the most enlightening ones. I am also grateful to Professor Martin Wong for the two courses through which I learned how to build solver engines of SPICE and draw the VLSI layout. These experiences expanded my knowledge in the field of integrated circuits, upon which I gradually develop an idea for how my research fits into the giant flow from silicon wafer all the way to electronic systems. I want to thank Professor Luke Olsen for his courses on numerical PDEs and multigrid algorithms. Through his two courses, I intensively practiced the different numerical schemes in solving PDEs, which paved my way later when I was exposed to different types of PDEs in the multiphysics simulation.

I would like to thank the former and current students in Professor Jin's group: Dr. Shih-hao Lee, Dr. Xiaolei Li, Peng Chen, Dr. Wang Yao, Dr. Mingfeng Xue, Dr. Huan-ting Meng, Dr. Su Yan, Jian Guan, Kedi Zhang, Yunjia Zeng, Yanan Liu, and Chao-Pao Chang as well as the visiting scholars Dr. Baolin Nie, Dr. Jin Ma, Dr. Xingchang Wei, Dr. Jian Li, and Geng Chen. They enlivened the fourth floor of Everitt Lab (the fifth floor of new ECE Building after 2014) with good electromagnetics and friendship. All the days

and nights we strived together to make a difference in our research will be long cherished. I would like to give special thanks to Dr. Shih-hao Lee since my work in this dissertation starts from digesting his codes and thesis on full-wave electromagnetic solvers. His codes are astonishingly beautiful and organized, which set a notably high standard to my own research in the following years. Special thanks goes to Dr. Xiaolei Li who in the early days tutored my research and constantly provided constructive suggestions. The best part of graduate school for me when the office lights were turned off are attributed to Dr. Mingfeng Xue, Dr. Wang Yao, Dr. Huan-ting Meng, Peng Chen, Dr. Baolin Nie, and Dr. Jin Ma and the good old times we spent together on soccer field, basketball court, swimming pools, and gyms. The habit of constantly hitting the gym to keep refresh from my research was developed with them.

Special thanks also goes to Dr. Zhiping Yang, Dr. Ken Wu, Dr. Shishuang Sun, and Dr. Hanfeng Wang for being great mentors during my internships at Google and Apple. They shared the invaluable opinions from the industry point of view to the multiphysics simulation in this dissertation. I have been given the highest level of freedom and respect in improving myself through the industrial attachments under their guidance. I would also like to thank Xu Chen for his generosity in sharing his seven years industrial experience at IBM with me and all the valuable discussions for my research.

Above all, I am indebted to my parents. I am extremely grateful for their endless love and support over the years while I was pursing my dreams. Although the completion of this dissertation has 12 years away from home, I will always believe that no matter where I am or how far away the journey in life takes me, I will remember the way back. Finally and most importantly, I would like to thank my wife, Feini, for her unwavering support, encouragement, patience, and love throughout our marriage. She is extremely patient and tolerant of my occasional vulgar mood. She taught me how to truly value myself as who I am. She built us a home wherever we lived. Thanks to Feini for being such a wonderful wife.

## TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION 1
1.1 Background and Motivation
1.2 Multiphysics Modeling and Simulation
1.3 Large-Scale Analysis
1.4 Organization 5
CHAPTER 2 EFFICIENCY ENHANCEMENT FOR LARGE-
SCALE ANALYSIS
2.1 Domain Decomposition $\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 7$
2.2 Parallel Implementation of FETI
2.3 Adaptive Time-Stepping $\ldots \ldots 12$
CHAPTER 3 ELECTRICAL-THERMAL CO-SIMULATION FOR
DC-IR DROP ANALYSIS
3.1 Problem Statement $\ldots \ldots 15$
3.2 Electrical-Thermal Co-Simulation
3.3 FETI-Enabled Parallel Computing
3.4 Three-Dimensional Power Delivery
CHAPTER 4 TRANSIENT ELECTRICAL-THERMAL CO-SIMULATION
FOR 3-D POWER DISTRIBUTION NETWORKS
4.1 Problem Statement
4.2 Electrical-Thermal Co-Simulation
4.3 Efficiency Enhancement $\dots \dots \dots$
4.4 Transient Electrical-Thermal Behaviors of PDNs 45
CUADTED 5 THEDMAL AWADE HICH EDECHENCY CUAD
ACTEDIZATION OF LADCE SCALE THOUSENED SILICON
ACTERIZATION OF LARGE-SCALE THROUGH-SILICON-
$VIA \text{ SIRUUTURES} \dots \dots$
$\begin{array}{cccccccccccccccccccccccccccccccccccc$
5.2 Electrical-Thermal Co-Simulation
5.3 Computation Information
5.4 Design Parameters Investigation

CHAPTER 6 ELECTRICAL-THERMAL CO-SIMULATION FOR ANALYSIS OF HIGH-POWER RF/MICROWAVE CIRCUITS 86
6.1 Problem Statement
6.2 Electrical-Thermal Co-Simulation
6.3 Model Verification and Demonstration
6.4 Co-Simulation Examples
CHAPTER 7 MULTIPHYSICS SIMULATION OF 3-D ICs WITH
INTEGRATED MICROCHANNEL COOLING
7.1 Problem Statement
7.2 Co-Simulation and Efficiency Enhancement
7.3 Numerical Example
7.4 Conclusion $\ldots \ldots 129$
CHAPTER 8 COUPLED ELECTRICAL-THERMAL-MECHANICAL
SIMULATION FOR THE RELIABILITY ANALYSIS OF LARGE-
SCALE 3-D INTERCONNECTS
8.1 Problem Statement
8.2 Coupled Simulation and Parallel Computing
8.3 Numerical Examples
8.4 Conclusion $\ldots \ldots 151$
CHAPTER 9 MULTIPHYSICS MODELING IN MICRO- AND
NANO-ELECTROMECHANICAL SYSTEMS
9.1 Problem Statement $\dots \dots \dots$
9.2 Simulation Methodology $\ldots \ldots 155$
9.3 Design of a Suspended Resistor
9.4 Results and Discussion
9.5 Discussion $\ldots \ldots 167$
CHAPTER 10 MULTIPHYSICS MODELING OF CAPACITOR
DERATING IN POWER INTEGRITY SIMULATION 169
10.1 Problem Statement
10.2 Capacitor Derating Mechanism
10.3 Capacitor Derating Models
10.4 System-Level Simulation and Measurement Correlation 183
10.5 $\tilde{\text{Conclusion}}$
CHAPTER 11 CONCLUSION AND FUTURE WORK
11.1 Conclusion
11.2 Future Work
11.3 Refinement on a Single Event
11.4 Improvement through the Couplings of Multiscale Events 196
11.5 Advancement from Simulation to Design
REFERENCES

# CHAPTER 1 INTRODUCTION

## 1.1 Background and Motivation

The advancement of the integrated circuit (IC) technology has been driven by the goals to increase performance and functionality and decrease cost and power. Over the years, the goal has been achieved by the continuous scaling of transistors and interconnects. The scaled transistors are faster due to their smaller intrinsic capacitance and the shorter channel length, whereas the scaled interconnects are longer and thinner. As with the continuous scaling of ICs, the increment of the interconnect delay tends to override the reduction of the transistor gate delay as shown in Fig. 1.1 such that there is improvement in the overall performance gained from scaling saturates. In addition, the continuous scaling increases the power density, and Fig. 1.2 shows that the power density of a CPU has reached  $100 \text{ W/cm}^2$ . An investigation [1] conducted with the 130 nm technology node shows interconnects contribute to approximately 50% of the power consumption of a microprocessor, which is expected to rise to 80% [2]. The same investigation [1] found out that 90% of the total power is consumed by 10% of the interconnects. The increased power density and the non-uniform distribution of power consumption function together and bring about serious thermal issues.

A remedy of the interconnect scaling problem is the three-dimensional (3-D) technology in which multiple planar devices are stacked up vertically to achieve a reduction of global interconnections. Other benefits of 3-D technology include increased packing density, reduced chip area, and the possibility of heterogeneous integration.

A key component of 3-D technology is the through-silicon-via (TSV) technique [3, 4, 5]. A schematic of a 3-D IC enabled by the TSV technique is shown in Fig. 1.3. However, the increased power density together with the



Figure 1.1: Gate and interconnect delay of different technology nodes from [6].

reduced chip area leads to a significant increase of heat power generation per unit area and poses great challenges for heat removal, which is further complicated by the interlayer dielectrics of very low thermal conductivity.

As the continuous scaling of ICs exacerbates the thermal issues, the thermal issues in return challenge the electrical design in many aspects. High temperature degrades the performance and longevity of devices, which may further result in malfunctions and deterioration in the performance of the entire system. Thermal stress due to the large temperature gradient and mismatch of thermal expansion coefficients between the package and chip may lead to mechanical failures, for example, chip cracks. The electromigration, recognized as the primary mode of metalization failure, has an exponential dependence on temperature. In addition, the increased temperature often impairs the quality of the signal and power and creates signal and power integrity concerns such as the thermally induced clock skew [7] and unexpected variance on voltage drops. Due to the ever increasing influence of thermal issues on electrical designs and the strong couplings between the two, it is necessary to have a multiphysics-based computer-aided design methodology that addresses the electrical designs and thermal issues simultaneously.



Figure 1.2: Power density of CPUs in the recent decades from [8].



Figure 1.3: Schematic of a 3-D integrated circuit.

## 1.2 Multiphysics Modeling and Simulation

A report presented to the U.S. Department of Energy (DOE) in 2008 [9] accentuates the aforementioned challenges with the following:

Today's problems, unlike traditional science and engineering, do not involve physical processes covered by a single traditional discipline of physics or the associated mathematics. Complex systems encountered in virtually all applications of interest to DOE involve many distinct physical processes.

... The issue of coupling models of different events at different scales and governed by different physical laws is largely wide open and represents an enormously challenging area for future research.

In this dissertation, we propose and implement a multiphysics simulation to address such a challenge in modern electronic systems. As illustrated by Fig. 1.4, the multiphysics simulation integrates an electrical analysis and a thermal analysis through an iterative scheme and targets the applications ranging from on-chip level, to package level, and to board level. The multiphysics simulation is based on the finite element method (FEM) for its unmatched capabilities in modeling complex geometries and material properties [10]. It is worth mentioning that a considerable amount of research has been devoted to searching for innovative heat removal techniques, for example, the integrated microchannel cooling [11, 12]. To that end, the multiphysics simulation incorporates the conjugate heat transfer and fluid analyses to take into account the forced convective liquid cooling by the integrated microchannels. Another important phenomenon that should never be neglected in an electrical design is the thermally-induced stress [13]. Thermal stresses of substantial amplitudes usually arise from large temperature gradients and mismatches of thermal expansion coefficients. Mechanical fatigue failures can occur because of repeated and concentrated high thermal stresses. Therefore, a mechanical analysis based on the theory of linear thermo-elasticity is also integrated into the multiphysics simulation. As the governing equations associated with individual physics and the coupling schemes among the multiphysics vary among the types of problems, the details of implementing the multiphysics simulation are provided based on the specific problems in the following chapters.



Figure 1.4: Multiphysics simulation with its existing capabilities and applications. The demonstrated applications on the bottom row from left to right include 3-D power distribution network, on-chip power grid, solder bump array, and substrate-integrated waveguide filter.

## 1.3 Large-Scale Analysis

In order to accurately resolve the 3-D geometrical complexities and the nonuniform material properties, three-dimensional simulations are preferred. In addition to the challenge resulting from the strong couplings among the multiphysics, the large problem size arising from the volumetric discretization to the circuit designs of escalating complexity creates another challenge in the modeling and simulation of modern integrated circuits. To enhance the efficiency and relieve the computation burden of a large-scale analysis, a domain decomposition scheme, parallel computing, and an adaptive time-stepping scheme are incorporated into the proposed multiphysics simulation.

## 1.4 Organization

The remainder of the dissertation is organized as follows. Chapter 2 provides detailed formulations of the efficiency-enhancement techniques adopted in the multiphysics simulation. Chapters 3 and 4 present the static and dynamic IR-drop analyses for power distribution networks (PDNs), in which thermal effects are taken into account. Chapters 5 and 6 demonstrate the

thermal-aware high-frequency characterizations in the multiphysics simulation through the applications of through-silicon-via structures and highpower RF/microwave circuits. Chapter 7 introduces a coupled electricalthermal-fluid simulation in analyzing three-dimensional ICs with integrated microchannel cooling. Chapters 8 and 9 explore the multiphysics simulation accounting for thermal stress, which enables the reliability analysis of interconnects on mechanical fatigue and failures and the mechanical characterization of micro- and nanoscale structures in electro-mechanical systems. Chapter 10 illustrates the derating issue of the decoupling capacitors and the proposed solution of incorporating multiphysics models of decoupling capacitors in the system-level simulation of PDNs. Along with the development of the multiphysics simulation, measurement correlations are provided for both uniphysics and multiphysics. Chapter 11 concludes the dissertation and proposes future work in this area.

## CHAPTER 2

## EFFICIENCY ENHANCEMENT FOR LARGE-SCALE ANALYSIS

The multiphysics simulation is based on the finite element method (FEM) because of its unmatched capabilities in modeling complex geometries and materials [10]. In addition, the accuracy of the FEM solution can be improved systematically by increasing the order of its basis functions. Although the FEM has many unique advantages, its volumetric discretization for many real and practical integrated circuit (IC) designs can easily result in a large linear system of millions of unknowns. Solving such a large system can be computationally intensive and the solution has to be performed repeatedly for many iterations in the co-simulation. In this dissertation, we employ two techniques in the co-simulation, namely, an adaptive time-stepping scheme in the transient thermal analysis and a domain decomposition scheme in both electrical and thermal analyses, in order to alleviate the computational burden for solving large-scale problems. The domain decomposition used in this dissertation is called the finite element tearing and interconnecting (FETI) [10, 14]. By applying the FETI, the entire computational domain is divided into smaller subdomains, which are completely decoupled initially and coupled only in the solution of the global interface system. Therefore, the FETI is highly suitable for parallel computing with multiple processors.

### 2.1 Domain Decomposition

The fundamental block in the multiphysics simulation is solving the matrix equation

$$[A] \{x\} = \{b\} \tag{2.1}$$

where [A] contains the information from both the stiffness and damping matrices,  $\{b\}$  includes the excitations in general or in the transient analysis the excitations of both the current time step and the information from the pre-

vious one, and  $\{x\}$  represents the unknown coefficients to be solved, which can be temperature  $\{T\}$  in the thermal analysis, or electric potential  $\{\phi\}$  in the IR-drop analysis, or electric field components  $\{E\}$  in the full-wave electromagnetic analysis. As conveyed by its name, the implementation of FETI can be generalized into three steps: tearing, interconnecting, and recovering subdomain solutions.

#### 2.1.1 Tearing

By applying FETI, the entire computation domain is torn into  $N_s$  nonoverlapping subdomains. The neighboring subdomains  $\Omega^i$  and  $\Omega^j$  share the interface denoted as  $\Gamma_{ij}$ . The continuity of field components has to be enforced at the subdomain interfaces. In the thermal analysis, the Dirichlettype and Neumann-type boundary conditions

$$T_i = T_j \tag{2.2}$$

$$\hat{n}^i \cdot \kappa_i \nabla T_i = -\hat{n}^j \cdot \kappa_j \nabla T_j = \Lambda_T \tag{2.3}$$

are applied to enforce the continuity of the temperature and the heat flux, where  $\kappa$  denotes the thermal conductivity,  $\hat{n}^i$  is the unit normal vector at  $\Gamma_{ij}$ pointing from  $\Omega^i$  to the exterior region, and  $\Lambda_T$  is the unknown Neumann boundary data associated with temperature. In the IR-drop analysis, the Dirichlet-type and Neumann-type boundary conditions

$$\phi_i = \phi_j \tag{2.4}$$

$$\hat{n}^i \cdot \sigma_i \nabla \phi_i = -\hat{n}^j \cdot \sigma_j \nabla \phi_j = \Lambda_\phi \tag{2.5}$$

are applied to enforce the continuity of the voltage and the current, where  $\sigma$  denotes the electrical conductivity and  $\Lambda_{\phi}$  is the unknown Neumann boundary data associated with electrical potential. In the full-wave electromagnetic analysis, the Dirichlet-type and the Neumann-type boundary conditions

$$\hat{n}^i \times (\hat{n}^i \times \vec{E}^i) = \hat{n}^j \times (\hat{n}^j \times \vec{E}^j)$$
(2.6)

$$\hat{n}^i \times \left(\frac{1}{\mu^i} \nabla \times \vec{E}^i\right) = -\hat{n}^j \times \left(\frac{1}{\mu^j} \nabla \times \vec{E}^j\right) = \vec{\Lambda}_E$$
 (2.7)

are applied to enforce the tangential continuity of the electric and magnetic field components, where  $\mu$  denotes the permeability and  $\vec{\Lambda}$  is the unknown Neumann boundary data associated with electric field.

Next, the finite element discretization is applied to the Neumann boundary conditions in Equations (2.3), (2.5), and (2.7), yielding the terms

$$\{\lambda_T^s\} = [B_T^s]^T \{\lambda_T\} = \int_{\Gamma^s} \{N^s\} \Lambda_T \,\mathrm{d}\Omega \tag{2.8}$$

$$\left\{\lambda_{\phi}^{s}\right\} = \left[B_{\phi}^{s}\right]^{T}\left\{\lambda_{\phi}\right\} = \int_{\Gamma^{s}}\left\{N^{s}\right\}\Lambda_{\phi} \,\mathrm{d}\Omega \tag{2.9}$$

$$\{\lambda_E^s\} = [B_E^s]^T \{\lambda_E\} = \int_{\Gamma^s} \{\vec{N}^s\} \cdot \vec{\Lambda}_E \, \mathrm{d}\Omega \tag{2.10}$$

where  $[B_T^s]^T$ ,  $[B_{\phi}^s]^T$ , and  $[B_E^s]^T$  are signed boolean matrices that extract the interface degrees of freedom (DOFs) of the s<sup>th</sup> subdomain;  $\{\lambda_T\}$ ,  $\{\lambda_{\phi}\}$ , and  $\{\lambda_E\}$  are the global dual variables, called the Lagrange multipliers, associated with temperature, electric potential, and electric field, respectively; and  $\{N^s\}$ and  $\{\vec{N}^s\}$  are column vectors in the s<sup>th</sup> subdomain, containing the scalar and vector interpolation functions, respectively. With the vector interpolation functions, the vector electric field in each element *e* can be expanded as

$$\vec{E}^{e} = \sum_{i=1}^{n} \vec{N}_{i}^{e} E_{i}^{e}$$
(2.11)

Similarly, with the scalar interpolation functions, the temperature in each element e can be expanded as

$$T^{e} = \sum_{i=1}^{n} N_{i}^{e} T_{i}^{e}$$
(2.12)

Note that in Equations (2.8), (2.9), and (2.10), even though the same notation  $\Gamma^s$  is employed to represent the subdomain interfaces for simplicity, the computation domain and the way to divide it into subdomains may differ among the different analyses.

#### 2.1.2 Interconnecting

With all the boundaries including the subdomain interfaces, the linear system of the  $s^{\rm th}$  subdomain becomes

$$[A^{s}] \{x^{s}\} = \{b^{s}\} - \{\lambda^{s}\}$$
(2.13)

The procedure of obtaining subdomain matrix  $[A^s]$  is identical to that in finding the global one in Equation (2.1). The only difference in terms of the expression between the global system in Equation (2.1) and the subdomain system in Equation (2.13) is the additional term  $\{\lambda^s\}$  in Equation (2.13), arising from the Neumann boundaries specified at the subdomain interfaces.

Finally, the system equations from all the subdomains can be rearranged into a global system as follows

$$\begin{bmatrix} [A^{1}] & \dots & 0 & [B^{1}]^{T} \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & [A^{N_{s}}] & [B^{N_{s}}]^{T} \\ [B^{1}] & \dots & [B^{N_{s}}] & 0 \end{bmatrix} \begin{bmatrix} \{x^{1}\} \\ \vdots \\ \{x^{N_{s}}\} \\ \{\lambda\} \end{bmatrix} = \begin{bmatrix} \{b^{1}\} \\ \vdots \\ \{b^{N_{s}}\} \\ 0 \end{bmatrix}$$
(2.14)

It is worth mentioning that the last equation in Equation (2.14)

$$[B^{1}] \{x^{1}\} + [B^{2}] \{x^{2}\} + \dots + [B^{N_{s}}] \{x^{N_{s}}\} = 0$$
 (2.15)

results from the enforcement of the continuity condition in Equations (2.2), (2.4), and (2.6) at the subdomain interfaces. Eliminating  $\{x^s\}$ ,  $s = 1, 2, ..., N_s$ results in an interface system of  $\{\lambda\}$  as

$$[F] \{\lambda\} = \{d\} \tag{2.16}$$

where

$$[F] = \sum_{s=1}^{N_s} [B^s] [A^s]^{-1} [B^s]^T$$
(2.17)

$$\{d\} = \sum_{s=1}^{N_s} [B^s] [A^s]^{-1} \{b^s\}$$
(2.18)

The thermal, IR drop, and electromagnetic analyses have individual La-

grange multipliers and hence different interface systems. Clearly, with the FETI, the original problem is reduced to an interface problem of a much smaller dimension. The interface system in Equation (2.16) can be solved using a Krylov subspace method such as the biconjugate gradient stabilized (BiCGSTAB) method [15] together with the Dirichlet preconditioner [16]. The Dirichlet preconditioner offers a good approximation of the inverse of [F] in Equation (2.16) as

$$\begin{bmatrix} F_D^{-1} \end{bmatrix} = \sum_{s=1}^{N_s} \begin{bmatrix} B^s \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & \begin{bmatrix} S_{II}^s \end{bmatrix} \begin{bmatrix} B^s \end{bmatrix}^T$$
(2.19)

where

$$[S_{II}] = [A_{II}^s] - [A_{IV}^s] [A_{VV}^s]^{-1} [A_{IV}^s]^T$$
(2.20)

and the subscripts V and I denote the volume and interface unknowns, respectively.

#### 2.1.3 Recovering Subdomain Solutions

The dual variables  $\{\lambda_T^s\}$ ,  $\{\lambda_{\phi}^s\}$ , and  $\{\lambda_E^s\}$  are obtained after iteratively solving the interface problems in Equation (2.16) associated with individual analyses. In the step of recovering subdomain solutions, the temperature, electric potential, and electric field unknowns of individual subdomains can be computed with the obtained Neumann-type boundary condition through Equation (2.13).

### 2.2 Parallel Implementation of FETI

In this section, the parallel implementation of FETI in the multiphysics simulation using Message Passing Interface (MPI) is provided. It has been shown that FETI is highly parallizable [17]. As described in the above section, three steps are involved in the implementation of FETI, namely, tearing, interconnecting, and recovering subdomain unknowns:

1. By tearing, the computation domain is divided into subdomains; similar numbers of unknowns over subdomains are preferred which will lead to balanced loads among processors. In the step of tearing, individual subdomains should also have the corresponding system matrix  $\{A\}$  and excitation vector  $\{b\}$  assembled; the factorization of the system matrices of individual subdomains should also be performed at this step. It can be seen that subdomains are completely independent of each other in the tearing step.

- 2. The interconnecting step deals with the interface system for solving the unknown Neumann-type boundary data. The interface system is solved with the BiCGSTAB method which consists of a series of matrix-vector products: subdomains should contribute to the matrix-vector product individually and the information should be broadcast to the rest immediately when all the subdomains have it ready. The interconnecting step can be achieved completely in parallel except at the broadcasting process.
- 3. As for the step of recovering, each subdomain is supposed to utilize its own boundary condition including the obtained Neumann-type boundary data, independently and completely in parallel.

The efficiency enhancement of FETI with multiple processors will be demonstrated with specific applications in the following chapters.

## 2.3 Adaptive Time-Stepping

It is proposed in [18, 19] that the step size selection in the numerical solution of ordinary differential equations can be viewed as an automatic control problem. The algorithm introduced by [18, 19] can be described as

$$h_n = \beta \left(\frac{tol}{r_n}\right)^{1/k} h_{n-1} \tag{2.21}$$

where h is the step size,  $r_n$  denotes the error associated with step n, tol is a user-input tolerance, and k is related to the order of the integration method. A major concern associated with the algorithm illustrated by Equation (2.21) is that when the stability limits the step size rather than the accuracy, this algorithm will produce a highly oscillating sequence of step size [20]. The step size selection based on the algorithm of proportional-integral (PI) control was



Figure 2.1: A PID controller in a feedback loop control.

proposed by [21] to resolve the aforementioned issue, which is described by

$$h_n = \left(\frac{tol}{r_n}\right)^{k_I} \left(\frac{r_{n-1}}{r_n}\right)^{k_P} h_{n-1} \tag{2.22}$$

where  $k_P$  and  $k_I$  are constant parameters related to the proportional and integral terms of a PI controller, respectively.

The adaptive time-stepping in this dissertation is based on the algorithm of proportional-integral-derivative (PID) control [22, 23]. A block diagram of a PID controller in a feedback loop is illustrated by Fig. 2.1. The PID control algorithm can be explained by

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt}$$
(2.23)

where u denotes the control signal, e represents the control error, and  $K_p$ ,  $K_i$ , and  $K_d$  are the coefficients of the proportional, integral, and derivate terms, respectively. The transfer function of a PID controller can thus be written as

$$G(s) = K_p + K_i \frac{1}{s} + K_d s$$
 (2.24)

It can be seen from the transfer function that the integral term contributes to the reduction of the steady-state errors by the low-frequency compensation and the derivative term contributes to the enhancement of the transient response by the high-frequency compensation.

Based on the PID control, the two consecutive time steps in the adaptive time-stepping scheme employed in this dissertation have the relation

$$\Delta t_{n+1} = \left(\frac{e_{n-1}}{e_n}\right)^{k_P} \left(\frac{1}{e_n}\right)^{k_I} \left(\frac{e_{n-1}^2}{e_n e_{n-2}}\right)^{k_D} \Delta t_n \tag{2.25}$$

The three constant parameters in Equation (2.25) are chosen as  $k_P = 0.075$ ,  $k_I = 0.175$ , and  $k_D = 0.01$ , following the studies in [22] and [23]. In Equation (2.25),  $e_n$  is the maximum relative change of temperature at time  $t_n$ , denoted as

$$e_n = \max(\{e_T\}_n) \tag{2.26}$$

where

$$\{e_T\}_n = \frac{1}{\tau_T} \frac{\|\{T\}_n - \{T\}_{n-1}\|}{\|\{T\}_n\|}$$
(2.27)

and  $\tau_T$  is the user input tolerance and its default value in the implementation is 1.0. There are two major advantages of this adaptive time-stepping. The first advantage is the automatic generation of non-uniform time steps. Because of the first-order nature of the thermal response, the time step enlarges itself as temperature approaches a steady state. Therefore, it requires many fewer time steps than the regular time-marching scheme with uniform time steps. The other advantage is that the adaptive time-stepping provides a very convenient way to detect a steady state. In the implementation, the steady state is declared and the simulation is terminated when the time step is greater than a certain portion of the specified simulation duration. The efficiency enhancement in the multiphysics simulation through the adaptive time-stepping scheme will be demonstrated in the following chapters with specific problems.

## CHAPTER 3

## ELECTRICAL-THERMAL CO-SIMULATION FOR DC-IR DROP ANALYSIS

### 3.1 Problem Statement

As current flows through the resistances of a power distribution network (PDN), a voltage drop takes place across the effective path between the voltage supply and a certain circuit component, which is also known as DC IR-drop. DC IR-drop is one of the primary concerns in the design of power distribution networks. Unintentional variances in the voltage drops may prevent the transistors from switching states and cause circuits failures [24], [25]. As the continuous scaling of interconnects, especially in the 3-D technology where the through-silicon-via (TSV) technique enables the vertical stack-up of multiple planar devices, the problems associated with DC IR-drop become more significant. On one hand, the supply voltage continues to drop with the increased scaling. The trend of the reduction of the supply voltage is plot in Fig. 3.1 based on the projection from the International Technology Roadmap of Semiconductors [26]. The decrease of supply voltage is accompanied by the reduction of noise margin such that the devices are more vulnerable to power noise [27], [28]. On the other hand, with the continuous integration of integrated circuits, the power density is expected to go beyond  $100 \text{ W/cm}^2$ in 2016 [26]. The decreased supply voltage and the increased power density function together and force a large amount of current flowing through the power distribution network. Joule heating effect is likely to cause a significant rise in temperature if the heat generated from the large amount of current flowing through interconnects is not removed efficiently [29], [30]. Because of the temperature-dependent resistivity, an accurate prediction of the DC IR-drop under the tightened noise margin requires considering the temperature influence. Therefore, it is necessary to have the electrical-thermal co-simulation for DC IR-drop analysis.



Figure 3.1: Power supply voltage by years projected by the International Technology Roadmap of Semiconductors [26].

The electrical-thermal co-simulation for DC IR-drop analysis has been realized through the finite volume method [31], the latency insertion method [32], and the equivalent circuit model [33]. In this dissertation, the finite element method (FEM) is adopted due to its capability of dealing with complex geometries and material properties [10]. This dissertation is not limited to the electrical-thermal co-simulation, but also aims to develop a more efficient approach for large-scale problems. One way of improving the efficiency while dealing with large-scale problems is to use domain decomposition; in this work, the finite element tearing and interconnecting (FETI) method [14, 34, 35, 36, 37] is employed. By FETI, the entire computational domain is decomposed into non-overlapping subdomains of smaller size which allows the possibility of parallel computing. The continuity enforcement at subdomain interfaces through Dirichlet and Neumann boundary conditions via Lagrange multipliers yields a reduced-order interface problem. The interface problem can be solved by a Krylov subspace method, and the obtained solutions can serve as the boundary condition to recover the unknowns within individual subdomains. Significant reduction in computation time can be achieved by utilizing the highly parallelizable nature of FETI through parallel computing with multiple processors

### 3.2 Electrical-Thermal Co-Simulation

Current flowing through interconnects obeys Ohm's law

$$\boldsymbol{J} = \sigma \boldsymbol{E} \tag{3.1}$$

where  $\sigma$  is the temperature-dependant electrical conductivity and the electric field  $\boldsymbol{E}$  can be expressed in terms of the electrical potential  $\phi$ 

$$\boldsymbol{E} = -\nabla\phi \tag{3.2}$$

The governing equation of the DC IR-drop analysis in terms of the electrical potential  $\phi$  can be obtained by substituting Equations (3.1) and (3.2) into

$$\nabla \cdot \boldsymbol{J} = 0 \tag{3.3}$$

The boundary-value problem of the DC IR-drop analysis consists of two parts: the governing equation

$$\nabla \cdot \sigma \nabla \phi = 0 \tag{3.4}$$

and the boundary conditions

$$\phi = \phi_c \quad \text{on } \Gamma_{\rm vc} \tag{3.5}$$

$$\sigma \frac{\partial \phi}{\partial n} = \frac{\phi}{RS} \quad \text{on } \Gamma_{\text{load}} \tag{3.6}$$

where  $\Gamma_{vc}$  is the Dirichlet boundary,  $\Gamma_{load}$  is the impedance boundary of an external load, and R and S are the resistance and the cross-sectional area of the impedance boundary, respectively.

The boundary-value problem of the steady-state thermal analysis consists of two parts: the governing equation

$$\nabla \cdot k \nabla T = -Q \tag{3.7}$$

where k is the thermal conductivity, T is the temperature distribution, and



Figure 3.2: Coupling between the voltage distribution and temperature profile in the electrical-thermal co-simulation for DC IR-drop analysis.

Q is the volumetric heat generation; and the boundary conditions

$$T = T_c \qquad \text{on } \Gamma_{\rm tc} \qquad (3.8)$$

$$k\frac{\partial T}{\partial n} = -h(T - T_a) \quad \text{on } \Gamma_{\text{conv}}$$
(3.9)

where  $\Gamma_{tc}$  is the isothermal boundary,  $\Gamma_{conv}$  is the convection boundary, h and  $T_a$  are convective heat transfer coefficient and ambient temperature, respectively.

The volumetric heat generation Q in Equation (3.7) includes the heat generated from Joule heating effect  $P_{\text{joule}}$  as

$$P_{\text{joule}} = \boldsymbol{J} \cdot \boldsymbol{E} = \sigma \boldsymbol{E}^2 \tag{3.10}$$

The temperature-dependant resistivity can be written as

$$\rho = \rho_0 \left[ 1 + \alpha (T - T_0) \right] \tag{3.11}$$

where  $\rho_0$  is the resistivity at temperature  $T_0$  and  $\alpha$  is the temperature coefficient of the material. The electrical analysis and thermal analysis are coupled through Equations (3.10) and (3.11) as shown in Fig. 3.2. The iterative scheme of solving Equations (3.4), (3.7), (3.10), and (3.11) concurrently is illustrated in Fig. 3.3.

A two-layered PCB board [31] in Fig. 3.4(a) is taken as a benchmark example to verify the implementation of heat conduction and convection, as well as the iterative scheme of electrical-thermal co-simulation. The same INPUT initial resistivity  $\rho_0$ , tolerance tol, number of iterations N OUTPUT voltage distribution and temperature profile

Step 1 Set i = 1 and  $\tilde{\rho} = \rho_0$ 

Step 2 While  $i \leq$  number of iterations do Steps 3-8

Step 3 Compute voltage via electrical analysis

- Step 4 Calculate power
- Step 5 Compute temperature via thermal analysis
- Step 6 Update resistivity  $\rho$
- Step 7 If  $|\rho \tilde{\rho}| < tol$

Output and Exit

- Step 8 Set i = i + 1 and  $\tilde{\rho} = \rho$
- Step 9 Output ("Not converged after N iterations")

Figure 3.3: The iterative scheme of electrical-thermal co-simulation.

example with its analytical solutions has been used in [31] as verifications for the electrical-thermal co-simulation implemented with the finite volume method. It is worth mentioning that the direct solver in Intel MKL PARDISO is employed whenever it is necessary to perform matrix factorization and solve linear systems. The planar dimension of the PCB board is  $5\times 2 \text{ cm}^2$ . The thicknesses of the top copper plane, the middle layer of FR-4, and the bottom copper plane are 0.05 mm, 0.5 mm, and 0.05 mm, respectively. Heat source of 50 W is attached on the top surface of the PCB board, where air convection also applies. The ambient temperature and the temperature on the bottom surface of the PCB board are set at 293 K. Since the heat source is attached on the entire top surface and the planar dimension is much larger than the vertical one, this problem can be treated as a one-dimensional heat transfer problem with heat conducting vertically. The analytical solution of the temperature T on the top surface of the PCB board can be obtained through

$$T = T_c + P \cdot R \tag{3.12}$$



Figure 3.4: Two-layered PCB: (a) geometry, (b) temperature varying with convection coefficients.

where  $T_c$  is the ambient temperature, P is the heat source, and R is the overall thermal resistance of the PCB board with the conduction resistance and the convection resistance in parallel to each other. The variation of the temperature on the top layer with different convection coefficients from both the simulation and analytical solution is plotted in Fig. 3.4(b).

As a further verification, the same structure as shown in Fig. 3.4(a) is considered and the planar dimension of the board is changed to  $10 \times 5$  cm<sup>2</sup> and the thicknesses of the copper and the dielectric layers to 36  $\mu$ m and 350  $\mu$ m, respectively. Voltage source is attached on one end of the top copper plane. Voltage distributions and temperature profiles of the top layer with and without Joule heating effect are compared with those in [31] and shown in Fig. 3.5.



Figure 3.5: Electrical-thermal co-simulation of two-layered PCB: (a) voltage distribution and (b) temperature profile.

## 3.3 FETI-Enabled Parallel Computing

In this section, the parallel implementation of FETI in the electrical-thermal co-simulation using the Message Passing Interface (MPI) is achieved. The efficiency of FETI with multiple processors is demonstrated with numerical







Figure 3.6: Voltage distribution at  $z = 30 \ \mu m$  of two-layered of TSV array: (a) the geometry of the two-layered TSV array, (b) IR Drop (unit in V) without considering Joule heating, and (c) IR Drop (unit in V) with Joule heating.

examples. All computations are performed with Intel Xeon 2.67 GHz hexcore processors.



Figure 3.7: Voltage distribution at  $z = 390 \ \mu m$  of a 12-layered TSV array: (a) the geometry of the 12-layered TSV array, (b) IR Drop (unit in V) without considering Joule heating, and (c) IR Drop (unit in V) with Joule heating.

#### 3.3.1 Two-Layered TSV Array

Consider the two-layered TSV array in Fig. 3.6(a). It is used to investigate the parallel efficiency of FETI with MPI. In Fig. 3.6(a), the copper-filled TSV has radius 8  $\mu$ m and height 30  $\mu$ m. The thickness of the copper plane is 5  $\mu$ m. The separation between two neighboring TSVs is 30  $\mu$ m along the x-axis and 34  $\mu$ m along the y-axis, both of which are measured between the centers of the vias. Assume that the bottom surface of the copper plane has constant temperature 313 K and is exposed to convection cooling with ambient temperature of 293 K and the convective heat transfer coefficient is 10 W/m<sup>2</sup>K. Voltage of 1.2 V is applied on the upper edge of the lower plane. Loads of 40  $\Omega$  are connected through the vias on the top layer. This example has a total of 2.03 million DOFs.

Table 3.1: Speed-up of FETI with MPI

Total Number of DOFs: $2.03 \times 10^6$					
Number of processors	4	8	16	32	
Number of subdomains	4	8	16	32	
Number of DOFs per subdomain	$5.1  imes 10^5$	$2.6\times 10^5$	$1.4\times 10^5$	$6.4\times10^4$	
Total computation time (seconds)	166.74	64.04	26.76	14.70	

Table 3.2: Unit-load-per-processor parallel efficiency of FETI with MPI

Total number of DOFs	$2.6  imes 10^5$	$5.1  imes 10^5$	$1.0  imes 10^6$	$2.0  imes 10^6$
Number of processors	4	8	16	32
Number of subdomains	4	8	16	32
Total computation time (seconds)	13.30	14.05	13.85	14.70

First, we explore the trend of the reduction in computation time with the increase in the number of processors. Four cases are generated for comparison where the entire structure is broken into 4, 8, 16, and 32 subdomains, respectively. Each subdomain is assigned to one processor. The computation time of all four cases is recorded in Table 3.1. The speed-up is defined as the ratio of the computation time of two different cases. It is observed that when the number of processors is doubled from 4 to 8, a speed-up of 2.6 is achieved. The reason that the speed-up is larger than 2 lies in the factorization of the subdomain matrices; in the case of 4 subdomains, the time taken to factorize the matrix is far more than twice of that for the case of 8 sub-domains. It is also observed that when the number of processors is doubled from 2.7 This is because with the increased number of subdomains, the size of the interface problem and the time consumed in communication over different processors increase accordingly. In this investigation, by continuously doubling the number of processors from 4 to 32 while maintaining the size of the problem, speed-up around 2 is achieved. This is due to the high scalability of FETI.

The second investigation of parallel efficiency is carried out based on the idea of unit load per processor where the number of DOFs assigned to individual processors is kept the same. There are 32 columns of vias along the x-axis. Each column of vias forms one subdomain and is assigned to one processor. Four cases are generated for comparison; the sizes of the problems and the computation time are recorded in Table 3.2. It is observed that the four cases of different numbers of DOFs take relatively the same amount of computation time. From the case of four subdomains to the case of 32 subdomains, the total number of DOFs is increased by eight times whereas the computational time has only an increment of 1.4 seconds. It can be drawn from Table 3.2 that the FETI-enabled parallel implementation with MPI is capable of solving a large-scale problem with a similar amount of time to that of a much smaller problem by simply employing more processors.

The voltage distributions at  $z = 30 \ \mu \text{m}$  with and without Joule heating effect are depicted in Figs. 3.6(b) and 3.6(c): without considering Joule heating, the voltage drop on the plane of  $z = 30 \ \mu \text{m}$  is 7 mV; with Joule heating, the voltage drop is 9 mV; the thermal effect on the voltage drop is 28.57%. As a reference, this two-layered TSV array problem is solved with FEM directly by the direct solver in Intel MKL PARDISO; the computation time is 573.65 seconds with one processor and 352.35 seconds with 12 processors where the 12 processors are used by Intel MKL PARDISO at the stage of solving the linear system.

#### 3.3.2 Twelve-Layered TSV Array

In this section, a 12-layered TSV array shown in Fig. 3.7(a) is used to demonstrate the overall performance of FETI with MPI. The dimensions of vias and the planes are identical to that in Fig. 3.6(a) except that the array in each layer has the size of  $10 \times 32$ . Resistive loads of 10  $\Omega$  are connected through vias on the top layer. The entire structure is divided into 32 subdomains with a total number of 4.1 million DOFs. The number of processors applied in this example is 32 and the total computation time is 28.39 seconds. The voltage distributions at  $z = 390 \ \mu \text{m}$  with and without the consideration of Joule heating are depicted in Figs. 3.7(b) and 3.7(c), where a voltage drop due to the thermal effect can be observed.

### 3.4 Three-Dimensional Power Delivery

A three-dimensional power delivery generally consists of global and local power distribution networks. The global network is responsible for delivering power to different IC strata and the local network feeds the power to devices in each IC stratum [38, 39, 40, 41]. In this section, large-scale local and global power delivery are taken as examples to explore their voltage distributions and fluctuations with Joule heating and the efficiency of cooling systems.

#### 3.4.1 On-Chip Power Grid

Figures 3.8(a) and 3.8(b) show a section of the on-chip power grid. Power and ground come in pairs and there are 24 pairs in total along the x-direction. The detailed information on the dimensions other than those illustrated in Fig. 3.8(b) includes: the cross-sectional area of the line of power and ground is  $1 \ \mu m \times 1 \ \mu m$ ; the separation between the power and ground line in each pair is 1  $\mu$ m; the height of vias is 2  $\mu$ m. The copper-made power grid is immersed in the silicon substrate. Heat sink and air cooling convection boundary are attached on the top and the bottom of the structure, respectively. The supply voltage is attached to the power lines on the bottom of the power grid. Voltage drops measured on the top of the power grid with different cooling temperatures of the heat sink and the air convection cooling are depicted in Fig. 3.9(a). With the load impedance remaining the same, the higher voltage drop over the interconnect should be observed in the case of the higher supply voltage, which is demonstrated in Fig. 3.9(a). Besides, based on the slopes of the four lines in Fig. 3.9(a), the voltage drop in the case of the higher supply voltage tends to be more sensitive with respect to the efficiency of the cooling system. Another issue to be addressed based on this analysis of the on-chip power grid is the relation between the thermal impact on the voltage drop and the supply voltage as shown in Fig. 3.9(b). In Fig. 3.9(b), the voltage drop at the room temperature of 25 °C corresponding to each



Figure 3.8: On-chip power grid from (a) the three-dimensional view and (b) the top view (vias are marked in red). Note only a section of the power grid under investigation is shown in Fig. 3.8.




Figure 3.9: Electrical-thermal co-simulation results of the on-chip power grid in Fig. 3.8: (a) voltage drop versus cooling temperature and (b) thermal impact on the voltage drop.

individual supply voltage is taken as the reference. As in the four cases of different supply voltages, the resistivity shares the same linear dependency on temperature, the four curves are expected to overlap with each other which is shown in Fig. 3.9(b).

It is worth mentioning that the spatial discretization to the on-chip power grid results in a total number of 1.29 million unknowns, which is divided into 12 subdomains with each paired up with one processor. With the 12 processors computing in parallel, it takes 56.5 seconds to obtain the voltage distributions for a single iteration of the electrical-thermal co-simulation.

#### 3.4.2 Redistribution Layer

Figure 3.10 shows an example of the redistribution layer (RDL). RDL generally consists of vias and horizontal traces and is used in the flip-chip assembly as interconnects between different functional layers. The details of the dimensions of the RDL in Fig. 3.10 are: the diameter and the height of the copper via are 40  $\mu$ m and 120  $\mu$ m, respectively; the via pitch is 60  $\mu$ m; the length, the width, and the thickness of the horizontal copper traces are 450  $\mu$ m, 60  $\mu$ m, and 4  $\mu$ m, respectively; the transition at the center of the traces has a length of 50  $\mu$ m; vias are immersed in the silicon substrate. The entire structure is divided into 60 subdomains with each domain consisting of one trace, two power vias, and two ground vias. Air convection cooling and heat sink are attached to the substrate on each end of the traces.

The total number of unknowns to be solved is 2.03 million. With a total number of 60 processors, the computation time of one iteration of the electrical analysis is 14.66 seconds. The residual tolerance of solving the interface problem generated by FETI is set as  $10^{-5}$  and it takes four iterations to converge in a single iteration of the electrical analysis.

The voltage difference between the two power vias varies. The voltage drop due to the thermal effect is plotted in Fig. 3.11. It is worth mentioning that the voltage drop is measured along the horizontal traces. Results obtained from the conventional FEM is plotted in the same figure for comparison. With the cooling temperature set at 313 K and the voltage difference between the two power vias set as 0.8 volt, it takes 12 iterations for the electricalthermal co-simulation to converge when the residual error drops below the



Figure 3.10: Redistribution layer.

predefined tolerance of  $10^{-4}$ . It can be seen from Fig. 3.11 that as the temperature increases from 313 K to 333 K the voltage along the horizontal traces decreases. This indicates that the vias are more vulnerable to the thermal effect rather than the horizontal traces in this example.

#### 3.4.3 Global Power Delivery

In Fig. 3.12(a), four layers of on-chip power grid, each having a dimension of  $512 \times 1536 \times 4 \ \mu m^3$ , stack up and connect to each other with TSVs (copper) and solders (carbon steel). The current starts from the solder at the bottom, flows through the vias in the Si interposer, to the micro-solders, and is drawn into the on-chip power grid. Heat sink and air cooling convection boundary are attached on the top and the bottom of the structure, respectively. Assume that the heat sink has the temperature of 40 °C and air cooling functions at the room temperature of 25 °C. Three floorplans shown in Figs. 3.12(b)-3.12(d) are applied where the red is set to be 0.9 V and the blue to 0.5 V. It is worth mentioning that in Figs. 3.12(b) and 3.12(c), the via pitch is 128  $\mu$ m and in Fig. 3.12(d), it is changed to 80  $\mu$ m.

The number of unknowns resulting from the spatial discretization is 1.21 million. With FETI, the computational domain is divided into 12 subdomains and handled by 12 processors in parallel. It takes 72.1 seconds to calculate the voltage distribution for a single iteration in the electrical-thermal co-simulation. It is worth mentioning that this example has a much larger interface problem to be solved than the one for the on-chip power grid in the Section 3.4.2, which also explains that with the similar number of unknowns, the same number of subdomains, and the same number of processors, it takes



Figure 3.11: Thermal effect in the redistribution layer with different cooling temperatures and voltage between power vias.

longer time to solve this example of the global power network.

The voltage distributions under the three different floorplans at the second layer to the bottom are shown in Fig. 3.13. It can be observed that Floorplan C produces the most uniform voltage distribution with the smallest difference between the maximum and the minimum, whereas Floorplan B delivers the highest voltage among the three types of floorplans.









Figure 3.12: Global power delivery network (a) and three different floor plans, namely, (b) Floorplan A, (c) Floorplan B, and (d) Floorplan C.









Figure 3.13: Voltage distributions on the top of the power grid of the second stratum to the bottom under (a) Floorplan A, (b) Floorplan B, and (c) Floorplan C.

## CHAPTER 4

# TRANSIENT ELECTRICAL-THERMAL CO-SIMULATION FOR 3-D POWER DISTRIBUTION NETWORKS

## 4.1 Problem Statement

In a three-dimensional (3-D) power distribution network (PDN), on-chip power grids of multiple dies form a vertical stack-up through micro-bumps and through-silicon vias (TSVs). The minimum diameter of a TSV can be 1-2  $\mu m$  in 2011-2014 and 0.8-1.5  $\mu m$  in 2015-2018 [26]. The continuous integration challenges the design of the power distribution network in many aspects. It causes the increase of the power density as interconnects consume about 50% of the microprocessor power in the 130-nm technology node (in the year 2002) [1]. The high power density exacerbates the thermal issues and results in very high temperature, which in turn degrades the performance of both transistors and interconnects. Additionally, the unintentional variance in the voltage drops associated with the temperature rise has to be taken into account especially with the lowering supply voltage and the reducing noise margin; otherwise it may prevent the transistors from switching states and causes circuits failures [24], [25]. Moreover, the continuous integration induces a large amount of current flowing through the interconnects, which may suffer from significant electromigration [42]. Because of the temperaturedependent material properties, the electrical-thermal co-simulation [3] has to address all the aforementioned issues simultaneously.

In Chapter 3, electrical-thermal co-simulation has been developed to analyze the static voltage drops with the finite element method in [43]. The dynamic variance in the supply voltage not only leads to high temperature but also produces a large temperature gradient. Therefore, it is critically important to understand the transient electrical-thermal behaviors of the power distribution networks. In this chapter, we develop the electrical-thermal cosimulation in the transient regime based on the finite element method (FEM)



Figure 4.1: Three-dimensional power delivery network with through-silicon vias.

[10] for its capability of modeling complex circuit geometries and materials[44], [45].

As we know, the distribution of the power has to be designed globally over the entire power distribution network. In the real world, the power distribution network may involve a large-scale on-chip power grid connected through massively coupled TSVs to the arrays of solder bumps, a portion of which is shown in Fig. 4.1. Accordingly, it is desired that the co-simulation is not limited on a single via [46] or device [47] but extended to solve large-scale problems. To deal with large-scale problems, we introduce a domain decomposition scheme called the finite element tearing and interconnecting (FETI) [14], [35]. The FETI algorithm is highly parallelizable in nature [17] and its parallel computing attains a very efficient co-simulation in the time domain, through which we can simulate the transient electrical-thermal behaviors of large-scale power distribution networks including on-chip power grids, solder bump arrays, and TSV-based power distribution network. With the obtained transient behaviors, we are able to detect the localized overheating and to insure the voltage drops fluctuate within design specifications. We can also investigate the impacts of different types of input pulses, power maps, and via pitches on the design of a PDN.

## 4.2 Electrical-Thermal Co-Simulation

In this section, the implementation of the electrical-thermal co-simulation is described, followed by the verification with two numerical examples.

#### 4.2.1 Formulation with the Finite Element Method

The continuity of the electric charge can be expressed by

$$\nabla \cdot \left( \vec{J} + \frac{\partial \vec{D}}{\partial t} \right) = 0 \tag{4.1}$$

and the conduction current  $\vec{J}$  is in the form of

$$\vec{J} = \sigma \vec{E} \tag{4.2}$$

where  $\sigma$  is the conductivity of the materials. The constitutive equation of a dielectric medium can be written as

$$\vec{D} = \epsilon \vec{E} \tag{4.3}$$

where  $\epsilon$  is the permittivity of the materials. For completeness, the displacement current representing the capacitive effect is included in the formulation. The electric field  $\vec{E}$  can be written in terms of the electric potential  $\phi$  as

$$\vec{E} = -\nabla\phi \tag{4.4}$$

With Equations (4.1) to (4.4), one obtains the governing equation of the electrical analysis

$$\nabla \cdot \epsilon \nabla \frac{\partial \phi}{\partial t} + \nabla \cdot \sigma \nabla \phi = 0 \tag{4.5}$$

The boundary conditions associated with the electrical analysis include the Dirichlet boundary

$$\phi = \phi_d \tag{4.6}$$

and the impedance boundary of an external load

$$\hat{n} \cdot \sigma \nabla \phi = \frac{\phi}{RS} \tag{4.7}$$

where R and S are the resistance and the cross-sectional area of the impedance boundary, respectively.

The governing equation associated with the transient heat conduction in

a three-dimensional solid body is

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot \kappa \nabla T + Q \tag{4.8}$$

where  $\rho$  is the density of the materials, c is the specific heat,  $\kappa$  is the thermal conductivity, and Q is the volumetric heat generation. The corresponding boundary conditions associated with the transient thermal analysis includes the Dirichlet boundary

$$T = T_d \tag{4.9}$$

and the air convection boundary

$$\hat{n} \cdot \kappa \nabla T = -h(T - T_a) \tag{4.10}$$

where h and  $T_a$  are the convective heat transfer coefficient and the ambient temperature, respectively.

Applying Galerkin's method to Equation (4.8), one obtains

$$[C_T]\left\{\dot{T}\right\} + [K_T]\left\{T\right\} = \{f_T\}$$
(4.11)

where  $[C_T]$  is the damping matrix and  $[K_T]$  is the stiffness matrix. The matrices and vectors can be written as

$$[C_T]_{i,j} = \int_V \rho c N_i N_j dV \tag{4.12}$$

$$[K_T]_{i,j} = \int_V \kappa \nabla N_i \cdot \nabla N_j dV + \int_{\Omega_c} h N_i N_j d\Omega \qquad (4.13)$$

$$\{f_T\}_i = \int_V QN_i dV + \int_{\Omega_c} hT_a N_i d\Omega \qquad (4.14)$$

where  $N_i$  represents the interpolating function such that

$$T = \sum_{i=1}^{n} N_i T_i \tag{4.15}$$

For the temporal discretization, the backward difference is used

$$\dot{T} = \frac{T(t) - T(t - \Delta t)}{\Delta t} + O(\Delta t)$$
(4.16)

which yields an unconditionally stable system, regardless of the size of  $\Delta t$ 

[10]. The spatial and temporal discretizations to Equation (8.2) in the electrical analysis follow similar procedures.

	$\begin{array}{c} \text{Electrical} \\ \text{conductivity} \\ (\text{S m}^{-1}) \end{array}$	$\begin{array}{c} \text{Temperature} \\ \text{coefficient} \\ (\text{K}^{-1}) \end{array}$	$\begin{array}{c} \text{Thermal} \\ \text{conductivity} \\ (\text{W m}^{-1}\text{K}^{-1}) \end{array}$	Density $(\text{kg m}^{-3})$	Specific heat capacity $(J kg^{-1}K^{-1})$
Aluminum	$5.98  imes 10^7$	$3.9  imes 10^{-3}$	400	$8.7  imes 10^3$	385
Copper	$3.77 \times 10^7$	$3.9  imes 10^{-3}$	160	$2.7 \times 10^3$	900
Nickel	$1.43 \times 10^7$	$6.0 \times 10^{-3}$	91	$8.9\times10^3$	440
Solder	$6.67 \times 10^6$	$4.2 \times 10^{-3}$	50	$9.0  imes 10^3$	150

Table 4.1: Material properties of metal

Table 4.2: Material properties of dielectric

	Relative permittivity	$\begin{array}{c} \text{Thermal} \\ \text{conductivity} \\ (\text{W m}^{-1}\text{K}^{-1}) \end{array}$	$\begin{array}{c} \text{Density} \\ \text{(kg m}^{-3}) \end{array}$	Specific heat capacity $(J kg^{-1}K^{-1})$
Substrate	12.1	Equation $(8.32)$	$2.3\times 10^3$	703
Insulation	4.2	Equation $(8.33)$	$2.2  imes 10^3$	730
Underfill	2.1	Equation $(8.33)$	$2.2 \times 10^3$	703

The electrical and thermal analyses are coupled through the dissipated power and the temperature-dependant material properties. The dissipated power that provides the source for Joule heating can be written as

$$P_{\text{joule}} = \boldsymbol{J} \cdot \boldsymbol{E} = \sigma \boldsymbol{E}^2 \tag{4.17}$$

and the material properties involved in the chapter are listed in Tables 4.1 and 4.2 for metal interconnects and dielectrics, respectively. The electrical conductivity in Table 4.1 is given at room temperature. The temperaturedependent resistivity of metal interconnects can be written as

$$\rho = \rho_0 \left[ 1 + \alpha (T - T_0) \right] \tag{4.18}$$

where  $\rho_0$  is the resistivity at temperature  $T_0$  and  $\alpha$  is the temperature coefficient of the metal interconnects. The temperature-dependent thermal conductivity of the silicon substrate is given in [48] as

$$\kappa_{\rm Si} = 2.475 \times 10^5 \ T^{-1.3} \ (273 {\rm K} \le T \le 1000 {\rm K})$$
(4.19)

Both the insulation layer attached to the TSVs and the underfill surrounding the solder bumps are considered as silicon dioxide and its temperaturedependent thermal conductivity is found in [47] as

$$\kappa_{\rm SiO_2} = 1.02278 + 0.00121 \ T \ (273 \text{K} \le T \le 600 \text{K})$$
 (4.20)

#### 4.2.2 Model Verification

Two examples are used to verify the proposed co-simulation model: one is the on-chip power grid in Fig. 4.2 and the other is the solder bump array in Fig. 4.4. Results obtained through COMSOL Multiphysics [49] are used as reference solutions.

#### **On-Chip** Power Grid

Figure 4.2 shows two segments of a on-chip power grid, which consists of three metal layers (M1, M2, and M3). All three metal layers are immersed in the silicon substrate. The width (a\_sub) and the length (b\_sub) of each segment are 80  $\mu$ m and 50  $\mu$ m, respectively. The detailed dimensions of the three metal layers are labeled in Fig. 4.2(b): the width (a\_M1) and the thickness (b\_M1) of M1 are 3  $\mu$ m and 1  $\mu$ m, respectively; both the width (a\_M2) and the thickness (b\_M2) of M2 are 1  $\mu$ m; the width (a\_M3) and the thickness (b\_M3) of M3 are 3  $\mu$ m and 1  $\mu$ m, respectively; the via connecting M1 to M3 has a height of 6  $\mu$ m; the via connecting M1 to M2 has a height of 3  $\mu$ m; the cross-section of the via is 1×1  $\mu$ m<sup>2</sup>. All three metal layers are made of copper. A Gaussian pulse in the form of

$$f(t) = f_0 e^{-(t-t_0)^2/\tau^2}$$
(4.21)

with  $f_0 = 0.8$  V,  $t_0 = 80$  ns, and  $\tau = 30$  ns is launched into the on-chip power grid through terminals 1 and 2 as labeled in Fig. 4.2(a), whereas terminals 3 and 4 are ground. It is worth mentioning that  $f_0$  is the peak voltage,  $t_0$  is the position of center of the peak, and  $\tau$  controls the pulse width. The maximum temperature in the structure is depicted in Fig. 4.3(a) as a function of time, and the power dissipation is shown in Fig. 4.3(b), both of which match very well with those obtained through COMSOL Multiphysics.



Figure 4.2: Two segments of a on-chip power grid: (a) 3-D view and (b) side view.

#### Solder Bump Array

Figure 4.4 shows a solder joint array. The electrodes are made of nickel, which are immersed in solders and connected through the copper-made redistribution layer. The detailed information of the solder joint array is: the radius (r\_bump) and the height (h\_bump) of the solder bump are 30  $\mu$ m and 40  $\mu$ m, respectively; the length (a\_electrode) and thickness (t\_electrode) of the electrode are 20  $\mu$ m and 10  $\mu$ m, respectively; the width (a\_trace), length (b\_trace), and thickness (t\_trace) of the trace are 60  $\mu$ m, 220  $\mu$ m, and 5  $\mu$ m, respectively. The width (a\_underfill) and length (b\_underfill) of the underfill are 160  $\mu$ m and 320  $\mu$ m, respectively. A rectangular pulse with the rising time of 20 ns and the falling time of 20 ns is sent into the structure through



(a)



Figure 4.3: On-chip power grid: (a) maximum temperature occurring in the structure and (b) power dissipation of the structure.



Figure 4.4: Geometry of a solder joint array.

terminal 2, which is labeled in Fig. 4.4, whereas terminals 1 and 3 are ground. The maximum and minimum of the rectangular pulse are 1.2 V and 0.2 V, respectively. The maximum temperature matches well with that obtained through COMSOL Multiphysics as shown in Fig. 4.5.

## 4.3 Efficiency Enhancement

The transient electrical-thermal co-simulation can be very computationally intensive: the time-marching procedures require fine time steps to accurately depict the transient behaviors of the system under given excitations; within each time step, it involves several Newton-type iterations to solve the nonlinear problem; and a single Newton-type iteration solves linear systems from both electrical and thermal analyses. In this section, we incorporate FETI into the transient co-simulation to enable parallel computing with multiple processors, which significantly reduces the computation time.

### 4.3.1 Verification of the Incorporation of FETI

The same on-chip power grid in Fig. 4.2 is used to verify the implementation of FETI. The Gaussian pulse involved in this example is in the form of Equation (4.21) with  $\tau = 20$  ns. The maximum temperatures obtained



Figure 4.5: Maximum temperature in the solder joint array.

through FETI with two subdomains and COMSOL Multiphysics are depicted in Fig. 4.6(a). Another example uses the same solder bump array in Fig. 4.4 with the rising time of the rectangular pulse changed from 10 ns to 20 ns. The maximum temperatures obtained through FETI with three subdomains and FEM are depicted in Fig. 4.6(b). Both examples show excellent agreement and hence verify the incorporation of FETI into the transient co-simulation.

#### 4.3.2 Parallel Efficiency Demonstration

The implementation of FETI consists of three steps: tearing, interconnecting, and recovering subdomain solutions. In the step of tearing, individual subdomains have the corresponding system matrix and excitation vector assembled and the system matrices factorized, which are completely independent to each other. The interconnecting step solves the interface system with the BiCGSTAB method, which consists of a series of matrix-vector products. Subdomains contribute to their own portion of the matrix-vector product and acquire the remaining portion through communication with other subdomains. The interconnecting step can be carried out completely in parallel except for the communication among subdomains. In the step of recover-





(b)

Figure 4.6: Verification of the implementation of FETI: (a) maximum temperature in the on-chip power grid in Fig. 4.2 with  $\tau=20$  ns and (b) maximum temperature in the solder bump array in Fig. 4.4 with the rising time of 20 ns.



Figure 4.7: On chip power grid consists of 32 segments, each of which has detailed geometrical information in Fig. 4.2(b).

ing, each subdomain utilizes its own boundary condition with the obtained Lagrange multipliers, independently and completely in parallel. The demonstration of the parallel efficiency is based on the idea of unit load per processor, under which the number of DOFs assigned to individual processors is kept the same. All computations are performed with Intel Xeon 2.67-GHz hex-core processors.

The on-chip power grid shown in Fig. 4.7 is taken as the example to demonstrate the parallel efficiency of FETI. Four cases are generated for comparison, namely, the power grids consisting of 4, 8, 16, and 32 segments. The computation time taken by factorization and four different time steps is recorded in Table 4.3. The total number of time steps remains 100 for all four cases. It is observed that as the total number of DOFs increases from  $1.8 \times 10^5$  to  $1.5 \times 10^6$ , the total computation time holds almost the same when the number of processors is increased from 4 to 32. The slight increase in the computation time is due to the necessary communication among the processors during the step of interconnecting. It can be drawn from Table 4.3 that the FETI-enabled parallel computing is capable of solving a very large-scale problem with a similar amount of time to that of a much smaller one by simply employing proportionally more processors.

## 4.4 Transient Electrical-Thermal Behaviors of PDNs

In this section, we characterize the transient electrical-thermal behaviors of various types of PDNs including large-scale on-chip power grid, solder bump arrays, and TSV-based PDN. With the 3-D co-simulation, we are able to

Total number of DOFs	$1.8  imes 10^5$	$3.6\times 10^5$	$7.3\times10^5$	$1.5\times 10^6$
Number of processors	4	8	16	32
Number of subdomains	4	8	16	32
Factorization in electrical analysis (sec)	7.2	6.8	7.2	7.1
Factorization in thermal analysis (sec)	7.2	7.1	7.1	7.2
$1^{\rm st}$ time step (sec)	5.1	5.1	5.4	5.5
$20^{\text{th}}$ time step (sec)	12.9	13.7	14.3	16.6
$50^{\text{th}}$ time step (sec)	13.0	14.5	15.4	17.6
$90^{\rm th}$ time step (sec)	10.3	10.3	12.7	12.7
Total computation time (min)	19.0	20.3	22.5	24.8

Table 4.3: Unit-load-per-processor parallel efficiency of FETI with MPI

visualize the thermal hot spots and depict the fluctuation of the voltage drops. We also investigate the impacts on the electrical-thermal behaviors from different types of voltage pulses, power maps, and via pitches.

#### 4.4.1 On-Chip Power Grid

In the Section 4.3, we demonstrated the parallel efficiency of FETI using the example of the on-chip power grid shown in Fig. 4.7. Here, we analyze the transient thermal behaviors of the on-chip power grid under various types of input and the corresponding thermal impact on the voltage drops. Temperatures and voltages at five points (from A to E as marked in Fig. 4.2(a)) in each segment are recorded. Points B and D are located at the two ends of the vias connecting M1 and M2, whereas points C and E are located at the two ends of the vias connecting M1 and M3. The notation A2 is adopted as a reference to point A of the second segment in Fig. 4.7.

#### Non-Uniform Temperature over Vias

A Gaussian pulse defined by Equation (4.21) with  $f_0 = 0.8$  V,  $t_0 = 80$  ns, and  $\tau = 40$  ns is used as the input. Even though the height of the via BD is only 3  $\mu$ m, a large temperature variance is observed. As shown in Fig. 4.8(a), the temperance difference over the via BD in segment 2 can be as large as 13.4 K, and that over the via BD in segment 7 is 10.7 K in maximum. The non-uniform temperatures over the vias are the potential causes of structural





(b)

Figure 4.8: Temperatures at two ends of (a) via BD and (b) via CE.

failures. As the via CE is placed further away from the injected Gaussian pulse where Joule heating is less severe, the temperature difference over the via CE is much smaller than that over the via BD. As shown in Fig. 4.8(b), the temperature difference over the via CE in both segments 2 and 11 is less than 2 K. Figure 4.9 captures the temperature profiles at the plane  $z = 52 \,\mu\text{m}$  at four instances, 55 ns, 77 ns, 99 ns, and 121 ns. The thermal hot spots are found at the via BD and part of M1. The temperature at the hot spots reaches its maximum around 100 ns.

#### Impact of the Pulse Width

Gaussian pulses of width  $\tau = 20$  ns and  $\tau = 40$  ns are launched into the on-chip power grid to investigate the corresponding impact on the transient thermal behaviors. The other parameters of the Gaussian pulses remain as  $f_0 = 1.0$  V and  $t_0 = 80$  ns. Temperatures at points D2 and D8 are recorded and depicted in Fig. 4.10(a). When the pulse width is reduced from 40 ns to 20 ns, less energy is injected to the system within the same period and the peak temperature is lowered down by 8 K as can be seen from Fig. 4.10(a). The voltage drops at points B, C, D, and E of segments 2, 8, and 12 under both types of pulse widths are recorded. The differences in voltage drops resulted from the variance in the pulse width are depicted in Fig. 4.10(b). It shows that the pulse width variance causes 6 mV difference in voltage drop at point B2 and around 7 mV at point D2.

#### Impact of the Pulse Rising Time

Two rectangular pulses of different rising time are injected into the system. One pulse rises at 10 ns and the other one rises at 20 ns, both of which reach 1.2 V at 40 ns. Both pulses fall at 60 ns, reach 0.2 V at 80 ns, and remain as 0.2 V throughout the rest of the time. The transient thermal behaviors at points D2 and D8 are depicted in Fig. 4.11(a). The rising time does not affect the temperature gradient because the slope of the rising temperature remains the same regardless of the rising time as shown in Fig. 4.11(a). However, the difference in the rising time does impact the peak temperature; for the pulse rising at 10 ns, more energy is injected to the system, which results in a higher temperature rise. The voltage drops associated with the temperature



Figure 4.9: Temperature profiles of plane  $z = 52 \ \mu m$  of the on-chip power grid at four instances (a) 55 ns, (b) 77 ns, (c) 99 ns, (d) 121 ns, and (e) colormap.





Figure 4.10: Impact of the pulse width on (a) temperature and (b) voltage drop.



(a)



Figure 4.11: Impact of the pulse rising time on (a) temperature and (b) voltage drop.



Figure 4.12: Solder bump array of 32 columns.

rise is shown in Fig. 4.11(b).

#### 4.4.2 Solder Bump Array

The second example is a solder bump array of 32 columns as shown in Fig. 4.12. In each column, the electrodes are immersed in solder bumps, which are connected through redistribution layers from both the top and the bottom. The dimensions of the solder bump array in each column are specified in Fig. 4.4: the radius (r\_bump) and the height (h\_bump) of the solder bump are 17  $\mu$ m and 12  $\mu$ m, respectively; the length (a\_electrode) and the thickness (t\_electrode) of the electrode are 10  $\mu$ m and 10  $\mu$ m, respectively; and the width (a\_trace), length (b\_trace), and thickness (t\_trace) of the trace are 30  $\mu$ m, 112  $\mu$ m, and 3  $\mu$ m, respectively. The width (a\_underfill) and the length (b\_underfill) of each segment of the underfill are 74  $\mu$ m and 156  $\mu$ m, respectively. In each column, power and ground are applied to the surfaces of the bottom RDL where the ground is attached to the middle one and power is attached to the ones on the sides. Heat sink and air convection cooling are applied to the two sides of the solder bump array. The ambient temperature is chosen as 293 K and the convection cooling coefficient is 10  $W/m^2K$ . Temperatures and voltages at three point, A, B, and C, as shown in Fig. 4.12 are recorded.



Figure 4.13: Temperature profile of (a) the top surface and (b) the cross section of column 20 in the solder bump array at t = 173 ns.

Thermal Hot Spots on the Top RDL

Figure 4.13 shows the temperature distribution of the solder bump array at 173 ns. As can be seen, the thermal hot spots occur at traces belonging to the top RDL and at the junctions between the top RDL and the bumps. As the electrical inputs are applied to the bottom RDL, the electric current crowds at the junctions between the solder bumps and the top RDL. The current crowding causes localized overheating. Even though points A and C are both on the top RDL, temperature at point C is higher than that at point A due to the presence of the additional trace near point C which sets a discontinuity along its the current path and causes the current crowding.

Impact of the Pulse Width and Magnitude

Gaussian pulses of different widths and magnitudes are applied to the solder bump array. Figure 4.14 shows the temperatures at point A24, B24, A30,



Figure 4.14: Temperature at point A and B in columns 24 and 30 of the solder bump array.

and B30 with the applied Gaussian pulse of peak magnitude of 1.0 V and 1.2 V, respectively. It is worth mentioning that A24 refers to point A of column 24 with column index increasing along the positive z-axis. It can be seen from Fig. 4.14 that even though point A and B share the same distance to both power and ground, temperature at point A is higher than that at point B since the extra trace near point B provides additional current paths and alleviates the current crowding and the corresponding Joule heating. It is also worth mentioning that instead of dropping quickly as the on-chip power grid in Fig. 4.8, the temperature on the top RDL stays flat above 300 K at the steady state. This is because the electrical input provides really high power as it passes through the solder bump array.

Temperatures at point A16 are depicted in Figure 4.15(a) under various types of pulses. By reducing the pulse width by 10 ns, a temperature decrease of more than 15 K is achieved. Reducing the pulse width by 10 ns or lowering the peak magnitude by 0.2 V have similar effects on the temperature decrease. Figure 4.15(b) shows the variances in the voltage drops with the pulse width reducing from 30 ns to 20 ns.





Figure 4.15: Impact of the pulse width and magnitude on (a) temperature at A16 and (b) voltage drops at A1, B1, A30, and B30 in the solder bump array.



Figure 4.16: TSV-based power distribution network.

#### 4.4.3 TSV-Based Power Distribution Network

Figure 4.16 shows a TSV-based PDN where the on-chip power grid are connected to a TSV array through landing pads. The on-chip power grid consists of two layers of metal, and the cross-sectional dimension of the metal is  $1 \times 1$  $\mu m^2$ . The vias connecting the two metal layers have the dimension of  $1 \times 1 \times 3$  $\mu m^3$ . The landing pads, immersed in the inter-metal layer, have the dimension of  $1 \times 1 \times 3 \ \mu m^3$ . The radius and the height of the TSV are 1.5  $\mu m$  and  $8 \ \mu m$ , respectively. The insulation layer surrounding the metal-filled via has a thickness of 0.5  $\mu$ m. The metal layers in the on-chip power grid, the landing pads, and the vias are all made of copper. Both power and ground are attached to the bottom surface of the TSVs. The ambient temperature is chosen as 293 K and the convection cooling coefficient is  $10 \text{ W/m}^2\text{K}$ . There is a total number of 32 columns of TSVs. Temperatures and voltages at five points as marked in Fig. 4.16 are recorded: point A is on the top metal layer of the on-chip power grid; point B is the on the via that connects the two metal layers; point C is on the bottom metal layer of the on-chip power grid; point D is on the landing pad; and point E is on the TSV. Also in Fig. 4.16, terminals on four vias are highlighted to differentiate the types of power maps. Three types of power maps are used in the example: PGPG, power on terminals 1 and 3 and ground on terminals 2 and 4; PPGG, power on terminals 1 and 2 and ground on terminals 3 and 4; PGGP, power on terminals 1 and 4 and ground on terminals 2 and 3.



Figure 4.17: Temperature at point C, D, and E in column 16.



Figure 4.18: Thermal hot spots on the landing pad.

Thermal Hot Spots on Landing Pad

With the power map of PGPG, a Gaussian pulse defined by Equation (4.21) with  $f_0 = 0.5$  V,  $t_0 = 80$  ns, and  $\tau = 30$  ns is sent into the structure. The temperatures at points C, D, and E of column 24 are depicted in Fig. 4.17. It can be seen that the thermal hot spot occurs on the landing pad. Because



Figure 4.19: Temperature at point D24 under various types of electrical input.

of the mismatch across the junction between the TSV and the landing pad, current crowding occurs and exacerbates the Joule heating effect, resulting in the localized overheating at the landing pad. Figure 4.18 captures the temperature profile on the cross-sectional area of column 12, which clearly indicates the thermal hot spots on the landing pads.

Impact of the Pulse Width and Magnitude

Gaussian pulses of width 30 ns and 20 ns and peak magnitude 0.5 V and 0.3 V are sent into the TSV-based PDN to examine the corresponding impact on the electrical-thermal behaviors. Because the landing pad has the most serious thermal issue, we pick point D24 and depict its transient thermal behaviors under the various types of pulses. It can be seen in Fig. 4.19 that by increasing the peak magnitude from 0.3 V to 0.5 V, a temperature increase of 60 K occurs. With the peak magnitude remaining the same, by reducing the pulse width from 30 ns to 20 ns, a temperature decrease of 15 K takes place.



(a)



Figure 4.20: Impact of the power maps on the temperature in column 3 at (a) points C and D and (b) points A and B.

#### Power Maps

With the input pulse and the number of terminals remaining the same, we adjust the power maps and examine their impact on the electrical-thermal behaviors. By changing the power map from PGPG to PGGP, a significant variance in temperature is observed at points C and D as shown in Fig. 4.20. There are three neighboring grounds for one single power in PGGP, whereas one power is paired up with one ground in PGPG. The three neighboring grounds in PGGP provides three possible current paths; the landing pad acts as a power hub where currents add up before branching out into individual paths. The huge amount of current added by three current paths exacerbates the Joule heating at the landing pad in PGGP, which explains the significant rise in temperature at point D. In contrast, the bottom metal layer, which used to be the only current path under PGPG, now becomes one of the three current paths under PGGP; consequently, the Joule heating at point C is alleviated, which explains the temperature drop in Fig. 4.20(a). By changing the power map from PGPG to PPGG, temperature rises at points A and B are observed in Fig. 4.20(b). From PGPG to PPGG, instead of flowing to the ground in the same column, the current detours and flows to the grounds in the neighboring columns. This alleviates the Joule heating on the bottom metal layer but exacerbates that on the top one. The voltage drop associated with different power maps are explained together with via pitches in the following.

#### TSV Pitch

Here, we study the voltage drops under different via pitches and power maps. To illustrate the findings, we pick two major components in this particular structure, namely, the landing pad and the TSV. Figure 4.21 shows the voltage drops over TSVs in columns 3 and 8 under different via pitches and power maps. The power map PGGP has three current paths, which not only leads to the most serious Joule heating but also the largest voltage drop along TSVs. With the same power map, as the TSV pitch is reduced, larger voltage drops over the TSVs are found. Similar phenomena are observed in the landing pads. Figure 4.22 shows the voltage drop over the landing pads in columns 7 and 16 under different via pitches and power maps.



Figure 4.21: Voltage drop under various power maps and TSV pitches in over (a) TSV in column 3 and (b) TSV in column 8.



Figure 4.22: Voltage drop under various power maps and TSV pitches in over (a) landing pad in column 7 and (b) landing pad in column 16.

## CHAPTER 5

# THERMAL-AWARE HIGH-FREQUENCY CHARACTERIZATION OF LARGE-SCALE THROUGH-SILICON-VIA STRUCTURES

### 5.1 Problem Statement

The through-silicon-via (TSV) interconnection technology enables a vertical stack-up of multiple planar devices. The short vertical interconnect ascribed to the TSV alleviates the interconnect delay problems; it also shrinks the chip area and smooths the path toward high packing density [26]. The TSV interconnection technology also finds its application in the 3-D system-on-chip (SOC), known as the heterogeneous integration, where disparate technologies of complex materials and processes are unified and accommodated into a single functional block on the silicon substrate [4], [5].

As the TSV becomes a key component to the 3D integrated circuit (IC) technology, a deep understanding of the electrical behaviors of the TSVs is essential in the low-cost and high-yield design. There have been attempts made to characterize the TSVs or the via-structures based on the lumped circuit models [50, 51, 52, 53, 54]. With the ever increasing clock frequency and continuous reduction in the feature size, the physically small TSV structures can be electrically large such that the full-wave electromagnetic analysis is preferred to attain certain accuracy [3], [55]. Attempts have also been made to extract the circuit models based on the measurement [56] or the full-wave electromagnetic analysis [57], many of which are devised for a single TSV or a few coupled ones. However, the density of the TSVs can be large in the real design; and additional loss, noise coupling, and signal distortions have been imposed to the massively coupled TSV structures by the lossy silicon substrate and the continuous reduction in the via pitch. In order to accurately predict the electrical behaviors, a large number of TSVs should be addressed simultaneously [58], [59]. This can be computationally intensive particularly when the accuracy is emphasized and the full-wave solvers are required. In

this chapter, we target on the characterization of large-scale TSV structures by introducing a highly efficient domain decomposition scheme into the fullwave analysis.

Thermal effect is another issue that should be considered in characterizing the electrical behaviors of the TSVs. High temperature increases the resistivity of metal interconnects and reduces the mobility of carriers in the silicon substrate. It has been demonstrated that the thermal issue is more severe in the 3D integration with vertical stack-ups than that over a single die [29], [60]. The adhesive layers in-between the vertically stacked-up dies are generally of poor thermal conductivity; and the 3D integration again increases the power density and degrades the temperature spreading. Therefore, thermal analysis cannot be neglected in nor separated from an accurate prediction of the electrical behaviors of the TSVs. In this chapter, the high-frequency characterization and the thermal analysis are coupled through an iterative manner to achieve electrical-thermal co-simulation.

The proposed electrical-thermal co-simulation in this chapter takes into account the thermal effects and at the same time is able to deal with large-scale TSV structures. We implement the co-simulation with the finite element method (FEM) [10] for its capability of modeling complex circuit geometries and materials [44], [45]. In order to deal with large-scale problems, we introduce a domain decomposition scheme called the finite element tearing and interconnecting (FETI) method [14], [35]. By utilizing the highly parallelizable nature of FETI [17] through parallel computing with multiple processors, we achieve a significant reduction in computation time and are able to analyze large-scale TSV structures in a highly efficient manner. Various designs parameters in typical TSV structures such as the TSV array in the silicon interposer and the TSV daisy chains are investigated with the proposed co-simulation.

## 5.2 Electrical-Thermal Co-Simulation

The electrical analysis of the TSV structures starts from solving the vector wave equation

$$\nabla \times \left(\frac{1}{\mu_r} \nabla \times \boldsymbol{E}\right) - k_0^2 \epsilon_r \boldsymbol{E} = -jk_0 Z_0 \boldsymbol{J}$$
(5.1)
where  $\epsilon_r$  and  $\mu_r$  are the relative permittivity and permeability, respectively;  $Z_0$  is the intrinsic impedance of free space; J is the density of the excitation current source; and  $k_0$  is the wavenumber of the field. Because the TSV structures are generally open or semi-open, absorbing boundary conditions are used to truncate the open domain into a finite computation domain. The first-order absorbing boundary condition adopted in this chapter is

$$\frac{1}{\mu_r}\hat{n} \times (\nabla \times \boldsymbol{E}) + jk_0 \sqrt{\frac{\epsilon_r}{\mu_r}}\hat{n} \times (\hat{n} \times \boldsymbol{E}) = 0$$
(5.2)

The electrical analysis also requires wave ports attached at the port surfaces to launch incident waves into the TSV structures. The wave port boundary condition on a port surface denoted as  $\Gamma$  can be expressed as

$$\hat{n} \times \nabla \times \boldsymbol{E} + P(\boldsymbol{E}) = \boldsymbol{U} \tag{5.3}$$

in which

$$P(\boldsymbol{E}) = \sum_{m=1}^{\infty} \frac{1}{j\omega\mu\kappa_m} (\gamma_m \vec{e}_{tm} + \nabla e_{zm})$$
(5.4)

$$\int_{\Gamma} \boldsymbol{E} \cdot (\gamma_m \vec{e}_{tm} + \nabla e_{zm}) \mathrm{d}\Gamma$$
(5.5)

$$\boldsymbol{U} = \hat{\boldsymbol{n}} \times \nabla \times \boldsymbol{E}^{\text{inc}} + P(\boldsymbol{E}^{\text{inc}})$$
(5.6)

where

$$\kappa_m = \frac{j}{\omega\mu} \int_{\Gamma} (\gamma_m \vec{e}_{tm} \cdot \vec{e}_{tm} + \vec{e}_{tm} \cdot \nabla_t e_{zm}) \mathrm{d}\Gamma$$
(5.7)

 $\vec{e}_{tm}$  and  $e_{zm}$  are the transversal and longitudinal components of the modal electric fields, respectively, and  $\gamma_m$  is the propagation constant.

The thermal analysis starts with solving the Poisson equation

$$\nabla \cdot k \nabla T = -Q \tag{5.8}$$

where k is the thermal conductivity, T is the temperature distribution, and Q is the volumetric heat generation. The isothermal boundary  $\Gamma_c$  is define as

$$T = T_c \tag{5.9}$$

and the convection boundary  $\Gamma_{\rm conv}$  is defined as

$$k\frac{\partial T}{\partial n} = -h(T - T_a) \tag{5.10}$$

where h and  $T_a$  are convective heat transfer coefficient and ambient temperature, respectively.

The electrical and the thermal analyses are coupled through the dissipated power and the temperature-dependant material properties as shown in Fig. 5.1: specifically, the electromagnetic fields are calculated in the electrical analysis, through which one obtains the power dissipation in the TSV structures; the dissipated power is regarded as the source of Joule heating, with which the temperature profile is updated; the updated temperature profile modifies the material properties prior to another round of electrical analysis; and this scheme repeats itself until the material properties converge with a preset tolerance. A pseudocode is provided in Fig. 5.2 to illustrate the detailed implementation of the co-simulation. It is worth mentioning that the dissipated power consists of both the conduction loss  $P_c$  denoted as

$$P_c = \int_v \sigma |E|^2 \mathrm{d}V \tag{5.11}$$

and the dielectric loss  $P_d$  denoted as

$$P_d = \omega \epsilon_0 \int_v \epsilon_r'' |E|^2 \mathrm{d}V \tag{5.12}$$

where  $\epsilon_r''$  is the imaginary part of the relative permittivity  $\epsilon_r$ . The temperaturedependent resistivity of metal interconnects can be written as

$$\rho = \rho_0 \left[ 1 + \alpha (T - T_0) \right] \tag{5.13}$$

where  $\rho_0$  is the resistivity at temperature  $T_0$  and  $\alpha$  is the temperature coefficient of the metal interconnects. A substrate of p-type bulk silicon is considered in this chapter. The hole mobility at room temperature, denoted as  $\mu_{p_0}$ , is a function of the dopant impurity concentration and is given by



Figure 5.1: Coupling between the electromagnetic fields and temperature profile in the electrical-thermal co-simulation.

[61] and [62] in the form of

$$\mu_{p_0} = \frac{\mu_{\max} - \mu_{\min}}{1 + \left(\frac{N_a}{N_{\text{ref}}}\right)^{\beta}} + \mu_{\min}$$
(5.14)

where  $N_{\rm a}$  is the concentration of dopant impurity and the rest are constants:  $\mu_{\rm max} = 495 \text{ cm}^2/\text{Vs}, \ \mu_{\rm min} = 47.7 \text{ cm}^2/\text{Vs}, \ \beta = 0.76, \text{ and } N_{\rm ref} = 6.3 \times 10^{16} \text{ cm}^{-3}$ . The carrier mobility is temperature dependent and is defined by [62] and [63]

$$\mu_p(T) = \mu_{p_0} \left(\frac{T}{300}\right)^{-\frac{3}{2}} \tag{5.15}$$

Before extending capability of the solver to deal with complex and largescale structures, let us use one example to verify the implementation of the solver itself. A rectangular box in Fig. 5.3 is taken as a benchmark example to verify the implementation of the co-simulation. Simulation results are compared with those obtained from COMSOL Multiphysics [49]. The rectangular box has a dimension of  $2 \times 0.5 \times 6$  cm<sup>3</sup>. Perfect electrical conductors are assumed to enclose the box except for one surface on which a wave port is attached. Convection cooling boundary is attached on the top surface of the box. The convection coefficient is chosen as 100 W/m<sup>2</sup>K and the ambient temperature is 300 K. The box is filled with p-type bulk silicon of relative permittivity 12.1 and conductivity 0.1 S/m. The input power of INPUT Initial resistivity  $\rho_0$ , tolerance tol, number of iterations N

OUTPUT Electromagnetic field components and temperature profile

Step 1 Set i = 1 and  $\tilde{\rho} = \rho_0$ 

Step 2 While  $i \leq$  number of iterations do Steps 3-8

- Step 3 Compute electromagnetic field components via electrical analysis
- Step 4 Calculate power
- Step 5 Compute temperature via thermal analysis
- Step 6 Update resistivity  $\rho$
- Step 7 If  $|\rho \tilde{\rho}| < tol$

Output and Exit

- Step 8 Set i = i + 1 and  $\tilde{\rho} = \rho$
- Step 9 Output ("Not converged after N iterations")

Figure 5.2: The iterative scheme of the electrical-thermal co-simulation.



Figure 5.3: Rectangular box with a wave port and a convection boundary.

1 W is used and the frequency of the excitation varies from 8 to 14 GHz. The dissipated power and the maximum value of the temperature measured in the box are depicted in Fig. 5.4. As we can see, the dissipated power and the maximum value of the temperature match well with those obtained with COMSOL Multiphysics. As a further verification, the trend of the maximum value of the temperature varying with the efficiency of the convection cooling is depicted in Fig. 5.5, where the input power of 1 W and 2 W are applied accordingly. The simulation results again match very well with those obtained with COMSOL Multiphysics.



Figure 5.4: Power dissipation and temperature rise in the rectangular box.



Figure 5.5: The maximum value of the temperature varies with the efficiency of the convection cooling in the rectangular box.

## 5.3 Computation Information

The parallel implementation of FETI is realized through the Message Passing Interface (MPI). The parallel efficiency of FETI with multiple processors is



Figure 5.6: TSV array in the silicon interposer: (a) 3D view of the TSV array; (b) dimensions of a single TSV.

demonstrated with the example of a silicon interposer shown in Fig. 5.6. All computations are performed with Intel Xeon 2.67-GHz hex-core processors. Consider the TSV array in the silicon interposer shown in Fig. 5.6: the radius of the via (rvia) is 10  $\mu$ m; the outer radius of the oxidation layer (roxc) is 11  $\mu$ m; the thickness of the metal layer (hm) is 2  $\mu$ m; the thickness of the oxidation layer (hoxc) is 2  $\mu$ m; the via pitch (p), measured between the centers of the vias, is 50  $\mu$ m along both the *x*-axis and the *y*-axis; the TSV is copper-filled with electrical conductivity of 5.8 × 10<sup>7</sup> S/m at room temperature, temperature coefficient of 0.0039, and thermal conductivity of 400 W/mK; the silicon substrate has the relative permittivity of 12.1, the electrical conductivity of 5 × 10<sup>3</sup> S/m at room temperature, and the thermal conductivity of 163 W/mK; and the oxidation layer has the relative permittivity of 4 and the thermal conductivity of 1.38 W/mK.





Figure 5.7: Electrical-thermal co-simulation of a  $3 \times 2$  TSV array by FEM and FETI: (a) scattering parameters and maximum value of temperature and (b) temperature profile on the top surface of the TSV array.

Before the investigation of the parallel efficiency of FETI with MPI, let us use one example of a  $3 \times 2$  TSV array, with the former indicating the number of vias along the x-axis and the latter along the y-axis, to benchmark the implementation of FETI. Two wave ports are defined with the same TSV, attached onto its top and bottom surface, respectively. The heat sink at room temperature is applied to the lower sidewall of the TSV array. Figure 5.7(a) depicts the scattering parameters and the maximum value of the temperature measured in the TSV array. It can be seen that the results from FEM and FETI match well with each other. Figure 5.7(b) portrays the temperature profile of the top surface in the TSV array where the wave port excitation at 3 GHz is placed on one end of the TSV.

Total number of DOFs	$3.3  imes 10^5$	$6.5  imes 10^5$	$1.3  imes 10^6$	$2.6  imes 10^6$
Number of processors	4	8	16	32
Number of subdomains	4	8	16	32
Factorization	26.1	27.1	26.3	27.1
Solvering interface problem	12.5	12.8	14.4	15.3
Extracting fields	0.89	0.90	0.90	0.91
Total computation time (seconds)	49.5	51.0	51.9	53.5

Table 5.1: Unit-load-per-processor parallel efficiency of FETI with MPI

Table 5.2: Speed-up of FETI with MPI

Total Number of DOFs: $1.31 \times 10^6$							
Number of processors	4	8	16	32			
Number of subdomains	4	8	16	32			
Number of DOFs per subdomain	$3.3  imes 10^5$	$1.6\times 10^5$	$8.2\times10^4$	$4.1 \times 10^4$			
Total computation time (seconds)	354.7	87.6	51.9	26.2			

Now, let us investigate the parallel efficiency based on the idea of unit load per processor, under which the number of DOFs assigned to individual processors is kept the same. Four cases are generated for comparison, namely, the TSV arrays of  $8 \times 4$ ,  $8 \times 8$ ,  $8 \times 16$ , and  $8 \times 32$ . The computation time of each individual step of FETI is recorded in Table 5.1 for the electromagnetic analysis during one iteration of the co-simulation. It is observed that as the total number of DOFs jumps from  $3.3 \times 10^5$  to  $2.6 \times 10^6$ , the total computation time holds almost the same when the number of processors is increased from 4 to 32. The slight increase in the computation time occurs during the step of interconnecting due to the necessary communication among the processors. It can be drawn from Table 5.1 that the FETI-enabled parallel implementation with MPI is capable of solving a large-scale problem with a similar amount of time to that of a much smaller problem by simply employing proportionally more processors.

As a further investigation of the parallel efficiency we explore the trend of the reduction in computation time with an increase in the number of processors. Four cases are generated for comparison, all of which deals with a TSV array of dimension  $8 \times 16$ . In the four cases, the entire structure is broken into 4, 8, 16, and 32 subdomains, respectively. Each subdomain is paired up with one processor. The computation time of all these four cases is recorded in Table 5.2. Again, the computation time is recorded for the electromagnetic analysis during one iteration of the co-simulation. The speed-up is defined as the ratio of the computation time of two different cases. It is observed that when the number of processors is doubled from 4 to 8, a speed-up of 4.0 is achieved. The reason lies in the factorization of the subdomain matrices; in the case of four subdomains, the time taken to factorize the matrix is about four times of that for the case of eight subdomains. It is also observed that when the number of processors is doubled from 8 to 16, the speed-up is 1.7, which is less than 2. This is because with the increased number of subdomains, the size of the interface problem and the time consumed in communication over different processors increase accordingly. Based on this investigation, it is seen that the optimal parallel performance is obtained when the number of DOFs per subdomain is in the range of  $2 \times 10^4$  to  $2 \times 10^{5}$ .

## 5.4 Design Parameters Investigation

With the proposed full-wave solver of which the efficiency is enhanced by the parallel implementation of FETI with MPI, we are now able to investigate large-scale TSV structures. We characterize the electrical behaviors of typical TSV structures including TSV arrays in the silicon interposer and TSV daisy chains by calculating the insertion loss, return loss, and crosstalks. We also

**Heat Sink** 



**Convection Cooling** 

Figure 5.8: Top view of the silicon interposer: port 2 is defined on the bottom and opposite to port 1; convection cooling and heat sink are applied on the two sidewalls.

study the influences on the signal transmission and couplings in the TSV structures exerted by the cooling temperature, the via-filling material, the doping concentration of silicon substrates, and the via pitch.

## 5.4.1 TSV Array in the Silicon Interposer

Figure 5.8 shows the top view of an  $8 \times 8$  TSV array in the silicon interposer, where a single TSV in the interposer is portrayed by Fig. 5.6(b). The geometrical information and the material properties of the TSV array are as follows: the radius of the via (rvia) is 10  $\mu$ m; the outer radius of the oxidation layer (roxc) is 11  $\mu$ m; the thickness of the metal layer (hm) is 2  $\mu$ m; the thickness of the oxidation layer (hoxc) is 2  $\mu$ m; the via pitch (p), measured between the centers of the vias, is 100  $\mu$ m along both the *x*-axis and the *y*-axis; the TSV is tungsten-filled with electrical conductivity of  $1.79 \times 10^7$  S/m at room temperature, temperature coefficient of 0.0045, and thermal conductivity of 173 W/mK; the silicon substrate has the relative permittivity of 12.1, the electrical conductivity of  $5 \times 10^3$  S/m at room temperature, and the thermal conductivity of 163 W/mK; and the oxidation layer has the relative permittivity of 4 and the thermal conductivity of 1.38 W/mK. Four ports are defined in Fig. 5.8 with Port 2 sitting on the bottom and opposite to Port 1. Air convection cooling and heat sink are attached onto the two sidewalls of the silicon interposer as shown in Fig. 5.8. The convection coefficient is chosen to be  $10 \text{ W/m}^2\text{K}$ .

#### Insertion Loss and Return Loss

We first assume that the ambient temperature is at 303 K and explore the thermal impact on the insertion loss and return loss by exciting Port 1 in Fig. 5.8. The insertion and return losses are characterized by  $S_{21}$  and  $S_{11}$ , respectively. Both the via-filling material (tungsten for this case) and the silicon substrate have the temperature-dependent material properties. Figure 5.9 depicts  $S_{21}$  and  $S_{11}$  over the frequency range from 16 to 34 GHz. The insertion loss of this particular structure, characterized by  $S_{21}$ , is more vulnerable to the thermal effects as a maximum deviation of 23.6% is observed. The thermal effects on the return loss, characterized by  $S_{11}$ , has the maximum deviation of 2.3% occurring around 34 GHz. It is also worth mentioning that when the thermal effects are taken into account, lower  $|S_{21}|$ and  $|S_{11}|$  are expected where the reduced portion of energy is converted into the Joule heat, which explains the conservation of energy. The deviations of the insertion and return losses due to the power dissipation again stress the importance of considering the thermal effects in predicting the electrical behaviors of the TSV structures.

#### Crosstalk

The near-end crosstalk is characterized by  $S_{31}$  and  $S_{41}$  where Port 1, Port 3, and Port 4 are defined in Fig. 5.8. The crosstalk in terms of  $S_{31}$  and  $S_{41}$  are portrayed in Fig. 5.10(a) and 5.10(b), respectively. As the cooling temperature rises from room temperature at 293 K to 333 K and then to 353 K, the near-end crosstalk becomes stronger. In particular, when the cooling temperature reaches 353 K, the near-end crosstalk is increased by more than 1 dB over the frequency band from 16 to 34 GHz.



(a)



Figure 5.9: Thermal impact on (a)  $S_{21}$  and (b)  $S_{11}$  of the tungsten-filled TSV array in the silicon interposer.



Figure 5.10: Thermal impact on the near-end crosstalk, characterized by (a)  $S_{31}$  and (b)  $S_{41}$ , of the tungsten-filled TSV array in the silicon interposer.



Figure 5.11: Near-end crosstalk with reduction in the via pitch of the tungsten-filled TSV array in the silicon interposer.

#### Via Pitch

It is generally desired to reduce the via pitch in order to improve the yield since a smaller via pitch leads to a higher integration density. However, the smaller via pitch also risks higher crosstalk and more severe signal distortions, especially in the massively packed TSV structures. The power density is likely to increase by reducing the via pitch, which exposes the TSV structures to more serious Joule heating effects. Figure 5.11 illustrates the variation of the near-end crosstalk with the reduction in the via pitch from 100 to 50  $\mu$ m; at 34 GHz,  $S_{31}$  and  $S_{41}$  are raised by 3.6 dB and 5.1 dB, respectively, which are obtained with the cooling temperature at 303 K.

#### Material Properties

Figure 5.12 illustrates the impact of the material properties on the insertion and return losses of the TSV array in the silicon interposer, characterized again by  $S_{21}$  and  $S_{11}$ , respectively. The changes made in the material properties include replacing the tungsten by copper as the via-filling material and decreasing the electrical conductivity of the silicon substrate by modifying its



(a)



Figure 5.12: Impact of the via-filling materials and the doping concentration of the silicon substrate on (a)  $S_{21}$  and (b)  $S_{11}$ .

doping concentration. The copper-filled TSV has the electrical conductivity of  $5.8 \times 10^7$  S/m at room temperature, temperature coefficient of 0.0039, and thermal conductivity of 400 W/mK. The electrical conductivity of the silicon substrate changes from  $5 \times 10^3$  S/m to 50 S/m. It can be seen from Fig. 5.12 that the copper-filled via has lower  $S_{21}$  and higher  $S_{11}$  than the tungstenfilled one in this example. For the tungsten-filled via, higher conductivity of the silicon substrate results in higher  $S_{21}$ .

The spatial discretization of the TSV array in the silicon interposer results in a total number of 1.43 million unknowns, which is divided into eight subdomains with each paired up with one processor. With eight processors computing in parallel, it takes 159.8 seconds for the electromagnetic full-wave analysis during one iteration in the co-simulation.

## 5.4.2 TSV Daisy Chain

Figure 5.13 shows a section of the TSV daisy chain [64], [65]. Two layers of TSV in separate dies are connected through metal lines and bumps in the redistribution layer (RDL). Instead of considering only the vertical vias and extracting the circuit parameters with quasi-static field solvers, we analyze the electromagnetic wave propagating inside these structures including the layers of vertical vias and the horizontal redistribution layers.

Consider the TSV daisy chain illustrated in Fig. 5.13: the diameter of the via (d\_TSV) is 30  $\mu$ m; the diameter of the oxidation layer (d\_OX) is 35  $\mu$ m; the thickness of the metal line in the redistribution layers (t\_RDL) is 5  $\mu$ m; the width of the metal line (w\_RDL) is 100  $\mu$ m for all the three redistribution layers; the length of the metal line in the top RDL (l1\_RDL), the middle RDL (l2\_RDL), and the bottom RDL (l2\_RDL) are 240  $\mu$ m, 140  $\mu$ m, and 70  $\mu$ m, respectively; bumps in the RDL are represented by frustums with the top diameter (d\_Bump1) of 30  $\mu$ m and the bottom diameter (d\_Bump2) of 60  $\mu$ m; the via pitch measured between the cental lines of the neighboring channels is 100  $\mu$ m; the height of the silicon substrates of the top die (h1\_Si) and the bottom die (h2\_Si) are both 100  $\mu$ m; and the bottom one is 130  $\mu$ m.

The material properties are given in the following: the TSV is tungstenfilled with electrical conductivity of  $1.79 \times 10^7$  S/m at room temperature,



Figure 5.13: Two layers of TSVs are connected through metal lines and bumps in the redistribution layer: (a) 3D view of a partial section; (b) side-view of one segment.



Figure 5.14: Top view of 16 coupled TSV daisy chain channels with definitions of five ports; Port 1 and Port 2 are defined on the eighth channel starting from the left.

temperature coefficient of 0.0045, and thermal conductivity of 173 W/mK; the silicon substrate has the relative permittivity of 11.9, the electrical conductivity of 1 S/m at room temperature, and the thermal conductivity of 163 W/mK; the oxidation layer has the relative permittivity of 4 and the thermal conductivity of 1.38 W/mK; and the redistribution layer has the relative permittivity of 2.6 and the thermal conductivity of 2.0 W/mK. A heat sink is attached on the bottom as shown in Fig. 5.13(b).

Figure 5.14 shows 16 coupled TSV daisy chain channels where five ports are defined. Investigations on the signal transmission, near-end and far-end crosstalks, and the impact of the material properties are performed with this particular TSV structure. Note that all the characterizations are carried out with the electrical-thermal co-simulation.

#### Insertion Loss

Port 1 is used as the excitation and  $S_{21}$  is used to characterize the insertion loss. As shown in Fig. 5.15(a), the higher ambient temperature leads to higher  $S_{21}$ . At 2 GHz and the ambient temperature of 353 K, the thermal effects on the insertion loss is around 26%. By replacing the via-filling material of tungsten with copper and changing the electrical conductivity of the silicon substrate from 1 S/m to 10 S/m, we further investigate the impact of material properties on the insertion loss. As shown in Fig. 5.15(b), for the copper-filled TSVs, the silicon substrate of lower conductivity has higher  $S_{21}$ . Further, for the same silicon substrate, when the operation frequency is higher than 1.5 GHz, the tungsten-filled TSVs have higher  $S_{21}$  than the copper-filled ones.

#### Near- and Far-End Crosstalks

Port 1 is taken as the excitation, and Port 3, Port 4, and Port 5 are used to measure the crosstalks where  $S_{31}$  characterizes the near-end crosstalk and  $S_{41}$  and  $S_{51}$  characterize the far-end crosstalk. Figure 5.16 depicts the nearand far-end crosstalks, which shows that crosstalks increase as the frequency goes higher. The near- and far-end crosstalks measured at Port 3 and Port 4 have a critical point around 1.2 GHz, lower than which the far-end crosstalk dominates. All these are calculated with a heat sink of 353 K attached on the bottom. Replacing the via-filling material of tungsten with copper and changing the electrical conductivity of the silicon substrate from 1 S/m to 10 S/m, we further investigate the impact of material properties on the nearand far-end crosstalks. With the doping concentration of the silicon substrate remaining the same, the tungsten-filled TSVs have both higher near- and farend crosstalks than the copper-filled ones when the operation frequency is larger than 4 GHz as shown in Fig. 5.17; for the copper-filled TSVs, higher electrical conductivity of the silicon substrate produces a higher near-end crosstalk over the frequency band studied in this example.

The spatial discretization of the daisy chain structure results in a total number of 1.1 million unknowns, which is divided into 16 subdomains with each paired up with one processor. With 16 processors computing in parallel, it takes 73 seconds for the electromagnetic full-wave analysis during one iteration of the co-simulation. The spatial discretization of the daisy chain structure of 32 coupled channels with a total number of 2.1 million unknowns is also calculated with 32 processors. It takes 63.3 seconds for the electromagnetic full-wave analysis during one iteration of the co-simulation.



Figure 5.15: Insertion loss, characterized by  $S_{21}$ , in the coupled TSV daisy chain structure with (a) various cooling temperatures and (b) varying material properties.



Figure 5.16: Near- and far-end crosstalks in the coupled TSV daisy chain structure.



Figure 5.17: Impact of the material properties on the (a) near-end crosstalk and (b) far-end crosstalk.

## CHAPTER 6

## ELECTRICAL-THERMAL CO-SIMULATION FOR ANALYSIS OF HIGH-POWER RF/MICROWAVE CIRCUITS

## 6.1 Problem Statement

Wireless infrastructure applications are driving RF/microwave circuit design into high frequencies and high power levels [66]. Continuous reduction in feature sizes and increase in complexities, together with the demand on higher frequencies and higher power levels, set a grand challenge for the design of RF/microwave circuits: on one hand, the distribution effect has to be considered at high frequencies; on the other hand, the thermal effect can no longer be neglected due to the reduced feature sizes and the increased power levels. Consequently, a multiphysics-based computer-aided design methodology is in demand to address the electrical and thermal issues simultaneously. In the past, the coupled electromagnetic and thermal analyses have been studied based on the finite difference time domain method (FDTD) for microwave heating in domestic ovens and microwave furnaces [67, 68, 69]. The electrical-thermal co-simulations based on the finite-element method (FEM) have been developed for DC IR drop analysis [43] and design of power distribution networks (PDNs) [70]. The coupled electromagnetic-thermal simulation based on the FDTD requires the electromagnetic steady state, which can be achieved through an adequate number of iterations. The convergence to electromagnetic steady state in FDTD depends on the Q-factor of the modeled structure [67, 68]. A lossy medium in general reduces the Q-factor and speeds up the convergence. In this chapter, we introduce an electrical-thermal co-simulation that integrates a full-wave steady-state electromagnetic analysis and a transient thermal analysis through an iterative scheme for the design of high-power RF/microwave circuits. The full-wave electromagnetic analysis is carried out in the frequency domain where the modeling of lossy and dispersive media in RF/microwave circuits and the achievement of an electromagnetic steady state come naturally. The FEM is employed for the co-simulation because of its unmatched capabilities in modeling complex geometries and materials [10]. In addition, the accuracy of the FEM solution can be improved systematically by increasing the order of its basis functions.

Although the FEM has many unique advantages, its volumetric discretization for many real and practical RF/microwave circuit structures can easily result in a large linear system of millions or even billions of unknowns. Solving such a large system can be very computationally intensive, and the solution has to be performed repeatedly for many iterations in the co-simulation. In this chapter, we employ two techniques in the co-simulation, namely, an adaptive time-stepping scheme in the transient thermal analysis and a domain decomposition scheme in both electromagnetic and thermal analyses, in order to alleviate the computational burden for solving large-scale problems. The adaptive time-stepping in this chapter is based on the algorithm of proportional-integral-derivative (PID) control [22, 23]. It automatically determines a set of non-uniform time steps for the transient analysis. The adaptive time-stepping can also detect when the transient temperature profile reaches a steady state and automatically terminates the co-simulation. The domain decomposition employed in this chapter is called the finite element tearing and interconnecting (FETI) [10, 14]. In the co-simulation, the FETI enhances the efficiency for solving the large FEM system associated with individual time steps. By applying the FETI, the entire computational domain is divided into smaller subdomains, which are completely decoupled initially and coupled only in the solution of the global interface system. Therefore, the FETI is highly suitable for parallel computing with multiple processors.

The capability and the efficiency of the co-simulation are demonstrated through analyzing matching networks in high-power RF amplifiers and substrate integrated waveguide (SIW) filters. In the high-power RF amplifiers, transistors are connected to the package leads through a large-scale matching network [71, 72]. The matching network is essentially an LC circuit with the inductance provided by bonding wire arrays and capacitance from metaloxide semiconductor (MOS) capacitors. The in-package matching network, whose characteristics are highly subject to the variances of temperature and fabrication [73], provides the necessary impedance transformation between the transistors and external circuits. A slight increase of temperature in the bonding wire arrays or the MOS capacitors may cause frequency shifts and alter the characteristics of the matching network. As a basic circuit component, RF filter is designed to suppress the undesired harmonics produced by amplifiers. Because of the low cost and the capability for high-density integration, the substrate integrated waveguide (SIW) [74] becomes a promising technique in the RF filter design. The high-power capability of the SIW interconnects has been demonstrated in [75]. However, the modern wireless infrastructures operate in an extremely crowded spectral environment. The limited bandwidth brings stringent design requirements to the RF filters such that the temperature stability becomes a critical concern. The proposed cosimulation is able to address and analyze the aforementioned issues in a very efficient manner.

## 6.2 Electrical-Thermal Co-Simulation

In this section, we present the details of the formulation and the implementation of the electrical-thermal co-simulation, including the adaptive timestepping and domain decomposition employed in the co-simulation.

## 6.2.1 Formulation with the Finite Element Method

The full-wave electromagnetic analysis solves the vector wave equation

$$\nabla \times \left(\frac{1}{\mu_{\rm r}} \nabla \times \vec{E}\right) - k_0^2 \epsilon_{\rm rc} \vec{E} = -jk_0 Z_0 \vec{J}$$
(6.1)

where  $\vec{E}$  represents the vector electric field;  $\mu_{\rm r}$  is the relative permeability;  $\epsilon_{\rm rc} = (\epsilon_{\rm r} - j\sigma/\omega\epsilon_0)$  is the relative complex permittivity, which includes the effects of electrical conductivity  $\sigma$  and angular frequency  $\omega$ ;  $Z_0$  is the intrinsic impedance of free space;  $\vec{J}$  represents the excitation current; and  $k_0$  is the wavenumber of the field. The first-order absorbing boundary condition

$$\frac{1}{\mu_{\rm r}}\hat{n} \times (\nabla \times \vec{E}) + jk_0 \sqrt{\frac{\epsilon_{\rm rc}}{\mu_{\rm r}}}\hat{n} \times (\hat{n} \times \vec{E}) = 0$$
(6.2)

is used to truncate the open region into a finite computational domain, where  $\hat{n}$  denotes a unit normal vector at the surface. Excitations including both

the wave ports and lumped ports are used in this chapter. The wave port can be expressed as

$$\hat{n} \times (\nabla \times \vec{E}) + P(\vec{E}) = \vec{U}$$
(6.3)

in which

$$P(\vec{E}) = \sum_{m=1}^{\infty} \frac{1}{j\omega\mu\kappa_m} (\gamma_m \vec{e}_{tm} + \nabla e_{zm})$$
(6.4)

$$\cdot \int_{S} \vec{E} \cdot (\gamma_m \vec{e}_{tm} + \nabla e_{zm}) \, \mathrm{d}S \tag{6.5}$$

$$\vec{U} = \hat{n} \times (\nabla \times \vec{E}^{\text{inc}}) + P(\vec{E}^{\text{inc}})$$
(6.6)

where

$$\kappa_m = \frac{\mathbf{j}}{\omega\mu} \int_S (\gamma_m \vec{e}_{\mathrm{tm}} \cdot \vec{e}_{\mathrm{tm}} + \vec{e}_{\mathrm{tm}} \cdot \nabla_{\mathrm{t}} e_{zm}) \,\mathrm{d}S \tag{6.7}$$

 $\vec{e}_{tm}$  and  $e_{zm}$  are the transversal and longitudinal components of the modal electric fields, respectively,  $\gamma_m$  is the propagation constant, and S denotes the port surfaces. The lumped port is considered as a constant line current running over the edges of finite elements. With the finite element discretization, the sub-vector associated with edge *i* contributed by the right-hand side excitation of Equation (7.1) can be written as

$$\{b\}_i = -jk_0 Z_0 I \int_{\text{edge } i} \vec{N}_i \cdot d\vec{l}$$
(6.8)

where I is the magnitude of the line current,  $d\vec{l}$  represents the parametrization of edge i, and  $\vec{N}_i$  denotes the vector basis functions.

The governing equation for the transient thermal analysis is

$$\rho c \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + Q \tag{6.9}$$

where T represents temperature,  $\rho$  is the density of the materials, c is the specific heat,  $\kappa$  is the thermal conductivity, and Q is the heat source. The corresponding boundary conditions in the transient thermal analysis includes the Dirichlet boundary

$$T = T_{\rm d},\tag{6.10}$$

where  $T_{\rm d}$  represents a constant temperature and the convection boundary

$$\hat{n} \cdot (\kappa \nabla T) = -h(T - T_{\rm a}) \tag{6.11}$$

where h and  $T_{\rm a}$  are the convective heat transfer coefficient and the ambient temperature, respectively.



Figure 6.1: The flow chart of the co-simulation algorithm:  $\Delta t$  represents the time step and  $\Delta t_{\text{max}}$  represents the predefined maximum time step.

The dissipated power calculated from the full-wave analysis contributes to the heat source in the thermal analysis, which consists of both the conduction loss

$$P_{\rm c} = \sigma |\vec{E}|^2 \tag{6.12}$$

and the dielectric loss

$$P_{\rm d} = \omega \epsilon_0 \epsilon_{\rm r}'' |\vec{E}|^2 \tag{6.13}$$

where  $\epsilon_{\rm r}''$  is the imaginary part of the relative permittivity. The procedure of the co-simulation algorithm is illustrated by Fig. 6.1. The full-wave and the

	Electrical conductivity $(0, -1)$	Temperature coefficient $(\mathbf{z}_{-1})$	Thermal conductivity $(W_{1}, -1) = 1$	$\begin{array}{c} \text{Density} \\ (\text{kg m}^{-3}) \end{array}$	Specific heat $(11 - 1)$
	$(S m^{-1})$	(K <sup>1</sup> )	$(W m^{-1}K^{-1})$		$(J kg^{-1}K^{-1})$
Copper	$5.8  imes 10^7$	$3.9  imes 10^{-3}$	400	$8.7\times10^3$	385
Gold	$4.1 \times 10^7$	$3.4 \times 10^{-3}$	315	$18.9  imes 10^3$	126

Table 6.1: Material properties of metal

	Relative	Thermal	Density	Specific
	permittivity	conductivity	$({\rm kg} {\rm m}^{-3})$	heat
		$(W m^{-1}K^{-1})$		$(\mathrm{J~kg^{-1}K^{-1}})$
Si	11.9	Eqn. (8.32)	$2.3  imes 10^3$	703
Silicon Dioxide	4.0	Eqn. (8.33)	$2.2 \times 10^3$	730
Alumina	9.4	Eqn. (8.34)	$3.9  imes 10^3$	900

Table 6.2: Material properties of dielectric

thermal analyses are coupled through the temperature-dependent material properties. The material properties involved in the chapter are listed in Table 6.1 for metal interconnects and Table 6.2 for dielectrics. The temperaturedependent material properties include electrical conductivity and permittivity in the electromagnetic analysis and thermal conductivities in the transient thermal analysis. The temperature-independent materials properties in Tables 6.1 and 6.2 are adopted from those in the material library of COMSOL Multiphysics [49]. The temperature-dependent resistivity or permittivity can be written as

$$\nu = \nu_0 \left[ 1 + \alpha (T - T_0) \right] \tag{6.14}$$

where  $\nu_0$  is the resistivity or permittivity at temperature  $T_0$  and  $\alpha$  is the temperature coefficient. Thermal conductivities of metal interconnects are assumed to be temperature-independent due to the very limited variance in the range of operating temperature. For example, the measured thermal conductivity of copper remains almost constant at 400 Wm<sup>-1</sup>K<sup>-1</sup> from 273 to 800 K [76]. In general, the thermal conductivity of silicon varies with dimensions as well as the manufacturing process [48]. For example, the thermal conductivity of thin silicon layers can be very different from that of bulk silicon substrates. In this chapter, the temperature-dependent thermal conductivity of the silicon substrate takes the empirical form provided in [48]

$$\kappa_{\rm Si} = 2.475 \times 10^5 \ T^{-1.3} \ (273 \rm K \le T \le 1000 \rm K)$$
(6.15)

which is based on the measurements from Volume 1 of [76]. The temperaturedependent thermal conductivity of silicon dioxide is found in [47] as

$$\kappa_{\rm SiO_2} = 1.02278 + 0.00121 \ T \ (273 \text{K} \le T \le 600 \text{K})$$
 (6.16)

of which the corresponding measurements can be found in Volume 2 of [76]. The temperature-dependent thermal conductivity of the alumina can be found in [77] as

$$\kappa_{\text{Al}_2\text{O}_3} = 5.5 + 34.5 \text{ e}^{-0.0033(T - T_0)}$$
(6.17)
(273K  $\leq T \leq 1000$ K)

The measured samples corresponding to Equation (6.17) belong to highalumina ceramics (at least 99.5%  $Al_2O_3$ ). Due to the nonlinearity brought by the temperature-dependent material properties, fixed-point iteration is employed at each time step of the transient analysis.

## 6.3 Model Verification and Demonstration

In this section, numerical examples are provide to verify the proposed cosimulation. The efficiency enhancement resulting from the adaptive timestepping and the FETI-enabled parallel computing is also demonstrated.

## 6.3.1 Model Verification: Three-Pole Chebyshev Filter

A three-pole Chebyshev filter [74] as shown in Fig. 6.2(a) has been designed and fabricated based on the SIW technique, which allows planar circuits and rectangular waveguides to be built on the same substrate. The dielectric substrate in the design has a relative permittivity of 2.2 (RT/Duroid 5880). The geometrical parameters labeled in Fig. 6.2(a) are summarized in Table 6.3. The Chebyshev filter is designed for the bandwidth of 1 GHz centered at 28 GHz. As shown in Fig. 6.2(b), the S-parameters from the full-wave analysis in the co-simulation match well with the measurement in [74].



Figure 6.2: Three-pole Chebyshev filter with two transitions of microstrip: (a) 3-D view and 2-D topology, and (b) measured and simulated S-parameters.

a	b	с	d	W
5.563	1.93	1.499	0.775	0.762
h	р	s1	s2	s3
0.787	1.525	4.71	5.11	1.01

Table 6.3: Dimensions (mm) of the SIW Chebyshev filter in Fig. 6.2(a)

# 6.3.2 Model Verification: Wire Bonding in the CPW Transition

Wire bonding is the most widely used interconnects in packaging owing to its convenience and low cost [78]. Figure 6.3(a) shows the wire bondings at the transition of two adjacent conductor-backed coplanar waveguides (CB-CPW) [79], of which the detailed geometrical information is summarized in Table 6.4. The gold bonding wire has a radius of 15.875  $\mu$ m and a height of 120  $\mu$ m above the alumina substrate. The gap width, denoted as d in Fig. 6.3(a), is not given in [79] and we choose it as 300  $\mu$ m. The alumina substrate on the package side has a loss tangent of 0.002 and the silicon substrate on the chip side has a resistivity of 2 k $\Omega$ ·cm. As shown in Fig. 6.3(b), the simulation results agree well with the measurement found in [79], which further verifies the implementation of the full-wave analysis in the co-simulation. The slight discrepancy is due to the lack of the geometrical information of the transition in [79].

Next, we assume that heat sink is attached at the bottom surface and air convective cooling is applied to the top surface of the CPW transition. The convective heat transfer coefficient is chosen as  $10 \text{ W/(m^2K)}$ . The signal trace on the chip side is considered as a constant heat source of 17.7 W, denoted as Q in Fig. 6.4(a). The ambient temperature is set at room temperature of 293 K. The maximum temperature varying with time in the CPW transition is depicted in Fig. 6.4(a), which agrees well with that obtained with COMSOL Multiphysics [49]. We also double the value of the heat source [denoted as 2Q in Fig. 6.4(a)] and observe the agreement between the obtained temperature profile and that from COMSOL Multiphysics as shown in Fig. 6.4(a), which further verifies the implementation of the transient thermal analysis.



(a)



Figure 6.3: Wire bondings in the coplanar waveguide transition between alumina and silicon substrates: (a) 3-D view in COMSOL Multiphysics [49] and 2-D topology, and (b) measured and simulated  $S_{21}$ .



Figure 6.4: Nonlinear transient thermal analysis with adaptive time-stepping: (a) maximum temperature varies with time in the CB-CPW structure in Fig. 6.3(a), and (b) the non-uniform time steps.

	а	b	w	$\mathbf{S}$	h
alumina	2045	3000	270	320	635
silicon	2100	3200	300	250	400

Table 6.4: Dimensions  $(\mu m)$  of the CB-CPW in Fig. 6.3(a)

## 6.3.3 Efficiency Enhancement by the Adaptive Time-Stepping

The efficiency enhancement of the co-simulation achieved by the adaptive time-stepping is demonstrated with the example of the wire bondings in the CB-CPW in Fig. 6.3(a). The predefined simulation time is 30 ms with an initial time step set as 0.3 ms. The conventional time-marching scheme is based on uniform time steps and it takes 100 time steps and a total number of 117 iterations to solve the nonlinear transient problem, whereas the adaptive time-stepping takes only 14 time steps and 20 iterations. Besides, the adaptive time-stepping gradually enlarges the time step due to the less significant variance in the transient temperature, which is shown in Fig. 6.4(b), such that the co-simulation can be automatically terminated when the time step exceeds a predefined portion of the simulation duration.

## 6.3.4 Efficiency Enhancement by the FETI-Enabled Parallel Computing

The bonding wire array in Fig. 6.5 is used to demonstrate the parallel efficiency of the FETI. A standard block or a subdomain in Fig. 6.5 consists of six bonding wires connecting a MOS capacitor to the package leads. The demonstration of the parallel efficiency is based on the idea of unit load per processor, under which the number of degrees of freedom (DOFs) assigned to each individual processor is kept the same. All computations are performed on the CISCO Arcetri cluster, with each node equipped with sixteen 2.70-GHz Intel Xeon E5-2680 processors. As shown in Table 6.5, when the total number of DOFs increases from 1.29 to 10.25 million, or the number of bonding wires increases from 48 to 384, the computation time remains similar if proportionally more processors are used. The application of the FETI also decreases the memory usage. For a problem with 2.03 million DOFs, the memory usage reduction by using the FETI in serial computing

is shown in Table 6.6, where the same problem is divided into 4, 8, 16, and 32 subdomains, respectively. More details on the reduction of memory usage and parallel efficiency can be found in [35].

Number of bonding wires	48	96	192	384
Total number of DOFs (million)	1.29	2.56	5.13	10.25
Number of processors	8	16	32	64
Number of subdomains	8	16	32	64
Pre-processing (s)	4.29	4.37	4.39	4.39
Factorization (s)	313.65	310.80	313.21	311.20
Solvering interface problem (s)	97.85	104.16	104.79	104.23
Extracting fields (s)	1.55	1.55	1.53	1.53
Total computation time (s)	417.34	420.88	423.92	421.35

Table 6.5: Unit-load-per-processor parallel efficiency of FETI with MPI

Table 6.6: Memory usage reduction by using FETI in serial computing

Number of subdomains	4	8	16	32
Peak memory usage (GB)	19.5	16.4	13.9	11.7

## 6.4 Co-Simulation Examples

### 6.4.1 Bonding Wire Array

Figure 6.5 shows a portion of an array of 384 bonding wires and 64 MOS capacitors, functioning as an in-package matching network in RF amplifiers. The geometrical information of the bonding wire is as follows: the gold wire has a diameter of 50  $\mu$ m and a maximum height of 350  $\mu$ m; the separation of two neighboring wires is 200  $\mu$ m; the span of the wire is 1000  $\mu$ m; each wire is positioned 100  $\mu$ m behind the edge of the package lead and the MOS capacitor. The MOS capacitor consists of four layers, in the top-down sequence as: a copper layer, an insulation layer of silicon dioxide, silicon substrate, and ground. The thickness and the width of the copper layer are 5 and 500  $\mu$ m,



respectively. The thickness and the width of the insulation layer are 5 and 800  $\mu$ m, respectively.

Figure 6.5: Bonding wire array with MOS capacitors.

## Full-Wave Analysis

As mentioned earlier, the matching network in Fig. 6.5 is essentially an LC circuit. By adjusting the capacitance from the MOS capacitors and the inductance from the bonding wire array, one can observe the shift of the cut-off frequency of the LC circuit. We choose the case of three bonding wires on each side of a MOS capacitor and the substrate with a thickness of 80  $\mu$ m as the reference. Adding another bonding wire in parallel reduces the total inductance. Increasing the height of the silicon substrate to 120  $\mu$ m reduces the capacitance. Through either of the two aforementioned approaches, the cut-off frequency is shifted to a higher value, which is demonstrated in Fig. 6.6.

#### Transient Thermal Analysis

The bonding wire array and the MOS capacitors are immersed in the molding compound of the package. We assume that the heat sink is attached to the bottom of the structure and air convection is applied to the top with the ambient temperature at 293 K. The metal layer of the MOS capacitor is considered as the heat source. The heat sources of 10 W and 15 W are alternatively placed in the MOS capacitors. The temperature profiles at 1


Figure 6.6: Shifts of the cut-off frequency of the matching network in Fig. 6.5 due to the variances of the inductance and capacitance.

ms and 40 ms are depicted in Fig. 6.7. Figure 6.8 shows the maximum temperature and the temperature in the middle of the bonding wire as functions of time.

#### Thermal Stability Analysis

We select the frequency band from 13 to 16 GHz with an insertion loss of 1.4 dB and investigate the thermal stability of the matching network. Assume there is a large biasing current flowing through the bonding wires, which is considered as the heat source. Figure 6.9 shows the variances on the S-parameters under different heat sources. The curves in Fig. 6.9 are obtained as follows: at each sampling frequency, the co-simulation is performed with the full-wave analysis set at that certain frequency; S-parameters are recorded as temperature reaches a steady state. The heat source of 8 W contributes to a temperature increase of 110 K at 14 GHz, which results in the variance of 6.5% for the insertion loss as shown in Fig. 6.9. At 16 GHz, the variance of the insertion loss is as large as 10%. The co-simulation of each sampling frequency takes 20 time steps and 47 iterations.



Figure 6.7: Temperature profiles of the bonding wire array in Fig. 6.5 at (a) 1 ms, and (b) 40 ms.

(b)

### 6.4.2 SIW Step Filter

Figure 6.10 shows a SIW step filter designed in [80], which consists of 12 stubs made of 349 vias. The geometrical information is provided in Table 6.7.

#### Full-Wave Analysis

The S-parameters simulated at the room temperature are shown in Fig. 6.11(a), which match well with the measured results from [80]. The frequency selectivity of the filter is subject to the temperature-dependent permittivity; therefore, the temperature stability analysis becomes a critical design fac-



Figure 6.8: Transient temperature response of the bonding wire array in Fig. 6.5 under different heat sources.



Figure 6.9: Variances in the insertion loss of the matching network in Fig. 6.5 under different heat sources.



Figure 6.10: SIW step filter with tapered microstrip lines.

tor. As shown in Fig. 6.11(b), 3.7% variance in the permittivity leads to a 25% shift of the passing band of the SIW step filter. However, the uniform variance of the permittivity in a complex structure is generally nonrealistic because of the non-uniform temperature profiles. Therefore, the electrical-thermal co-simulation is in need to make a more accurate prediction of the temperature stability.

Table 6.7: Dimensions (mm) of the SIW step filter in Fig. 6.10: r is the radius of the via, p is the separation measured between the center of two adjacent vias, and  $W_l$  is the width of the microstrip line

$L_1$	$L_2$	$L_3$	$L_4$	$L_5$	$W_i$
12.565	15.451	16.128	15.451	12.565	4.0
$H_1$	$H_2$	$H_3$	$H_4$	$H_5$	$H_6$
2.9365	4.5075	5.034	5.034	4.5075	2.9365
$L_a$	$W_a$	$W_{\rm siw}$	r	p	$W_l$
19.0	9.6	21.5	0.25	1.0	2.76

#### Transient Thermal Analysis

The heat source in the thermal analysis is the dielectric loss of the step filter, which is determined by the loss tangent of the substrate and the field distribution under the operation frequency. The snapshots in Figs. 6.12 and 6.13 capture the temperature profiles at two instants. The temperature profile should agree with the field distribution under the operation frequency. As the input power is increased by 20 W, the temperature increase is larger than 10 K, which is shown in Fig. 6.14.

#### Thermal Stability Analysis

The temperature profile is largely determined by the field distributions under a certain power level and its frequency, which explains the non-uniform shift of the passing band of the step filter in Fig. 6.15. Similar procedures described earlier to obtain the curves in Fig. 6.9 are followed to obtain the passing band results in Fig. 6.15. The S-parameters are recorded when the temperature reaches a steady state. It can be seen from Fig. 6.15 that the insertion loss is about 1.5 dB in the passing band from 5.95 to 6.45 GHz. A frequency shift of 40 MHz is observed when the temperature coefficient of the permittivity is 300 ppm/°C. When the temperature coefficient is increased to 500 ppm/°C and the input power is kept as 100 W, the passing band narrows down by 30 MHz.



(a)



Figure 6.11: S-parameters of the step filter in Fig. 6.10: (a) simulated at room temperature and compared with the measurement, and (b) shift under the variance of the substrate permittivity.



```
(a)
```



(b)

Figure 6.12: Temperature profiles of the SIW step filter in Fig. 6.10 with input of 40 W and 6.5 GHz at (a) t = 1 s and (b) t = 39 s.



Figure 6.13: Temperature profiles of the SIW step filter in Fig. 6.10 with input of 40 W and 5.75 GHz at (a) t = 1 s and (b) t = 39 s.



Figure 6.14: Maximum temperature varying with time under different power levels of 6.5 GHz input.



Figure 6.15: The passing band of the SIW step filter altered by temperature.

## CHAPTER 7

# MULTIPHYSICS SIMULATION OF 3-D ICs WITH INTEGRATED MICROCHANNEL COOLING

### 7.1 Problem Statement

Three-dimensional (3-D) integrated circuits (ICs) are regarded as promising solutions to overcome the bottleneck imposed by interconnect scaling problems [81, 82]. In 3-D ICs, multiple layers of planar devices are stacked up vertically to achieve the reduction of global interconnects, in addition to other benefits such as the increased packing density, the reduced chip area, and the enabled possibility of heterogeneous integration. However, the increased power density together with the reduced chip area leads to a significant increase of heat power generation per unit area and poses great challenges for heat removal, which is further complicated by the interlayer dielectrics of very low thermal conductivity. It is known that high temperatures not only tend to degrade the performance of devices or the entire system but may also cause reliability issues or even failures [83]. As temperature becomes a critical factor in analyzing and optimizing the performance and reliability of an electrical design especially at the early stage in a design flow, it is necessary to employ the concept of thermal-ware design [84] and develop the corresponding multiphysics-based computer aided design (CAD) methodologies that address the electrical and thermal aspects of a design and their couplings simultaneously. In the past, electrical-thermal co-simulation has been developed for the design of power distribution networks [43, 31, 70], the thermal-aware electromagnetic analysis of through-silicon-via (TSV) structures [85, 86], the investigation of carbon nanotube contacted phase-change memory arrays [87], and the temperature stability analysis of high-power RF/microwave circuits [88].

Along with the development of multiphysics-based CAD methodologies for thermal-aware design, a considerable amount of research has been devoted to searching for innovative heat removal techniques for 3-D ICs. One such promising techniques is the integrated microchannel cooling [11, 12]. Compared to the conventional back-side heat removal techniques such as heat sinks and air cooling, the integrated microchannel layer has low thermal resistance [11] and it provides a direct path of heat dissipation within individual device layers. A schematic of a 3-D IC with integrated microchannel cooling is provided in Fig. 7.1. In this chapter, a multiphysics-based co-simulation technique is proposed for the characterization of such 3-D ICs. The co-simulation integrates three types of analysis into one single scheme: (1) full-wave electromagnetic analysis based on Maxwell's equations, (2) fluid analysis based on conservation of mass and momentum, and (3) transient conjugate heat transfer analysis based on conservation of energy. The co-simulation considers the motion of fluid flow and takes into account the thermal effects of forced convective liquid cooling with integrated microchannels. The conjugate heat transfer analysis in the co-simulation provides temperature profiles of both solid and fluid with a sufficient accuracy of a 3-D model. Therefore, the co-simulation enables thermal-aware full-wave electromagnetic characterization of a 3-D IC by considering the thermal influence on the electromagnetic characteristics from both convectional back-side and integrated microchannel cooling. The effectiveness of heat removal through the integrated microchannel cooling and its impact on the electromagnetic characteristics of 3-D ICs are demonstrated through numerical examples. Note that the conjugate heat transfer analysis is performed in the transient regime such that the temporal evolution of temperature is available for dynamic thermal management [89].

The co-simulation is based on the finite element method (FEM) [10] for its unmatched capability in handling complex structures and material properties. As the volumetric discretization in FEM for practical 3-D IC designs may lead to system matrices with millions of unknowns, it is critical to enhance the efficiency of the co-simulation in order to deal with large-scale problems. In this chapter, this enhancement is achieved through an adaptive time stepping scheme, a domain decomposition scheme called the finite element tearing and interconnecting (FETI), and FETI-enabled parallel computing. The adaptive time stepping scheme enables the automatic generation and utilization of non-uniform time steps, as well as the automatic detection of thermal steady state. The domain decomposition scheme divides the computational domain into non-overlapping subdomains so that the assembly



Figure 7.1: A schematic of a 3-D IC with integrated microchannel cooling [11].

and factorization of the system matrices of individual subdomains can be performed independently and in parallel with multiple processors at a very high parallel efficiency. These efficiency enhancement methods have been developed earlier for electrical-thermal co-simulation [88]. In this chapter, we extend the capability of the co-simulation and apply these two methods to the electrical-thermal-fluid simulation of large-scale problems. In the proposed co-simulation, individual analyses are allowed to perform on different portions of a 3-D IC design utilizing a different number of decomposed subdomains and correspondingly a different number of processors. Numerical examples are provided to demonstrate the capability and the efficiency of the proposed co-simulation technique. In passing, we note that researchers have also used non-conformal domain decomposition methods [90, 91] for thermal analysis and model order reduction techniques [92] for coupled electricalmechanical simulation to improve the computational efficiency.

### 7.2 Co-Simulation and Efficiency Enhancement

#### 7.2.1 Formulation of the Co-Simulation

Consider the typical problem illustrated in Fig. 7.1. For the full-wave electromagnetic analysis, the vector wave equation

$$\nabla \times \left(\frac{1}{\mu_{\rm r}} \nabla \times \vec{E}\right) - k_0^2 \epsilon_{\rm rc} \vec{E} = -j k_0 Z_0 \vec{J}$$
(7.1)

is solved subject to various boundary conditions including absorbing boundary conditions (ABCs)

$$\frac{1}{\mu_{\rm r}}\hat{n} \times (\nabla \times \vec{E}) + jk_0 \sqrt{\frac{\epsilon_{\rm rc}}{\mu_{\rm r}}}\hat{n} \times (\hat{n} \times \vec{E}) = 0$$
(7.2)

for truncating an infinite computation domain and wave port boundary conditions (WPBCs) for modeling excitations in the circuit structures. In Equations (7.1) and (7.2),  $\mu_{\rm r}$  represents the relative permeability,  $\vec{E}$  denotes the vector electric field,  $k_0$  is the wavenumber of the field,  $\epsilon_{\rm rc} = (\epsilon_{\rm r} - j\sigma/\omega\epsilon_0)$ represents the relative complex permittivity with the electrical conductivity  $\sigma$  and angular frequency  $\omega$  embedded,  $Z_0$  is the intrinsic impedance of free space,  $\vec{J}$  denotes the excitation current, and  $\hat{n}$  is a unit normal vector defined at the boundary surface.

The governing equations of fluid flow are derived based on conservations of mass and momentum. From conservation of mass, the continuity equation can be expressed as

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{u}) = 0 \tag{7.3}$$

where  $\vec{u}$  represents the velocity field and  $\rho$  denotes the density. Incompressible flow is assumed in this chapter where the density remains constant so that Equation (7.3) is reduced to

$$\nabla \cdot \vec{u} = 0 \tag{7.4}$$

From conservation of momentum, the Navier-Stokes equation is obtained,

which has a simplified form for incompressible flow as

$$\rho\left(\frac{\partial}{\partial t} + \vec{u} \cdot \nabla\right) \vec{u} = \rho \vec{g} - \nabla p + \nu \nabla^2 \vec{u}$$
(7.5)

where p is the pressure,  $\nu$  is the viscosity of fluid, and  $\vec{g}$  represents the vector acceleration due to gravity. In this chapter, the flow is assumed to be fully developed such that the velocity field distribution remains purely axial and has no variations along the direction of flow. For a fully developed flow in a rectangular microchannel with the corresponding geometrical parameters highlighted in Fig. 7.2, the velocity field distribution obtained by solving Equations (7.4) and (7.5) subject to the no-slip condition can be expressed as

$$\vec{u}(x,z) = \hat{y} \frac{4H^2 \Delta p}{\pi^3 \nu L} \sum_{n=1,3,5...}^{\infty} \frac{1}{n^3} \sin \frac{n\pi z}{H} \left[ 1 - \frac{\cosh \frac{n\pi x}{H}}{\cosh \frac{n\pi W}{2H}} \right]$$
(7.6)

\_

where  $\Delta p$  is the pressure drop in the channel, and W, H, and L are the width, height, and length of the microchannel, respectively. An integration of the velocity field distribution over the cross-section of the channel yields the expression of the flow rate

$$Q_{\rm r} = \int_{S} \vec{u} (x, z) \cdot d\vec{S}$$

$$= \frac{W H^{3} \Delta p}{12\nu L} \left[ 1 - \sum_{n=1,3,5...}^{\infty} \frac{1}{n^{5}} \frac{192}{\pi^{5}} \frac{H}{W} \tanh \frac{n\pi W}{2H} \right]$$
(7.7)

The average flow velocity of the incompressible flow can thus be obtained as

$$U = \frac{Q_{\rm r}}{WH} \tag{7.8}$$

The governing equations associated with conjugate heat transfer are derived based on conservation of energy. For heat transfer in solid, the governing equation is

$$\rho c_{\rm p} \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + q \tag{7.9}$$

where T represents temperature,  $\rho$  is the density of the materials,  $c_{\rm p}$  is the

specific heat capacity,  $\kappa$  is the thermal conductivity, and q represents the rate of volumetric heat generation. The governing equation for heat transfer in fluid is

$$\rho c_{\rm p} \frac{\partial T}{\partial t} + \rho c_{\rm p} \vec{u} \cdot \nabla T = \nabla \cdot (\kappa \nabla T) + q \qquad (7.10)$$

The governing equations of conjugate heat transfer are solved under the Dirichlet boundary condition

$$T = T_{\rm d} \tag{7.11}$$

where  $T_{\rm d}$  represents a constant temperature, and air convection boundary condition

$$\hat{n} \cdot (\kappa \nabla T) = -h(T - T_{\rm a}) \tag{7.12}$$

where h and  $T_{\rm a}$  are the convective heat transfer coefficient and the ambient temperature, respectively.

It is worth mentioning that for an incompressible flow and temperatureindependent density and viscosity assumed in this chapter, the governing equations of motion, to be specific, Equations (7.4) and (7.5), are decoupled from the temperature field [40, 93]. After the velocity field is solved from Equations (7.4) and (7.5), it can be used to solve for the temperature field through a conjugate heat transfer analysis, governed by Equations (7.9) and (7.10).

The electromagnetic and conjugate heat transfer analyses are coupled through temperature-dependent material properties, which makes the problem nonlinear in nature. At each step in the time-marching scheme, the fixed-point iteration technique is utilized to solve the nonlinear problem. The volumetric heat generation in Equations (7.9) and (7.10) consists of two parts: the power dissipation calculated from the electromagnetic analysis and the heat generation from active device layers. In the conjugate heat transfer analysis, one solves for temperature distributions in both solid and fluid. Based upon the temperature distributions, material properties are updated. With the updated material properties, electromagnetic and conjugate heat transfer analyses are performed in another iteration. The same mechanism is carried out through all the iterations until a convergence is achieved and through all time steps until a thermal steady state is obtained.

The materials and the corresponding properties employed in this chapter are shown in Table 7.1. The temperature-dependent resistivity or permittiv-



Figure 7.2: Geometry of a rectangular microchannel.

	Temperature	Thermal	Density	Specific heat
	coefficient	conductivity	$({\rm kg} {\rm m}^{-3})$	capacity
	$(K^{-1})$	$(W m^{-1} K^{-1})$		$(J kg^{-1}K^{-1})$
Copper	$3.9  imes 10^{-3}$	400	$8.7  imes 10^3$	385
Silicon	$1.3  imes 10^{-3}$	Equation $(8.32)$	$2.3  imes 10^3$	703
Silicon Dioxide	N/A	Equation $(8.33)$	$2.2 \times 10^3$	730
Water	N/A	0.6	$1.0 \times 10^3$	4200

 Table 7.1: Material properties

ity share the same expression as

$$\zeta = \zeta_0 \left[ 1 + \alpha (T - T_0) \right]$$
(7.13)

where  $\zeta_0$  is the resistivity or permittivity at temperature  $T_0$  and  $\alpha$  is the temperature coefficient. The temperature coefficient of copper in Table 7.1 is associated with its resistivity, which has a value of  $1.7 \times 10^{-8}$   $\Omega$ m at the room temperature. The temperature coefficient of silicon in Table 7.1 is associated with its permittivity, which is extracted from the measured data in [94]. At the room temperature, the relative permittivity of silicon is 11.65 and its effective loss tangent is 0.005 [94]. According to the measured data in [76], the thermal conductivity of copper remains very close to 400 Wm<sup>-1</sup>K<sup>-1</sup> from the room temperature and all the way up to 800 K. The temperature dependent thermal conductivities of the silicon and silicon dioxide take the

empirical forms as

$$\kappa_{\rm Si} = 2.475 \times 10^5 \ T^{-1.3} \ (273 {\rm K} \le T \le 1000 {\rm K})$$
(7.14)

$$\kappa_{\rm SiO_2} = 1.02278 + 0.00121T \,(273K \le T \le 600K)$$
 (7.15)

which are based on the measurements in [76].

### 7.2.2 Validation of the Modeling of Integrated Microchannel Cooling

To validate the model of the integrated microchannel cooling, we consider a problem described in [95] and compare the results obtained through the proposed simulation against the experimental studies in [95]. The microchannels under consideration are machined from a copper block and deionized water is driven through the channels by pressured nitrogen gas. There are five different test sections of such microchannels fabricated with different widths [95] and we pick one to perform the validation, which has dimensions as follows:  $W = 194 \ \mu m$ ,  $H = 884 \ \mu m$ , and  $L = 25.4 \ mm$ . An average heat flux of 45 W/cm<sup>2</sup> is applied at the base of the channel and the same value is maintained during the experiment.

In order to perform the comparison, a few parameters are introduced in the following. The Reynolds number is a dimensionless quantity defined as the ratio of inertial force and viscous force. For low Reynolds numbers, laminar flow occurs where the viscous force dominates. The Reynolds number is defined by

$$Re = \frac{UD_{h}}{\nu} \tag{7.16}$$

where the hydraulic diameter  $D_{\rm h}$  is given by

$$D_{\rm h} = \frac{4 \times (\text{cross-sectional area})}{\text{wetted perimeter}} = \frac{2WH}{W+H}$$
(7.17)

The hydraulic diameter associated with this particular example is  $D_{\rm h} = 318 \ \mu {\rm m}$ . The Nusselt number is another dimensionless quantity, which represents the enhancement of heat transfer through convection resulting from a fluid flow in comparison with conduction across the fluid. The average

Nusselt number is defined by

$$Nu = \frac{hD_{h}}{\kappa} \tag{7.18}$$

where h is the average heat transfer coefficient defined as

$$h = \frac{Q}{L(W + 2H)(T_{\rm w} - T_{\rm m})}$$
(7.19)

where  $T_{\rm w}$  is the average wall temperature,  $T_{\rm m}$  is the mean fluid temperature, and Q represents the rate of total heat removed by microchannel cooling. The mean fluid temperature at the outlet is defined by

$$T_{\rm m,o} = \frac{\int_S T \vec{u} \cdot d\vec{S}}{Q_{\rm r}}$$
(7.20)

With these parameters, the average Nusselt number as a function of the Reynolds number is computed and plotted in Fig. 7.3. It can be seen that the simulation results agree very well with those from the measurement. The steady-state temperature distributions at the inlet, outlet, and the middle of the microchannel are depicted in Fig. 7.4. As the incompressible fluid flows through the channels, its temperature increases by absorbing heat from the surroundings. The development of a thermal boundary layer in the fluid is shown clearly in Figs. 7.4 and 7.5. As can be seen, the conjugate heat transfer analysis resolves the exponential decay of the temperature field in the downstream flow of liquid as the thermal boundary layer becomes fully developed, which is known as the thermal-wake phenomenon [96].

#### 7.2.3 Efficiency Enhancement of the Co-Simulation

Because the volumetric discretization in FEM for practical problems may lead to system matrices with millions of unknowns, it is necessary to enhance the efficiency of the co-simulation so that it can be used effectively for analyzing and optimizing practical and large-scale 3-D IC designs. In this chapter, the efficiency of the co-simulation is enhanced through the incorporation of two schemes: the adaptive time stepping scheme for the overall efficiency of the transient analysis and the domain decomposition scheme and its enabled parallel computing for the efficiency associated with each itera-



Figure 7.3: Average Nusselt numbers obtained through the proposed simulation and compared with measurement in [95].



Figure 7.4: Steady-state temperature distributions at (a) the inlet, (b) the middle, and (c) the outlet of the microchannel.



Figure 7.5: Three-dimensional view of the steady-state temperature distribution and the development of the thermal boundary layer.

tion within a time step. The adaptive time-marching scheme is based on the algorithm of proportional-integral-derivative (PID) control and the details for its incorporation into the electrical-thermal co-simulation can be found in [88]. The adaptive time stepping enables the automatic generation of non-uniform time steps and the automatic detection of thermal steady state by utilizing the first-order nature of the thermal response. The domain decomposition scheme divides the computational domain into non-overlapping subdomains and enforces the continuity of field quantities at the subdomain interfaces based on the FETI method. The assembly and factorization of the system matrices of individual subdomains and the calculation of the field in the individual subdomains can be performed independently and in parallel. The iterative solution of the reduced system at the interfaces can also be parallelized with a good parallel efficiency. The domain decomposition scheme thus enables the parallel computing with multiple processors at a very high parallel efficiency, which has been demonstrated in [43], [70], and [86]. All computations are performed on the CISCO Arcetri cluster, with each node equipped with sixteen 2.70-GHz Intel Xeon E5-2680 processors.

### 7.3 Numerical Example

The numerical example used to demonstrate the capability and the efficiency of the developed co-simulation is shown in Fig. 7.6. A substrate-integratedwaveguide (SIW) step filter together with its tapered feeding lines is embedded in a silicon substrate. Another two layers of silicon are placed above



Figure 7.6: Geometry of a 3-D structure including a substrate-integrated-waveguide (SIW) step filter and integrated microchannels.

and below the layer of SIW step filter, respectively. The adjacent layers of silicon are bonded by one interlayer of dielectric made of silicon dioxide. A microchannel layer is integrated in the same layer of the SIW step filter. For simplicity, only the SIW filter layer, the integrated microchannels, and the bottom layer of silicon are shown in Fig. 7.6. The detailed geometrical information of the structure is provided in Table 7.2, where H is the height of the silicon substrate enclosing the SIW filter; W is the width of entire the structure;  $w_{ch}$ ,  $h_{ch}$ , and  $p_{ch}$  are the width, height, and pitch of the integrated microchannels;  $H_d$  is the thickness of the silicon layers on the top and bottom;  $l_i$  (i = 1, 2, ..., 5) is the separation between adjacent stubs of the SIW filter; r denotes the radius of the vias; p is the separation of two neighboring vias and is measured between the centers;  $w_1$  is the width of the input line; and  $w_a$  and  $l_a$  are the width and length of the tapered line, respectively. The height and width of the microchannel are 1.8 and 1.5 mm, respectively.

### 7.3.1 Thermal Analysis with Integrated Microchannel Cooling

To study the problem, the bottom silicon substrate is considered as the layer that accommodates active devices. Due to the heat generation, an input heat flux of  $35 \text{ W/cm}^2$  is applied at the base of the silicon layer that encloses the SIW filter. The average temperatures of the SIW filter layer without

Н	W	$w_{ m ch}$	$h_{\rm ch}$	$p_{ m ch}$	$H_{\rm d}$
3.78	19.58	1.42	1.42	6.75	0.94
$l_1$	$l_2$	$l_3$	$l_4$	$l_5$	$w_i$
5.9294	7.2913	7.6108	7.2913	5.9294	1.8876
$h_1$	$h_2$	$h_3$	$h_4$	$h_5$	$h_6$
1.3857	2.1271	2.3755	2.3755	2.1271	1.3857
la	$w_{\mathrm{a}}$	$w_{\rm siw}$	r	p	$w_{\mathrm{l}}$
8.9661	4.5302	10.1458	0.1180	0.4719	1.3024

Table 7.2: Geometrical information of the structure shown in Fig. 7.6; all the dimensions are in terms of mm



Figure 7.7: Average temperature of the SIW filter layer with and without integrated microchannel cooling.

integrated microchannels and with integrated microchannels of different flow velocities are plotted in Fig. 7.7. Note that the adjacent channels have opposite directions of fluid flow in order to achieve a more uniform temperature distribution at the SIW filter layer. In the situation where no integrated microchannel cooling is employed, the thermal steady state is reached at around 25 s when the average temperature of the SIW filter layer reaches its maximum of 390 K. When eight microchannels are integrated into the SIW filter



Figure 7.8: Top views of the steady-state temperature profiles of the cross section over the microchannel layer with (a) U = 0.01 m/s, (b) U = 0.1 m/s, and (c) U = 0.3 m/s.

layer with an inlet temperature at 293 K and flow velocity of 0.1 m/s, the thermal steady state is achieved at about 12 s and the average temperature of the SIW filter layer is lowered to 342 K. Finally, the integrated microchannels with an average flow velocity of 0.3 m/s lowers the temperature by 52 degrees compared to the situation when no microchannels are employed.



Figure 7.9: Size of time step in the adaptive time stepping scheme under spatially time-varying power.

With different average flow velocities, the temperature development of the fluid differs especially at the interface between fluid and walls, which can be seen from Fig. 7.8. Even though the temperature can be lowered down by injecting coolant into the channels at a higher velocity (0.3 m/s), the microchannels of a lower average velocity (0.1 m/s) is more effective in terms of cooling. The effectiveness of fluidic cooling is proportional to the temperature difference between the outlet and the inlet.

In all four cases shown in Fig. 7.7, adaptive time stepping is utilized to obtain the transient thermal response. The predefined simulation time is set as 40 seconds. In the case with an average flow velocity of 0.1 m/s and the initial time step of 0.1 s, it takes 42 time steps and a total number of 142 iterations to automatically detect the thermal steady state. As the temperature approaches steady state, the time step enlarges itself and the time step size reaches 3.17 s at the last step of the time marching. As a comparison, with the identical setup, the convectional time stepping scheme with a uniform time step size would take 400 time steps.

The adaptive time-stepping scheme has the capability of dealing with spatially time-varying power. In the implementation, a predicted time step should never go beyond the transition point where the power changes. In that case, the transient temperature profile resulting from a power change can be accurately captured. Figure 7.9 shows two cases where the power of the central part changes at 20 seconds, including an increase to  $52 \text{ W/cm}^2$  and a decrease to  $25 \text{ W/cm}^2$ . It can be seen from Fig. 7.9 that the time step is reset to the initial value upon the power change. The total number of time steps increases due to the power change but it is still much smaller than that in the convectional time stepping with a constant time step.

### 7.3.2 Thermal Impact on the Electromagnetic Characteristics

The SIW step filter has a 3 dB passing band between 5.95 and 6.6 GHz. As the temperature increases toward steady state, it alters the material properties, which further impact the electromagnetic characteristics of the filter. As shown in Fig. 7.10(a), as the temperature increases, the passing band shifts toward lower frequencies. The maximum shift of 0.3 GHz occurs at 25 s when the temperature reaches steady state. Out of the passing bandwidth of 0.65 GHz, the shift takes up 46%. However, the instability of the filters due to the temperature increase can be mitigated by integrated microchannel cooling, which is shown in Fig. 7.10(b), where the shift of the passing band is reduced to 0.15 GHz. A slight change of the insertion loss due to temperature rise is shown in Fig. 7.10(c), whereas the bandwidth shows little variance. This is because the temperature beneath the region of the SIW is relatively uniform as shown in Fig. 7.11, which results in the shift of the passing band whereas the bandwidth remains basically unaffected. The sensitivity of the SIW filter resulting from the temperature increase is shown in Fig. 7.12.

#### 7.3.3 Thermal Hotspots Cooling

In the above situations, the SIW filter layer is considered as victim due to the heat generation at the bottom active device layer. However, in the high-power applications, the filter itself may consume a significant amount of power and cause serious thermal issues as well. In this example, in addition to the input heat flux of  $35 \text{ W/cm}^2$  at part of the active device layer, an input of 200 W is injected to the 3-D structure through the wave port of



(a)



(b)

Figure 7.10: Passing band of the SIW filter (a) without and (b) with integrated microchannels as temperature approaches steady state and the zoomed-in view (c). The average fluid velocity is chosen as 0.3 m/s.



Figure 7.10: (cont.)



Figure 7.11: Temperature profile of the SIW filter of the lateral dimension.



Figure 7.12: Sensitivity of the SIW filter resulting from temperature increase over time.

the SIW step filter. Due to the lossy substrate and the corresponding power dissipation in the SIW filter layer, a significant increase of temperature takes place within the structure. In addition, the localized over heating known as thermal hotspots is also observed, which is shown in Fig. 7.13(a). The temperature profile in Fig. 7.13(a) is associated with an input signal of 6.5GHz. At the room temperature, the insertion loss denoted by S21 is 3.3 dB, whereas the insertion loss is increased to 4.8 dB due to the thermal hotspots. To alleviate this problem, the integrated layer of microchannels is adopted to cool the thermal hotspots and mitigate their impact on the electromagnetic characteristics. For example, we can use two microchannels at the locations where thermal hotspots appear, which not only lower the temperature as shown in Fig. 7.14 but also achieve a more uniform temperature distribution as shown in Figs. 7.13(a) and (b). When eight microchannels are employed, the temperature is decreased by 30 degrees, which can be seen in Fig. 7.14. The mitigation of the thermal impact on the electromagnetic characteristics by using the integrated microchannel cooling is shown in Fig. 7.15, for which the average flow velocity is chosen as 0.3 m/s.





Figure 7.13: Temperature distribution with thermal hotspots in the integrated microchannel layer (a) without microchannels, (b) with two microchannels, and (c) with eight microchannels. A slant cut is made across the second microchannel in Fig. 7.13(c) to show the fluid flow in the microchannel.

(c)



Figure 7.14: Average temperature of blocks 1 and 2 with and without the integrated microchannel cooling. Note that block 1 denotes the region surrounding the 7th microchannel and block 2 denotes that surrounding the 8th microchannel.

Table 7.3: Computation characteristics of the co-simulation

Solver	Total	Number	Number	Time	Time of	Time of
	number	of	of	of	solving	extracting
	of DOFs	subdomains	processors	factorization	interface	fields
	(million)				problem	
EM	1.3	9	9	$7.47 \mathrm{\ s}$	20.68 s	$0.92 \mathrm{~s}$
Thermal	1.5	18	18	2.85 s	28.13 s	0.79 s

### 7.3.4 Computational Information

The computational information of the co-simulation is provided in Table 7.3. In the co-simulation, each individual analysis deals with different portions of the design and utilizes a different number of processors. In this particular example, the electromagnetic analysis is only performed on the SIW filter layer, the fluid analysis is applied to the microchannels, and the conjugate heat transfer analysis is applied to the entire structure. The electromagnetic analysis utilizes nine processors and the conjugate heat transfer analysis takes 18 processors. The time recorded in Table 7.3 is for one iteration in solving



Figure 7.15: Impact on the insertion loss from the thermal hotspots and the integrated microchannel cooling.

the nonlinear problem within each time step. It takes at most five iterations for the nonlinear problem to achieve a convergence with the residual reduced below  $10^{-8}$ . As a reference, the factorization in the conventional FEM by a direct solver with nine processors takes 78.2 seconds. The FETI-enabled parallel computing outstands in dealing with large-scale problems due to its high parallel efficiency and scalability. By utilizing proportionally more processors and subdomains, a large problem can be solved within a similar amount of time to that of a much smaller one.

### 7.4 Conclusion

In this chapter, a multiphysics-based co-simulation technique is developed based on the finite element method. The co-simulation integrates full-wave electromagnetic, fluid, and transient conjugate heat transfer analyses into a single scheme in order to characterize a 3-D IC with integrated microchannel cooling. By solving the governing equations of fluid flow subject to the noslip condition, the velocity field is obtained analytically for the incompressible and fully developed flow. Based on the obtained motion of fluid flow, the conjugate heat transfer analysis provides the temperature profiles of both solid and fluid and the corresponding temporal evolution. Due to the temperaturedependent material properties, the full-wave electromagnetic and conjugate heat transfer analyses are coupled through an iterative scheme and a nonlinear problem is solved at each step in the time-marching scheme. As for many practical designs, a 3-D simulation has to deal with very large matrix systems resulting from the volumetric discretization, which can be very computationally intensive. To this end, the adaptive time stepping, domain decomposition, and parallel computing are incorporated in the co-simulation to enhance the computational efficiency. Consequently, the accuracy of a 3-D modeling is attained and at the same time a significant reduction in computation time is achieved. The capability and efficiency of the co-simulation are demonstrated through a 3-D IC design with multiple stacked-up layers. The effectiveness of the integrated microchannel cooling in heat removal from both adjacent active device layers and localized over-heating is illustrated through the numerical example. The thermal impacts on the high-frequency electromagnetic characteristics of the design have also been shown with the numerical example.

## CHAPTER 8

# COUPLED ELECTRICAL-THERMAL-MECHANICAL SIMULATION FOR THE RELIABILITY ANALYSIS OF LARGE-SCALE 3-D INTERCONNECTS

### 8.1 Problem Statement

The advancement of semiconductor technology has been driven by the continuous scaling of process technologies to achieve further reductions of feature sizes and increases in packing densities. However, the continuous scaling results in significant increases of heat power generation per unit area and poses great challenges for heat removal. Under high operating temperatures, reliability of interconnects becomes one of the major concerns in integrated circuit designs. It has been demonstrated that an increase of the operating temperature from room temperature (25 °C) to 100 °C leads to a 30% increase of the time delay [97] because of the temperature-dependent resistivity of metal interconnects. Besides, thermally-induced voltage drops may prevent transistors from switching states and cause malfunctions or even failures of the entire system. Electro-migration is another persistent reliability concern that is temperature-dependent. As described in Black's equation, the mean time to failure (MTF) associated with electro-migration has an exponential dependence on temperature. Another category of reliability concerns is related to the situations when electrical designs undergo thermal loadings. Thermal stresses of substantial amplitudes often arise from large temperature gradients and mismatches of thermal expansion coefficients and tend to concentrate at very small regions during both manufacture and operation [13]. Mechanical fatigue failures can occur because of repeated and localized high thermal stresses. For example, Fig. 8.1(a) shows the lift-off of an aluminum bonding wire occurring at the wire-chip interface under power cycling [98]



Figure 8.1: Mechanical fatigue failures of (a) a bonding wire [98] and (b) a solder joint [99].

and Fig. 8.1(b) shows a crack initiated and propagating within a solder joint on an organic substrate under thermal cycling [99].

Figure 8.1 shows that reliability concerns of interconnects involve multidisciplinary, specifically, electrical, thermal, and mechanical, aspects of a design. Because of the continuous reduction of feature sizes, it becomes more and more challenging to obtain precise values of deformations and stresses through direct measurements. Experimental methods employed to retrieve information of deformations and stresses include but are not limited to X-ray diffraction (XRD) methods based on the changes in lattice spacing for metal lines and wafer curvature methods based on the changes in the curvatures of substrates, both of which can merely provide averaged values of stresses rather than the detailed distributions and localized high values of stresses [100]. Mechanical failures can also be captured through imaging methods such as those with scanning electron microscopy (SEM) and cautiously prepared samples, which, nevertheless, only deliver qualitative information for a further reliability analysis. The retrieval of temperature distribution through measurements can be difficult and challenging too. For example, infrared thermal CCDs are usually employed for measuring temperatures of bonding wires, whereas the image pixel resolution is sometimes greater than the dimension of the wires themselves [101]. To achieve a higher accuracy of temperature measurements for thin bonding wires, optical devices have been considered as alternatives [102].

Comparing to the experimental methods, three-dimensional (3-D) numerical simulations possess unique advantages for providing detailed distributions of temperature, deformations, and stresses in the interconnect structures. Among the many well-known numerical methods, the finite element method (FEM) prevails owing to its unmatched capability of modeling complex geometries and material properties [10]. Therefore, the multi-physics simulation based on FEM becomes an important approach to analyzing the electrical, thermal, and mechanical aspects and the ever increasing couplings among the multidisciplinary aspects of interconnects. In the past, the coupled electrical-thermal-mechanical simulation based on FEM has been proposed and applied to bonding wire structures [103] and through-silicon-via (TSV) structures [104]. Electrical-thermal co-simulations based on FEM have also been proposed for high-frequency characterization of TSV structures [86], power distribution networks (PDNs) analysis [43, 70], and characterization of carbon-based heterogeneous interconnects [105].

It is well known that for large-scale 3-D interconnect structures, the volumetric discretization in FEM may result in system matrices with millions of unknowns. Consequently, despite the existence of the coupled electricalthermal-mechanical simulation, the bottleneck imposed by the computational burden prevents the 3-D simulation from gaining further popularity. Therefore, in this chapter, a coupled electrical-thermal-mechanical simulation with an improved computational efficiency is devised for analyzing practically large-scale problems. The enhanced efficiency in dealing with large-scale problems is realized through utilizing a domain decomposition scheme, parallel computing, and the localized nature of thermal stresses in interconnect structures. The domain decomposition scheme used is called the finite element tearing and interconnecting (FETI), which divides the computational domain into small non-overlapping subdomains so that the assembly and factorization of the system matrices of individual subdomains can be performed independently and in parallel with multiple processors at a very high parallel efficiency. The steady-state temperature distributions associated with individual subdomains are taken as input to the mechanical analysis to obtain distributions of deformations, stresses, and strains. The mechanical analysis is performed within individual subdomains and in parallel. Consequently, the proposed coupled simulation is able to simulate and analyze practically largescale problems in a highly efficient manner. To validate the implementation and demonstrate the capability and efficiency of the proposed simulation, numerical examples involving arrays of solder bumps and bonding wires are provided.

### 8.2 Coupled Simulation and Parallel Computing

This section provides detailed formulations associated with individual physics, the finite element implementation, and the coupling and parallel computing schemes in the multi-physics simulation.

### 8.2.1 Formulations of Individual Physics

The electrical analysis starts with the current continuity equation, which is expressed as

$$\nabla \cdot \left( \vec{J} + \frac{\partial \vec{D}}{\partial t} \right) = 0 \tag{8.1}$$

The current continuity equation can be re-written in terms of electrical potential  $\phi$  as

$$\nabla \cdot \left( \sigma_E \nabla \phi + \epsilon_E \nabla \frac{\partial \phi}{\partial t} \right) = 0 \tag{8.2}$$

through the relations  $\vec{J} = \sigma_E \vec{E}$ ,  $\vec{D} = \epsilon_E \vec{E}$ , and  $\vec{E} = -\nabla \phi$ , where  $\epsilon_E$  and  $\sigma_E$  are the permittivity and electrical conductivity of the materials, respectively. The inclusion of the displacement current is to account for the capacitive effect. In the electrical analysis, Equation (8.2) is solved subject to the Dirichlet boundary condition

$$\phi = \phi_{\rm d} \tag{8.3}$$

and the impedance boundary condition

$$\hat{n} \cdot \sigma_E \nabla \phi = \frac{\phi}{RS} \tag{8.4}$$

where R and S are the resistance and the cross-sectional area of the impedance boundary, respectively.

The governing equation of the thermal analysis is based on conservation of energy and can be written as

$$\rho c_{\rm p} \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + q \tag{8.5}$$

where  $\rho$  is the mass density,  $c_{\rm p}$  is the specific heat capacity,  $\kappa$  is the thermal conductivity, T denotes temperature, and q represents the rate of volumetric heat generation. In the thermal analysis, Equation (8.5) is solved with the Dirichlet boundary condition

$$T = T_{\rm d} \tag{8.6}$$

where  $T_{\rm d}$  denotes a constant temperature, and air convection boundary condition

$$\hat{n} \cdot (\kappa \nabla T) = -h(T - T_{\rm a}) \tag{8.7}$$

where h and  $T_{\rm a}$  are the convective heat transfer coefficient and the ambient temperature, respectively.

The mechanical analysis in the coupled simulation is based on the theory of linear thermo-elasticity, which consists of the infinitesimal straindisplacement relations

$$\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{8.8}$$

the constitutive relation

$$\sigma_{ij} = C_{ijkl} \left( \epsilon_{kl} - \alpha \delta_{kl} \Delta T \right) \tag{8.9}$$

and the conservation of momentum (Newton's law)

$$\rho \frac{\partial v_i}{\partial t} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \tag{8.10}$$

where  $x_i(i = x, y, z)$  denotes the space coordinate,  $u_i$ ,  $v_i$ ,  $f_i$  are the components of the displacement vector, velocity vector, and body force per unit volume, respectively,  $\epsilon_{ij}$  and  $\sigma_{ij}$  represent the components of the strain and stress tensors, respectively,  $C_{ijkl}$  denotes the stiffness coefficient,  $\alpha$  represents the thermal expansion coefficient, and  $\Delta T$  is the temperature change from a reference, in this chapter, the room temperature, at which the structure is regarded as free of stress. Note that steady-state assumption is made in this chapter such that the left-hand side of Equation (8.10) becomes zero. However, we still keep the transient term in Equation (8.10) to make a more general formulation. By substituting Equations (8.8) and (8.9) into Equation (8.10), one arrives at the governing equation of the thermal-stress analysis,
which can be written as

$$\frac{1}{2}\frac{\partial}{\partial x_j} \left[ C_{ijkl} \left( \frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right) \right] = -\frac{\partial}{\partial x_j} \left( C_{ijkl} \alpha \delta_{kl} \Delta T \right) - f_i$$
(8.11)

In the thermal-stress analysis, Equation (8.11) is solved under the displacement boundary condition

$$u_i = u_d \tag{8.12}$$

where  $u_{\rm d}$  denotes a constant displacement component.

#### 8.2.2 Finite Element Implementation

In the finite element implementation of the thermal-stress analysis, the computational domain is discretized into tetrahedrons. The displacement vector  $\{u\}$  at an arbitrary point within a tetrahedron can be related to the displacement components at the nodes of the tetrahedron through interpolation, which can be written as

$$\{u\} = [N] \{q\} \tag{8.13}$$

where  $\{q\}$  contains the displacement components at all the nodes and [N] containing the interpolation functions is in the form of

$$[N] = \begin{bmatrix} N_1 & 0 & 0 & \cdots & N_i & 0 & 0 & \cdots \\ 0 & N_1 & 0 & \cdots & 0 & N_i & 0 & \cdots \\ 0 & 0 & N_1 & \cdots & 0 & 0 & N_i & \cdots \end{bmatrix}$$
(8.14)

According to the infinitesimal strain-displacement relations in Equation (8.8), the strain vector  $\{\epsilon\}$  can be represented by the displacement components at all the nodes as

$$\{\epsilon\} = [D] \{u\} = [D] [N] \{q\} = [B] \{q\}$$
(8.15)

The strain vector  $\{\epsilon\}$  includes all the six different strain components, which can be explicitly written as

$$\{\epsilon\} = \left\{\epsilon_x \quad \epsilon_y \quad \epsilon_z \quad \gamma_{xy} \quad \gamma_{yz} \quad \gamma_{zx}\right\}^T \tag{8.16}$$

and the differential operator [D] is expressed as

$$[D] = \begin{bmatrix} \frac{\partial}{\partial x} & 0 & 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial z} \\ 0 & \frac{\partial}{\partial y} & 0 & \frac{\partial}{\partial x} & \frac{\partial}{\partial z} & 0 \\ 0 & 0 & \frac{\partial}{\partial z} & 0 & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \end{bmatrix}^{T}$$
(8.17)

Similarly, the stress vector  $\{\sigma\}$ , containing all the six different stress components, can be written as

$$\{\sigma\} = \left\{\sigma_x \quad \sigma_y \quad \sigma_z \quad \tau_{xy} \quad \tau_{yz} \quad \tau_{zx}\right\}^T \tag{8.18}$$

According to the constitutive relation in Equation (8.9), the stress vector can also be written in terms of the displacement components at all the nodes as

$$\{\sigma\} = [E_{\rm c}] (\{\epsilon\} - \{\epsilon_T\}) = [E_{\rm c}] [B] \{q\} - [E_{\rm c}] \{\epsilon_T\}$$
(8.19)

It is worth mentioning that  $\{\epsilon_T\}$  represents the strain due to temperature change, which is expressed as

$$\{\epsilon_T\} = \left\{ \alpha \Delta T \quad \alpha \Delta T \quad \alpha \Delta T \quad 0 \quad 0 \quad 0 \right\}^T \tag{8.20}$$

The constitutive matrix  $[E_c]$  can be written as

$$[E_{c}] = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{bmatrix}$$
(8.21)

where

$$\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}$$
(8.22)  
$$\mu = \frac{E}{2(1+\nu)}$$
(8.23)

Note in Equations (8.22) and (8.23), E is Young's modulus and  $\nu$  is Poisson's ratio. Through the finite element implementation together with Equations

(8.13), (8.15), and (8.19), a matrix equation can be obtained as

$$[K] \{q\} = \{f_e\} + \{f_T\}$$
(8.24)

where

$$[K] = \int_{V} [B]^{T} [E_{c}] [B] dV \qquad (8.25)$$

$$\{f_e\} = \int_{V} [N]^T \{f_V\} \, \mathrm{d}V + \int_{S} [N]^T \{f_S\} \, \mathrm{d}S$$
(8.26)

$$\{f_T\} = \int_V [B]^T [E_c] \{\epsilon_T\} \,\mathrm{d}V \tag{8.27}$$

Note that in Equation (8.26),  $\{f_V\}$  and  $\{f_S\}$  are vectors containing the body and surface forces, respectively. Upon obtaining the values of displacement fields, one can calculate strains and stresses through Equations (8.15) and (8.19). In order to determine if the computed stresses will bring out reliability concerns, the Von Mises yield condition is adopted in this chapter, which is based on the distortion energy theory [13]. The Von Mises stress is defined by the second invariant of the stress tensor and can be written as

$$\sigma_{\text{Von}} = \left\{ \frac{1}{2} \left[ (\sigma_x - \sigma_y)^2 + (\sigma_y - \sigma_z)^2 + (\sigma_z - \sigma_x)^2 \right] + 3 \left( \tau_{xy}^2 + \tau_{yz}^2 + \tau_{zx}^2 \right) \right\}^{\frac{1}{2}}$$
(8.28)

The detailed finite element implementations of the electrical and thermal analyses can be found in [70].

#### 8.2.3 Coupling and Parallel Computing Schemes

The multi-physics simulation consists of electrical-thermal and thermal-stress simulations. In the electrical-thermal simulation, the dissipated power calculated in the electrical analysis is taken as the heat source in the thermal analysis and the material properties in both the electrical and thermal analyses are updated from the temperature distributions. The dissipated power can be written as

$$P_{\text{Joule}} = \sigma |\vec{E}|^2 \tag{8.29}$$

Due to the nonlinearity brought by the temperature-dependent material properties, fixed-point iteration is used at each time step of the transient analysis. The mechanical coupling term between strain and temperature in the heat conduction equation, Equation (8.5), is neglected by assuming

$$T\frac{\partial \dot{\sigma}_i}{\partial x_i} = 0 \tag{8.30}$$

such that the coupling in the thermal-stress simulation becomes one-way and only from temperature to stress.

In order for the multi-physics simulation to become powerful and practical, its computational efficiency has to be improved while dealing with large-scale problems. This can be achieved by utilizing a highly efficient domain decomposition scheme, parallel computing, and the localized nature of thermal stresses in the interconnect structures. The domain decomposition scheme used in the electrical-thermal simulation is called FETI, which divides the computational domain into non-overlapping subdomains and stitches the subdomains back through the enforcement of continuity of voltage, current, temperature, and heat flux at the subdomain interfaces. As such, the factorization of one single large system matrix is avoided and the assembly and factorization of much smaller system matrices belonging to individual subdomains can be performed independently and in parallel. The calculation of voltage and temperature can also be performed completely in parallel and within individual subdomains. Therefore, FETI enables the parallel computing with multiple processors at a very high parallel efficiency, which has been demonstrated in [35] and [106]. Besides, unlike in the electrical analysis where electrical fields propagate through the entire computational domain, thermal stresses in the interconnect structures are often localized and tend to concentrate at very small regions of large temperature gradients and mismatches of thermal expansions. The localized nature of thermal stresses can be utilized in the multi-physics simulation and stresses can be calculated based on the steady state temperature distributions of individual subdomains, which is also in parallel with multiple processors. All computations are performed on the CISCO Arcetri cluster, with each node equipped with sixteen 2.70-GHz Intel Xeon E5-2680 processors.

#### 8.3 Numerical Examples

In this section, numerical examples are provided to verify the implementation of the multi-physics simulation and to demonstrate its capability and efficiency. The material properties appearing in the numerical examples are listed in Table 8.1 for metal interconnects and Table 8.2 for dielectrics. The temperature-dependent resistivity of metal interconnects can be written as

$$\zeta = \zeta_0 \left[ 1 + \alpha_{\rm T} (T - T_0) \right] \tag{8.31}$$

where  $\zeta_0$  denotes the resistivity at temperature  $T_0$  and  $\alpha_T$  is the temperature coefficient. The temperature-dependent thermal conductivities include

$$\kappa_{\rm Si} = 2.475 \times 10^5 \ T^{-1.3} \ (273 {\rm K} \le T \le 1000 {\rm K})$$
(8.32)

for silicon substrate [70, 107], and

$$\kappa_{\rm SiO_2} = 1.02278 + 0.00121T \,(273K \le T \le 600K) \tag{8.33}$$

for silicon dioxide [70, 108], and

$$\kappa_{\text{Al}_2\text{O}_3} = 5.5 + 34.5 \text{ e}^{-0.0033(T - T_0)}$$

$$(273\text{K} \le T \le 1000\text{K})$$
(8.34)

for alumina [77].

#### 8.3.1 Validation of Thermal-Stress Modeling

The validation of the electrical-thermal part in the multi-physics simulation can be referred in [43, 70, 86]. The beam shown in Fig. 8.2 is used to validate the thermal-stress part in the simulation. The copper-made beam has the dimension of  $30 \times 0.5 \times 0.5 \ \mu \text{m}^3$ . The top and bottom surfaces of the beam are held at 333 K and 293 K, respectively, such that the variation of temperature is linear along the z-axis. First, one end of the beam is fixed to form a cantilever. Due to the linear variation of temperature along the z-axis, deflection is observed and shown in Fig. 8.2. Note that the blurred region in Fig. 8.2 represents the initial location of the cantilever before deflection.



Figure 8.2: Temperature profile of a cantilever before and after a thermally-induced deflection.



Figure 8.3: Deflections of a cantilever and a two-fixed-ended beam.

	Electrical conductivity $(S m^{-1})$	$\begin{array}{c} \text{Temperature} \\ \text{coefficient} \\ (\text{K}^{-1}) \end{array}$	Thermal expansion coefficient $(10^{-6} \text{ K}^{-1})$	Young's modulus (GPa)	Poisson's ratio
Copper	$5.98 \times 10^7$	$3.9 \times 10^{-3}$	17	110	0.35
Nickel	$1.43  imes 10^7$	$6.0  imes 10^{-3}$	13.4	210	0.31
Solder	$6.67  imes 10^6$	$4.2 \times 10^{-3}$	21	10	0.4

Table 8.1: Material properties of metal

	Relative permittivity	$\begin{array}{c} \text{Thermal} \\ \text{conductivity} \\ (\text{W } \text{m}^{-1}\text{K}^{-1}) \end{array}$	Thermal expansion coefficient $(10^{-6} \text{ K}^{-1})$	Young's modulus (GPa)	Poisson's ratio
Si	11.9	Equation $(8.32)$	2.6	170	0.28
Silicon Dioxide	4.0	Equation $(8.33)$	0.5	70	0.17
Alumina	9.4	Equation $(8.34)$	8	300	0.222

 Table 8.2: Material properties of dielectric

The deflection of a cantilever under the linear variation of temperature has an analytical solution. A good agreement between the analytical solution and that obtained from the proposed simulation is demonstrated in Fig. 8.3. For a further validation, both ends of the beam are fixed. Under the linear variation of temperature along the z-axis, the deflection at the top surface of the beam is depicted in Fig. 8.3. It can be seen that expansion at the ends of the beam first occurs and compression takes place toward the center of the beam. As a comparison, the solution obtained through COMSOL Multiphysics [49] is also shown in Fig. 8.3, which agrees well with that obtained with the proposed simulation.

### 8.3.2 Large-Scale Analysis of Solder Bump Array

Since first introduced by IBM in 1960s as a means of interconnection between die and thick film metalization on alumina, solder bumps have been widely used in the flip-chip technology. Fatigue failure arising from fluctuating temperatures and alternating stresses is considered as the most serious threat to the integrity of solder bumps [109]. The fatigue failure analysis of solder bumps normally involves multidisciplinary aspects of the interconnects as well as a large problem size, which is a perfect test case to demonstrate the capability of the proposed large-scale simulation. Figure 8.4 shows an array of solder bumps, consisting of a total number of 96 solder bumps connected through redistribution layers on the top and bottom and grouped into 16 columns. Each column forms a daisy chain structure. The detailed geometrical information is as follows: the length of the nickel-made electrodes (a in Fig. 8.4) is 130  $\mu$ m; the height of the solder bump (h in Fig. 8.4) is 240  $\mu$ m; the radius of the solder bump is 170  $\mu$ m; the width of the copper-made redistribution layer is 300  $\mu$ m; and the pitch measured between the centers of two adjacent columns is 740  $\mu$ m. At the two ends of each column, terminals are defined as power and ground, respectively. The bottom surface of the structure is assumed to have the constant room temperature of 293 K.

A trapezoidal pulse train is injected into the solder bump array through the power terminals. As shown in Fig. 8.5, the pulse has a rising time of 3 ms and falling time of 2 ms. The maximum temperature within the solder bump array varying with time is depicted in Fig. 8.5. The temperature stabilizes at 5.5 ms, reaching 375 K. The temperature profiles at 1 ms and 5.5 ms are captured in Fig. 8.6, which shows localized overheating occurring within the solder bumps and the nearby redistribution layer. The dissimilar materials in the structure expand at very different rates due to the large temperature gradient and the mismatches of thermal expansion coefficients. Therefore, large stresses take place as shown in Fig. 8.7. Figure 8.7(a) is a 3-D view of the distributions of Von Mises stresses where the interfaces of the dissimilar materials have localized high values of stresses. The maximum stress of 220 MPa occurs around the edges of the solder bumps and the middle portion of the redistribution layer. It can be seen from Fig. 8.7(b) that the stresses concentrate in small areas and decrease rapidly away from those locations. As a reference, the plastic yield strength of a thin film copper is 225 MPa for a thickness of 3.015  $\mu$ m and 300 MPa for a thickness of 0.885  $\mu$ m [100]. The plastic yield strength also varies with dimensions and grain sizes; for example, the yield strength of micromachined copper thin film beams is estimated in the range from 2.8 to 3.09 GPa based on measured inelastic deformation [100].



Figure 8.4: Geometry of a solder bump array.



Figure 8.5: Maximum temperature of the solder bump array in Fig. 8.4 under the stress of a trapezoidal pulse train.



Figure 8.6: Temperature distribution of the solder bump array in Fig. 8.4 at (a) 1 ms and (b) 5.5 ms.



Figure 8.7: Distribution of Von Mises stresses within the solder bump array in Fig. 8.4 of (a) 3-D view and (b) top-view at the surface between the redistribution layer and the substrate.



Figure 8.8: Geometry of an array of bonding wires.

#### 8.3.3 Large-Scale Analysis of Bonding Wire Array

Bonding wire is another form of widely used interconnections in the semiconductor industry. Despite the fast growing flip-chip technology and its many variations, wire bondings are still irreplaceable in the foreseeable future [78]. It is well known that thermal and mechanical reliabilities are major concerns in the wire bonding interconnects. Therefore, the array of bonding wires in Fig. 8.8 is taken as the second example to demonstrate the capability of the proposed simulation. There are a total number of 32 bonding wires, which are grouped into 16 columns. The detailed geometrical information is as follows: the height of the silicon substrate (hs in Fig. 8.8) is 2 mm; the height of the alumina substrate (ha) is 4 mm; the heights of the two copper-made bonding wires (h1 and h2) are both 3.2 mm; the width of the copper-made landing pad is 1 mm; the radius of the bonding wire is 0.15 mm; and the pitch measured between two adjacent bonding wires in the neighboring columns is 5 mm. Note that in Fig. 8.8, each wire has one end of ball bonding and the other end of wedge bonding. The bottom surface of the structure is assumed to have the room temperature of 293 K. Within each column, the power terminals are defined on the alumina side and the ground terminals on the die side.

A trapezoidal pulse train is injected into the bonding wire array through the power terminals. As shown in Fig. 8.9, both the rising and the falling time of each pulse are 5 ms. Under the stress of the pulse train, the tem-



Figure 8.9: Maximum temperature of the bonding wire in Fig. 8.8 under the stress of a trapezoidal pulse train.

perature within the structure quickly ramps up, reaching 561 K at 120 ms. The maximum temperature varying with time is depicted in Fig. 8.9. The temperature distributions at 30 ms and 120 ms are captured in Fig. 8.10. As shown in Fig. 8.10(b), the maximum temperature occurs at the wedge bond heel and near the peak of the wires. Over the long run as the device is repeatedly heated up and cooled down under operations, the huge temperature fluctuations and the thermally-induced excursions at the wedge bond heel will often result in failures. Similar to the example of the solder bump array, distributions of Von Mises stresses of the bonding wire array are provided in Fig. 8.11 with a 3-D view and two zoomed-in views at the ball and wedge bonds, respectively. As shown in Fig. 8.11(b), large stresses take place at the interfaces between the ball pad and the ball neck and between the ball pad and the substrate, which are the primary causes of wire lift-offs. The large stresses at the wedge bond and wedge bond heel due to thermally induced flexing will also cause cracks to initiate and propagate and eventually lead to failures. All the aforementioned failure modes are successfully predicted by the proposed simulation with detailed temperature and stress distributions.



Figure 8.10: Temperature distribution of the bonding wire structure in Fig. 8.8 at (a) 30 ms and (b) 120 ms.

(b)



(a)

(b)



Figure 8.11: Distribution of Von Mises stresses within the bonding wire array in Fig. 8.8 of (a) 3-D view, (b) zoomed-in view at the ball bond, and (c) zoomed-in view at the wedge bond.

Example	Solver	DOFs (million)	Number	Number	Time	Time	
			of	of	of	of	
			subdomains	processors	factorization	one step	
Solder bump	Electrical	5.3	16	16	$23.5~\mathrm{s}$	58.6 s	
	Thermal	5.3	16	16	24.8 s		
	Mechanical	15.9	16	16	$257.5~\mathrm{s}$	$287.5~\mathrm{s}$	
Bonding wire	Electrical	1.3	16	16	2.9 s	10.8 s	
	Thermal	1.3	16	16	3.1 s		
	Mechanical	3.9	16	16	11.5 s	$20.3 \mathrm{~s}$	

Table 8.3: Computational information of the proposed simulation

#### 8.3.4 Computational Information

The computational information for the large-scale analysis of the solder bump and bonding wire arrays is provided in Table 8.3. In both numerical examples, a total of 16 processors are employed. The time of one step in the electrical-thermal simulation includes a few Newton-like iterations because of the nonlinear nature of the problem. It takes around five iterations for the nonlinear problem to achieve a convergence with the residual reduced below  $10^{-8}$ . Within each Newton-like iteration, it takes around four iterations for the interface problem [35, 70, 106] arising from FETI to achieve a convergence with the residual reduced below  $10^{-6}$ . As the steady-state assumption is made in the mechanical analysis, the time of one step provided in Table 8.3 is the total time. It is worthy pointing out that factorizing the linear system of 1 million degrees of freedoms (DOFs) takes 257.5 seconds whereas with FETI-enabled parallel computing the factorization time of the linear system of 5.3 million DOFs is less than 25 seconds.

## 8.4 Conclusion

In this chapter, a coupled electrical-thermal-mechanical simulation is developed based on the finite element method. The proposed simulation has the capabilities of providing detailed distributions of temperature, deformations, and stresses. It integrates electrical-thermal and thermal-stress simulations into a single scheme in order to characterize the multidisciplinary aspects of interconnects in the reliability analysis. The coupling scheme adopted in the thermal-stress simulation is one-way based on the assumption that the mechanical coupling between temperature and strain in heat conduction is negligible. Due to the temperature-dependent material properties, the fixed-point iteration is employed in the electrical-thermal simulation to solve the nonlinear problem at each step in the transient analysis. To improve the computational efficiency of the coupled simulation for solving practically large-scale problems, FETI, parallel computing, and the localized nature of thermal stresses in the interconnect structures are utilized. Consequently, the advantages of a 3-D modeling are attained and significant reductions in computation time are achieved. The capability and efficiency of the proposed simulation are demonstrated through large-scale analysis of solder bump and bonding wire arrays, both of which are not only major forms of interconnects but also are attracting intensive concerns of reliability. The coupled simulation is able to provide detailed distributions of temperature, deformations, and stresses in a highly efficiently manner. As shown in the numerical examples, concentrated stresses of large amplitude associated several failure mechanisms have been successfully predicted by the proposed simulation.

## CHAPTER 9

# MULTIPHYSICS MODELING IN MICRO-AND NANO-ELECTROMECHANICAL SYSTEMS

## 9.1 Problem Statement

Many micro- and nano-scale devices have been developed such as ultraviolet nanowire lasers [110], nanowire arrays [111], carbon nanotubes [112], and graphene-based devices [113] for a broad range of applications including electronics, optoelectronics, thermoelectronics, electromechanics, and sensing. As the dimension of the devices continues being scaled at the micrometer and even nanometer ranges, thermal characterization becomes critical in evaluating their performance and resolving reliability concerns. It is also important to understand the thermal behaviors of the micro- and nano-scale devices in order to develop novel materials, for example, in designing thermal insulation and achieving thermoelectric energy recovery. Therefore, many studies have recently been carried out to characterize the thermal responses of materials, devices and structures [110, 111, 112, 113, 114, 115, 116, 117] related to micro- and nano-electromechanical systems (MEMS and NEMS).

However, the characterization of thermal response in MEMS and NEMS is by nature a multiphysics problem, which should not be limited to the understanding of heat transfer only. For example, the heat source triggering the thermal response in MEMS and NEMS is often a result from electrical current flowing or electromagnetics wave propagating through the system. Mechanical stress is often induced in MEMS and NEMS due to large temperature gradient and mismatches of thermal expansion coefficients, which is critical in both designing the functionality and understanding the reliability of a micro- or nano-scale electromechanical device. Therefore, we apply the multiphysics simulation to characterize the thermal responses of MEMS and NEMS in this chapter. The multiphysics simulation consists of the electrical-thermal [43, 70, 86, 88] and the thermal-elastic simulations. In the electrical-thermal simulation, the dissipated power under a fixed direct current (DC) input voltage is calculated in the electrical analysis and is considered as the heat source in the thermal analysis. The electrical-thermal simulation is formulated in the transient regime such that both temporal and spatial profiles of temperature can be obtained. Because of temperaturedependent material properties, the electrical-thermal simulation deals with a nonlinear problem within each individual time step. The thermal-elastic simulation takes the steady-state temperature profile as input and calculates the displacement fields within a certain structure. It is necessary to have the multiphysics simulation formulated in three dimensions because both the lateral and vertical directions are critical paths of heat conduction in the suspended metal resistive bridge. The 3-D simulation is implemented based on the finite element method (FEM) because of its unmatched capability of modeling complex geometries and material properties [10].

To benchmark the proposed multiphysics simulation, a multilayer suspended resistor is fabricated and the thermal measurement technique based on diffraction phase microscopy (DPM) is performed on the fabricated suspended resistor. DPM is a real-time quantitative phase imaging (QPI) method [118], which is single-shot and non-destructive. DPM has been applied in various situations to accurately monitor nanoscale dynamics in-situ [119, 120, 121, 122, 123, 124, 125]. To be specific, the DPM technique in this chapter is epi-illumination diffraction phase microscopy (epi-DPM), which is to reconstruct three-dimensional shape of reflective structures. In this chapter, we applied epi-DPM to characterize the suspended metal resistive bridge undergoing thermal expansion, of which the accuracy of height measurement in nanometer scale is achieved. It is worth mentioning that the structure of the suspended metal resistive bridge has wide applications as a heating element for tuning photonic membrane devices [126, 127] and for actuating MEMS devices [128, 129].

Other than the epi-DPM method adopted in this chapter, scanning thermal microscopy (SThM) is one of the promising scanning probe microscopy (SPM) based measurement techniques, which has been used to investigate nanoscale thermal phenomena [130, 131] over the past 20 years. Huber and co-workers investigated a 15 nm transient thermal expansion of a scanning tunneling microscope (STM) tip based on a combination of an atomic force microscope (AFM) and a laser as an external heating source [132]. A group led by Meriles and Riedo used a nitrogen-vacancy center in diamond attached to the apex of a silicon thermal tip as a local scanning temperature sensor [133]. They obtained nanometer-resolved thermal conductivity maps by combining magnetic resonance and AFM. These methods rely on built-in thermal sensors on STM or AFM probes, which result in two major challenges: one challenge is the uncertainty of the heat conduction between the tip and substrate; and the other challenge comes from the fabrication of the nanoscale tip and temperature sensor, which requires advanced nanofabrication techniques. Other nanoscale temperature measurement techniques based on optical probing methods have non-contact and non-destructive probing features. However, conventional optical microscopy methods with visible light cannot achieve a lateral resolution better than a few hundred nanometers. The near-field scanning optical microscopy (NSOM) methods, such as nanoscale fluorescence thermometry [134, 135] and Raman thermometry [117, 136], can successfully overcome the diffraction limit and produce an enhanced optical field within a very small region. However, the localized heating induced by the strong optical field can influence the temperature of the device under test or be destructive to the device. The aforementioned scanning methods also share a common limitation, which is the low speed in acquiring images over a large area. Due to the corresponding limitations of the aforementioned techniques, we apply the epi-DPM technique on the fabricated sample to perform the benchmark.

## 9.2 Simulation Methodology

The governing equations of the electrical-thermal simulation include the current continuity equation

$$\nabla \cdot \left( \vec{J} + \frac{\partial \vec{D}}{\partial t} \right) = 0 \tag{9.1}$$

and the heat conduction equation based on conservation of energy

$$\rho c_{\rm p} \frac{\partial T}{\partial t} = \nabla \cdot (\kappa \nabla T) + q \qquad (9.2)$$

where  $\vec{J}$  represents the conduction current density,  $\vec{D}$  denotes the electric displacement field, T denotes the temperature,  $\rho$  represents the mass density,  $c_{\rm p}$  is the specific heat capacity, and  $\kappa$  denotes the thermal conductivity. The rate of volumetric heat generation q in Equation (9.2) is associated with the volumetric power density dissipated in the gold layer, which can be written as

$$P_{\rm Joule} = \sigma |\vec{E}|^2 \tag{9.3}$$

where  $\sigma$  is electrical conductivity and  $\vec{E}$  denotes the electric field. Convection heat transfer is described through the boundary condition

$$\hat{n} \cdot (\kappa \nabla T) = -h \left( T - T_{\rm a} \right) \tag{9.4}$$

where  $\hat{n}$  is a unit normal vector pointing outward to the surface, h is the convection heat transfer coefficient, and  $T_{\rm a}$  denotes the ambient temperature. Radiation heat transfer is also taken into account through a boundary condition, which can be written as

$$\hat{n} \cdot (\kappa \nabla T) = -\epsilon_{\rm rad} \sigma_{\rm rad} \left( T^4 - T_{\rm a}^4 \right) \tag{9.5}$$

where  $\epsilon_{\rm rad}$  is the surface emissivity and  $\sigma_{\rm rad}$  is the Stefan-Boltzmann constant. The resistivity of gold is temperature dependent

$$\rho_{\rm Au} = \rho_{\rm Au0} \left[ 1 + \alpha_{\rm T} (T - T_0) \right] \tag{9.6}$$

where  $\rho_{Au0}$  is the resistivity of gold at reference temperature  $T_0$  and  $\alpha_T$  is the temperature coefficient of the resistivity. The governing equations of the thermal-elastic simulation [13] consist of the conservation of momentum

$$\rho \frac{\partial v_i}{\partial t} = \frac{\partial \sigma_{ij}}{\partial x_j} + f_i \tag{9.7}$$

the infinitesimal strain-displacement relations

$$\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \tag{9.8}$$

and the constitutive equation

$$\sigma_{ij} = C_{ijkl} \left( \epsilon_{kl} - \alpha \delta_{ij} \Delta T \right) \tag{9.9}$$

where  $u_i$ ,  $v_i$ ,  $f_i$  represent the components of the displacement vector, the velocity vector, and the vector of body force per unit volume, respectively, and  $x_i(i=x, y, z)$  are the space coordinates,  $C_{ijkl}$  and  $\alpha$  are the stiffness and thermal expansion coefficient, respectively,  $\Delta T$  is the temperature change from a reference, in this chapter, the room temperature, and  $\epsilon_{ij}$  and  $\sigma_{ij}$  denote the component of the strain and stress tensors, respectively. In the thermal-elastic simulation, the displacement fields are calculated at thermal steady state when the temperature distribution stabilizes. The simulation volume includes the sample and the air above it. A mixed boundary condition consisting of convection and radiation is applied at the sample-air interface. An adiabatic boundary condition is used on the bottom sample interface with the tape. For simplicity, adiabatic boundary conditions were also applied on the four sidewalls of the sample because there is negligible heat transfer at these interfaces given their small areas.

## 9.3 Design of a Suspended Resistor

As shown in Fig. 9.1, a multilayer suspended resistor was fabricated on a 500  $\mu$ m thick Si substrate with a 1024 nm thick SiO<sub>2</sub> insulator layer. The resistor consists of a 110 nm thick nickel/gold/nickel layer that was deposited on a low stress 1000 nm thick SiN<sub>x</sub> film that sits on an 80 nm thick germanium sacrificial layer to have strong thermal expansion properties (methods). Then the sample was affixed to a printed circuit board (PCB) with double-sided Scotch tape and electrical connections were made between the sample and the PCB by wire bonding. When a voltage is applied to the PCB, Joule heating in the sample causes its temperature to increase. In order to simply the analysis, we fabricated the suspended bridge with much larger (micrometer-size) lateral dimensions than its vertical dimensions (nanometer-size). Therefore, vertical expansion of the sample is expected to dominate. The material properties used in the simulation is shown in Table 9.1.

	$\begin{array}{c} \text{Thermal} \\ \text{conductivity} \\ (\text{W } \text{m}^{-1}\text{K}^{-1}) \end{array}$	Density (kg m <sup>-3</sup> )	$\begin{array}{c} \text{Heat} \\ \text{capacity} \\ (\text{J } \text{kg}^{-1}\text{K}^{-1}) \end{array}$	Thermal expansion coefficient $(10^{-6} \text{ K}^{-1})$	Young's modulus (GPa)	Poisson's ratio
Au	317	$19.3 \times 10^3$	129	14.2	70	0.46
$\mathrm{Si}_3\mathrm{N}_4$	20	$3.1  imes 10^3$	700	2.3	250	0.26
Ge	58	$5.3  imes 10^3$	310	5.9	103	0.28
$SiO_2$	1.4	$2.2 \times 10^3$	730	0.5	70	0.17
Si	130	$2.3  imes 10^3$	700	2.6	170	0.28
Air	0.024	1.1	101	0.0	103	0.26
(a)	Ni/Au/Ni	(1	D) 500µm	/	(c)	
_	SI3IN4		<b>I I I I I I I I I I</b>		60	

Table 9.1: Material properties in the simulation



Figure 9.1: (a) Side view of the suspended resistor showing the material and device structure. The heights of Ni/Au/Ni, Si<sub>3</sub>N<sub>4</sub>, Ge, SiO<sub>2</sub> and Si are 110 nm, 1000 nm, 80 nm, 1024 nm and 500  $\mu$ m, respectively. (b) Top view of the suspended resistor. The yellow area in the top view indicates the pattern of the Ni/Au/Ni layer and the orange area indicates that the SiO<sub>2</sub> layer is exposed. (c) Zoomed-in view of the rectangular region inscribed in the red circle in (b). This rectangular area represents the field of view for the epi-DPM images.

## 9.4 Results and Discussion

#### 9.4.1 Simulation Results

Figure 9.2 shows the structure in the simulation with zoomed-in views of the suspended bridge and its spatial discretization. The entire simulation domain is  $10 \times 7 \times 0.5 \text{ mm}^3$ . The zoomed-in portion at the suspended bridge has a dimension of  $2 \times 1.5 \times 0.5 \text{ mm}^3$ . As shown in Fig. 9.3, the power dissipated in the sample increases and the temperature rises as the input voltage is increased. The increase in temperature creates a vertical



Figure 9.2: The 3D view of the geometry in the simulation and zoomed-in views showing the suspended resister and its spatial discretization.

deflection of the suspended bridge as is shown in Fig. 9.3(b). Because of the temperature-dependent resistivity of the gold layer, all the aforementioned quantities exhibit nonlinear variations with the input voltages.

The detailed 3-D distribution of the temperature obtained through the coupled simulation under the input of 1 V is provided in Fig. 9.4. Interestingly, because of the geometrical dimensions and material properties, the primary heat removal path is vertical conduction through the 80 nm thick air layer between the suspended resistor bridge and the oxide layer. An additional key heat path is laterally from the bridge to the side contact pads. The wings of the suspended resistor are a third path and explain why the temperature is not maximum at the very center of the bridge. Lateral heat conduction can be identified in Fig. 9.4 (b) by finding places that have a temperature gradient.

The distribution for the vertical component of the displacement field is shown in Fig. 9.5. The bridge deflects upward several nanometers because of the significant heating of the bridge. The oxide layer also heats up due to heat conduction from the bridge above and expands vertically. In the experiment, we will measure the relative height change, i.e., the difference in



Figure 9.3: Simulated variations in the sample parameters versus voltage applied to the PCB include (a) power dissipation in the sample and resistance of the sample, and (b) temperature and relative height change. Note that because the PCB board and bonding wires have 3  $\Omega$  of resistance, the power dissipated in the sample is about half the total power supplied. The voltage across the sample is about half the applied voltage.





Figure 9.4: Simulation of (a) the steady-state temperature profile and (b) its zoomed-in view at the center of the suspended bridge of the resistor under the input voltage of 2 V.



Figure 9.5: Simulation of (a) the distribution of z-displacement profile and (b) its zoomed-in view at the center of the suspended bridge of the resistor under the input voltage of 2 V.

these two deflections.

#### 9.4.2 Experimental Benchmark

We used a conventional epi-DPM setup [120, 122], including a 405 nm single mode fiber (SMF) coupled laser diode (QFLD-405-20S, QPhotonics), a UV transmission grating (300 lines/mm, 8.6° blaze angle), 4f lens system, a 10  $\mu$ m diameter pinhole filter and a charge-coupled device (CCD) camera (Hamamatsu Orca ER C4742-80). The 405 nm laser is temperature controlled and electrically driven by a laser diode controller (ILX Lightwave LDC-3722), which enables a highly stable laser wavelength and low-noise output power. The beam is imaged at the sample plane by a microscope objective (Zeiss Plan-Apochromat 20X, NA 0.8) forming a field of view (FOV) of 81  $\mu$ m X 62  $\mu$ m. The reflected light goes through the UV transmission grating, which separates the imaging field into multi-order copies. Only the first and zeroth orders are picked up by the pinhole filter and by the circular aperture, respectively, in the Fourier plane (FP); these orders served as the reference field and the image field, respectively. Compared with the previous DPM setup [122] which had a visible transmission grating, the diffraction efficiency of the zeroth and first order increases from 4% and 47% to 26% and 65%, respectively. Therefore, the instrument noise per pixel is reduced from 3.0 nm to 2.0 nm. The exposure time is correspondingly reduced because of the higher intensity. Compared to our previous epi-DPM system [120, 122], we exchanged the positions of the pinhole filter and the circular aperture and selected the stable zeroth order as the reference field. The first order then served as the image field. Because the sizes of the lenses are large, spherical aberrations are negligible. The more stable reference signal helped to further reduce the instrument noise per pixel from 2.0 nm to 1.5 nm. The details on the epi-DPM methods and phase retrieval are discussed in the previous works [120, 122, 137].

The PCB with the sample mounted on it was fixed to the xyz translation stage of the DPM system. The resistance of the sample, bonding wires, and PCB was 5.6  $\Omega$ , of which 3  $\Omega$  was attributed to the bonding wires and PCB. Figure 9.1 (b) shows the top view of the sample. The cross pattern was suspended and its two ends were connected to two rectangles (200  $\mu$ m × 250  $\mu$ m). Two bigger square pattern (500  $\mu$ m × 500  $\mu$ m) were used as contact pads for wire bonding. Gold wires were bonded between the sample and the PCB holes and the electric wires were soldered on the PCB holes to make connections to the electrical voltage source. The central bridge region of the resistor is observed by using a 20X objective and the observed FOV is shown in Fig. 9.1 (c).

A background phase image was first collected without an applied voltage and subtracted from each subsequent phase image. This is commonly done with the DPM method to reduce background noise [118, 119, 120, 121, 122, 123, 124, 125] and enables us to obtain an image sequence of the change in phase versus time. To reduce the propagation of phase errors during unwrapping, we temporally unwrapped the image sequence at each pixel [138]. Because this procedure directly results in the phase change at each pixel, spatial unwrapping is not needed. Epi-DPM phase images were collected at a rate of 1.8 frames/s. Phase images were converted to height images by multiplying by  $\lambda/4\pi$ , where  $\lambda$  is the wavelength of the light source. The height change of the central bridge of the suspended resistor, consisting of all six layers, was measured by averaging over the yellow area in Fig. 9.1 (c), and the height change of the bottom two layers, was measured by averaging around four corners, which are the orange area in Fig. 9.1 (c). Their difference,  $\Delta h$ , which is defined as the relative height change, describes the height change of the top four layers. To further reduce temporal noise, these instantaneous relative height change images were averaged for 9 frames. As shown in Fig. 9.6 (a), the relative height change of the suspended resistor is plotted as a function of time. The thermal expansion of the suspended resistor is thereby measurable with high accuracy.

Figure 9.6 (a) shows the average relative height changes for the central bridge of the resistor as a function of time and the applied voltage as a function of time. After application of the 2 V voltage pulse, the sample started to expand and its relative height increased from 140 s to 195 s. At t=140 s, we obtained a relative height change of 1.37 nm, which was defined as the initial relative height change. The predicted value from Fig. 9.3 (b) is 1.65 nm under an input voltage of 2 V. Then the entire sample including the substrate reached thermal steady state between 195 s and 310 s and the relative height change became maximum. After that, the voltage supply was turned off. The sample gradually returned to its initial state and its corresponding relative height change decreased from 315 s to 360 s. Figures 9.6 (b)-(j) present the frames and their relative height change curves at five key moments: before expansion (t=15 s), at the beginning of expansion (t=140 s, i.e. at thermal steady state for the suspended bridge), at the thermal steady state for the substrate (t=195 s), at the beginning of recovery (t=310 s), and after the recovery (t=370 s). From these frames, we find that the thermal expansion process of a suspended resistor can be reconstructed using epi-DPM and the coupled model can predict the initial height change when the suspended bridge is in thermal steady state.

In order to quantify the dependency of thermal expansion on voltage, the relative height change measurement was repeated for various DC voltages. As shown in Fig. 9.7, the steady-state relative height changes are plotted



Figure 9.6: Relative height change measurements throughout the thermal expansion process. (a) Left axis: average relative height changes of a suspended resistor versus time; right axis: applied voltage versus time. (b)-(f) Selected frames and their relative height changes at various times.



Figure 9.6: (cont.)



Figure 9.7: Relative height changes versus input voltage.



Figure 9.8: Simulation and experimental results of the resistance of the sample, bonding wires, and PCB versus time.

as a function of applied voltage. These parameters increase as the applied voltage increases.

The resistance of the sample, bonding wires, and PCB was also measured during the thermal expansion process. Figure 9.8 shows that the resistance increases as a function of time. Three different DC voltage were applied. For t<0 s, we plotted resistance of the sample, bonding wires, and PCB without heating. This value was 5.6  $\Omega$ . At t=0 s, the voltage was applied and the resistance increased instantly. Then the resistance increased slowly and became stable in the last 60 s. The initial jump in resistance is attributed to the heating of the suspended bridge and the slow increase in resistance is attributed to the gradual heat spreading across the large area substrate. The simulation results match well with the experimental results.

#### 9.5 Discussion

We applied epi-DPM to observe the nanoscale thermal expansion of a suspended bridge resistor. We quantified the relative height change with subnanometer accuracy. The results show that even under a low dissipated power of a few hundred milliwatts, there can be several nanometers of relative expansion. In some nanoscale device applications, this expansion can be a source for performance degradation or complete device failure.

We presented a coupled electrical-thermal-mechanical model to understand the underlying physics and be able to predict the 3D temperature profile, the change in resistance, and the relative height expansion under different applied voltages. The simulated and experimental results for both the thermal expansion and the change in resistance showed good agreement. We must point out that choosing appropriate boundary conditions that faithfully represent the experimental conditions are critical for accurate simulation results. For example, although modeling only the zoomed-in region of the sample near the suspended resistor is adequate for capturing the fast dynamics of the initial height jump, this approach fails to capture the slower dynamics caused by substrate heating. Another related example is that a large thermally conductive mass (e.g. the substrate on the PCB) is often approximated by an isothermal boundary. However, the presence of tape between the substrate and the PCB creates an adiabatic thermal boundary condition there because the tape is thick and has very low thermal conductivity. Including an isothermal boundary in the simulation domain can erroneously reduce the time scale for the expansion dynamics by a few orders of magnitude.

With the growing interest in predicting and engineering the coupled dynamics of micro- and nano-scale systems, we anticipate that this model and measurement technique will have many applications. We presented a coupled electrical-thermal-mechanical model to describe the physics of the thermal expansion of a suspended bridge resistor. We applied epi-DPM to observe the nanoscale thermal expansion and quantify the relative height change. Overall, there was good agreement between theory and experiment. We anticipate that this model and measurement technique will be applied to study coupled thermal dynamics in many other micro- and nano-scale systems.

## CHAPTER 10

# MULTIPHYSICS MODELING OF CAPACITOR DERATING IN POWER INTEGRITY SIMULATION

## 10.1 Problem Statement

The design of power distribution networks (PDNs) has become increasingly important in a modern electronic system especially under the trend of system miniaturization and continuous reduction of voltage noise margin. A power distribution network is responsible for delivering clean voltage and current from a voltage regular module (VRM) all the way to transistors through a hierarchy of interconnects. The target impedance [139] based on Ohm's law and calculated from power supply voltage and switching current is considered as one of the most useful quantities in evaluating a PDN. The target impedance of a PDN in a microprocessor-based system often falls into the range of milliOhms. For example, the target impedance of an Itanium processor functioning with 150 W and 1.2 V can be as low as  $1 \text{ m}\Omega$  [28]. Decoupling capacitors play a critical role in a PDN to maintain the low-impedance path from the VRM to the processor over a broadband [139, 28, 140, 141]. In addition, decoupling capacitors can also suppress excessive noise on the power rails. However, as the frequency increases beyond a self-resonant frequency, the series loop inductance of a decoupling capacitor dominates and the decoupling capacitor becomes inductive and hence less effective in maintaining a low impedance. Another problem associated with the decoupling capacitors is the derating issue that the decoupling capacitors often function at lower capacitance than their rated specifications. Derating of capacitors takes place under normal working conditions associated with temperature, voltage, and aging, etc. For example, Fig. 10.1(a) demonstrates that a decoupling capacitor of 1 uF derates as the DC bias increases and Fig. 10.1(b) shows that a decoupling capacitor of 2.2 uF derates when the temperature is raised up from the room temperature of 25 °C. It can also be seen from Fig.

10.1 that the DC bias and temperature can function together and cause the derating issues.

Capacitor derating and its modeling has drawn industry-wide concerns [142]. It is known that in order to speed up the design cycles, advanced modeling and simulation techniques [10, 143] must be employed at the system level. The incorporated models of decoupling capacitors in the systemlevel simulation of a PDN have significant impacts on the simulation results, which is shown in Fig. 10.2. The system-level simulation to calculate PDN impedance takes three types of capacitor models, namely, the ideal model treating each capacitor as a lumped component with its rated capacitance, the derating model based on the operation condition of 1 V and 50  $^{\circ}$ C. and a second derating model based on the operation condition of 1.5 V and 85 °C. As shown in Fig. 10.2, the maximum impedance between the ideal model and the derating model of 1 V and 50  $^{\circ}$ C has a difference of 26% at 10 kHz, whereas at the same frequency the two derating models differ by 16%. The large variance of impedance invoked by capacitor models further demonstrates the importance of incorporating accurate models of decoupling capacitors in the system-level simulation of a PDN.

In this chapter, we propose a simulation methodology that incorporates a library of capacitor deraing models. The capacitor derating library, associated with operation temperature and voltage and based on impedance measurement and curve fitting, can be repeatedly used for PDN simulation and optimization at different design cycles and for different products. Three methods of impedance measurement with individual characterization fixtures and de-embedding techniques are compared and the most accurate one is selected to construct the derating library. Since circuit models are often preferable in the system-level simulation, a curve fitting method is used to convert the measured impedance into SPICE models. The conversion from impedance network parameters to SPICE models is achieved through gradient-based optimization. Even though rated with the same specifications including capacitance, package size, temperature characteristics, and operation voltage, capacitors from vendors may have distinct impedance. Vendor information has also been taken into account in designing the structure of the derating library especially in determining the nominal case. Systemlevel simulations utilizing the capacitor derating library are carried out for calculating the impedance of PDNs in real products. The accuracy of the



Figure 10.1: Derating of (a) a 1 uF capacitor with an increasing DC bias and (b) a 2.2 uF capacitor with the temperature being raised up from the room temperature.


Figure 10.2: Simulating the impedance of a PDN by using ideal and derating models of decoupling capacitors. The ideal model treats the capacitor as a single lumped component with its rated capacitance. The simulation with derating models is proposed in this chapter.

proposed simulation methodology is verified through good correlations with measurement.

## 10.2 Capacitor Derating Mechanism

In many situations, temperature, DC/AC bias, aging, material formulations, thickness actually function together and cause derating. In this section, the derating mechanism of ceramic capacitors associated with temperature, DC bias, and aging is illustrated with details. The three factors are considered as the most influential ones in the capacitor derating mechanism. Barium titanate (BaTiO<sub>3</sub>) is the first known ferroelectric ceramics [141, 144] and is among the most widely accepted materials used for manufacturing ceramic capacitors [145, 146, 147] owing to its high dielectric constant and low loss. Upon cooling down from the Curie point of 130  $^{\circ}$ C, a unit cell of BaTiO<sub>3</sub> undergoes a transition from the paraelectric cubic phase to the ferroelectric tetragonal phase as shown in Fig. 10.3. In the cubic phase of  $BaTiO_3$  illustrated by Fig. 10.3(a), one titanium atom is symmetrically coordinated by six oxygen atoms. When the temperature drops below the Curie point, the caxis stretches and the cubic structure becomes a tetragonal one, in which the titanium atom shifts away from the body center of the unit cell and creates a permanent dipole as indicated in Fig. 10.3(b). The polarization arising from the asymmetry exists without the application of an external electric field, which is called the spontaneous polarization and is the root cause of capacitor derating.

#### 10.2.1 Temperature

Figure 10.4 illustrates the derating mechanism of decoupling capacitors associated with temperature. In the temperature range below the Curie point, the dielectric constant of pure BaTiO<sub>3</sub> increases as the temperature increases. This temperature characteristics is modified from pure BaTiO<sub>3</sub> by the addition of shifters and depressors. The shifters normally pushes the peak value of the dielectric constant to room temperature of around 25 °C, which is the reason why we often see the ceramic capacitors derate as temperature increases. The depressors are responsible to flatten out the dielectric constant over a



Figure 10.3: The unit cell of  $BaTiO_3$  in (a) the cubic paraelectric phase and (b) the tetragonal ferroelectric phase, respectively.

certain temperature range. The addition of depressors achieves the thermal stability but sacrifices the peak value of dielectric constant. Through the addition of shifters and depressors, the ceramics capacitors fall into different categories of thermal characteristics such as NP0, Z5U, and X7R according to the EIA standard [148].

#### 10.2.2 DC Bias

Capacitor derating with DC bias can be understood through the experimental study in [149], in which the dielectric constant of the ferroelectric material barium strontium titanate (BST) is measured under different conditions of temperature and voltage. It can be seen from Fig. 10.5 that by increasing the DC bias, the dielectric constant of BST decreases. As indicated in Fig. 10.3(b), the ferroelectric material contains permanent dipoles. The applied DC bias locks some of the permanent dipoles and prevents them from reorienting with an external AC field, which explains why the ceramic capacitors derate with an increasing DC bias. For ferroelectric materials in the paraelectric state, the relation between the dielectric constant  $\epsilon$  and an applied static electric field of magnitude E can be derived from Devonshire's theory [150] as

$$\epsilon/\epsilon_0 \approx \left[1 + a\epsilon_0^3 E^2\right]^{-1/3} \tag{10.1}$$



Figure 10.4: The temperature characteristics of pure barium titanate is altered by the addition of shifters and depressors.

where a is a constant and  $\epsilon_0$  is the dielectric constant without bias. In Fig. 10.5, an increase of the fraction of strontium in the BST compound shifts the Curie temperature to a lower magnitude, which is exactly how the addition of shifter functions. Strontium titanate (SrTiO<sub>3</sub>) is paraelectric and of very high thermal stability, which is categorized as Class 1 ceramics together with titanium oxide (TiO<sub>2</sub>) and calcium titanate (CaTiO<sub>3</sub>). Class 2 ceramics have significantly higher dielectric constants than Class 1 ceramics and are usually based on BaTiO<sub>3</sub>.

### 10.2.3 Aging

Ceramic capacitors also derate with time. A study in [151] reveals that the aging rate can be as high as 8% per decade-hour for a Y5V capacitor. Upon cooling down from a temperature above the Curie point, a unit cell of  $BaTiO_3$  reverts to a tetragonal structure from a cubic one, during which a strain is developed. The strain relaxes itself over time and the dielectric constant



Figure 10.5: The impact of DC bias on the dielectric constant of barium strontium titanate [149], of which the chemical formula is  $Ba_{1-x}Sr_xTiO_3$  and x is the molar fraction of strontium from 0 to 1.

decays accordingly [152], which causes the derating of ceramic capacitors with time. It is worth mentioning that the soldering process used to mount the ceramic capacitors onto a circuit board provides temperatures above the Curie point for a brief period, which produces a partial reset of the capacitance.

### 10.3 Capacitor Derating Models

The proposed system-level simulation incorporates the derating models of decoupling capacitors, which are built upon impedance measurement and curving fitting method. In this section, detailed information of the impedance measurement and fitting methods is provided.

#### 10.3.1 Impedance Measurement

Three impedance measurement methods of decoupling capacitors are compared, including the method proposed by Intel [153], the approach used by a component manufacturer named vendor A in this work, and the solution



Figure 10.6: Impedance measurement set-up consists of Keysight E5061B network analyzer, RF probes, precision positioners, microscope, cables, and PCB holders.

from a third-party consultant. Three types of capacitors from vendor A are taken as measurement samples, namely, 0201 1 uF, 0402 10 uF, and 0603 22 uF. Through the comparison, we would like to select the most accurate impedance measurement method to construct the capacitor derating library.

The measurement set-up is illustrated in Fig. 10.6, which consists of a Keysight E5061B network analyzer, two RF probes of 0.4 mm pitch and two TP150 precision positioners from PacketMicro [154], a microscope, cables, and PCB holders. The SOLT calibration is performed with CalKit TCS60 [154] shown in Fig. 10.7, which achieves the calibration till the tips of the RF probes.

All three impedance measurement methods under comparison are based on the concept of S-parameter two-port shunt-thru methodology, which is devised for high-frequency PDNs of very low impedance. With the twoport shunt-thru method, the measured impedance within the range of 10% accuracy can be as low as 1 m $\Omega$  up to about 3 GHz [155]. In the two-port shunt-thru method, one port launches current to the DUT and the other port measures the voltage across it. The impedance of the DUT  $Z_{\text{DUT}}$  can

		58 OHM			-	-
-	-			-	-	
					188 DHM	
					188 OHM	
-	-	58 OHM	58 DHM	-	-	
	-			-		

Figure 10.7: CalKit TCS60 from PacketMicro [154].



Figure 10.8: Test fixtures from Intel for impedance measurement of capacitors. The ground pad is connected to the ground trace through vias. In the short fixture, the power and ground pads to mount the component are shorted together.

be derived in terms of the measured  $S_{21}$ 

$$Z_{\rm DUT} = 25 \frac{S_{21}}{1 - S_{21}} \tag{10.2}$$

assuming that the port impedance is 50  $\Omega$  [155, 156].

To perform the impedance measurement, capacitors have to be mounted on some test fixtures. The test fixture of such purpose from Intel is shown in Fig. 10.8. In order to obtain the impedance of the capacitor itself, deembedding has to be performed to remove the effects of traces and vias in the test fixture. The de-embedding of the Intel approach is realized through a short fixture as shown in Fig. 10.8. Two measurements are performed, one for the impedance of the entire structure with capacitor mounted on as  $Z_{\text{tot}}$ and one for the impedance of the short fixture as  $Z_{\text{short}}$ . The impedance of the capacitor itself  $Z_{\text{cap}}$  can thus be obtained through a subtraction

$$Z_{\rm cap} = Z_{\rm tot} - Z_{\rm short} \tag{10.3}$$

It is worth mentioning that if the 3-D geometry of the test fixture for mounting capacitors is available, the de-embedding can be achieved through a fullwave electromagnetic simulation, which saves the effort of fabricating a short fixture. A wave port can be attached at the location where the capacitor is mounted. The obtained S-parameters from the simulation can then be converted to Z-parameters, with which the de-embedding is performed together with the measured  $Z_{\text{tot}}$ .

Figure 10.9 shows the measured impedance versus frequency of the sample capacitors by using the methods from Intel, vendor A, and the third-party consultant. It can be seen from Fig. 10.9(a) that the in-house measurements using Intel and vendor A approaches achieve good agreement. However, there is a large discrepancy of the measured impedance based on the third-party approach from the other two. The large discrepancy is believed to arise from lacking verifications in the de-embedding test fixtures. Due to the large discrepancy, the impedance curve measured using the third-party approach is only shown in Fig. 10.9(a) for the 0603 22 uF capacitor.

In Fig. 10.9, beyond the self-resonant frequency, the measured impedance between Intel and vendor A approaches does not agreement as well as that in the low frequency range. The difference at high frequencies results from the over-subtraction of inductance in the de-embedding procedure of the Intel approach. Note that in Fig. 10.8 there is a piece of metal shorting together the power and ground pads in the short fixture, whereas there is no connection between the power and ground pads in the fixture for mounting the capacitors. As a result, after de-embedding, the inductance associated with the piece of metal shorting together the power and ground pads in the short fixture is over-subtracted. However, in the approach from vendor A, the de-embedding is performed through electromagnetic simulation described earlier and there is no over-subtraction. It explains why beyond the selfresonant frequency the impedance obtained from the vendor A approach is larger than that from the Intel approach. The self-partial inductance



Figure 10.9: Measured impedance with three different approaches of three sample capacitors (a) 0603 22 uF, (b) 0402 10 uF, and (c) 0201 1 uF.



Figure 10.9: (cont.)

associated with the piece of metal in the short fixture can be calculated through

$$L = \frac{\mu_0}{2\pi} \frac{1}{w^2} \left[ lw^2 \ln\left(\frac{l}{w} + \sqrt{\left(\frac{l}{w}\right)^2 + 1}\right) + l^2 w \ln\left(\frac{w}{l} + \sqrt{\left(\frac{w}{l}\right)^2 + 1}\right) + \frac{1}{3} \left(l^3 + w^3\right) - \frac{1}{3} \left(l^2 + w^2\right)^{\frac{3}{2}} \right]$$
(10.4)

in which w and l are the width and length of the rectangular piece of conductor shorting power and ground pads [157] with the assumption that the thickness of the perfect electric conductor is negligible. To compensate the over-subtracted inductance, the calculated inductance from Equation (10.4) should be added backed to the circuit model obtained from the fitting method in the following section.

#### 10.3.2 Curve Fitting Method

Circuit models are often preferred in the system-level simulation of PDNs. Through the curve fitting method illustrated in this section, the measured impedance of capacitors can be converted to SPICE models. There are three steps involved in the curve fitting method: fixing a circuit topology; tuning the major RLC branch for a good initial guess; and building an object function and selecting the most appropriate optimization method. The fixed circuit topology is shown in Fig. 10.10, which consists of a major RLC (in series) branch, five RC (in shunt) sections, four RL (in shunt) sections, and four RLC (in shunt) sections. It is known that the simplest SPICE model of a decoupling capacitor consists of three lumped components: the capacitance, the equivalent series resistance (ESR), and the equivalent series inductance (ESL) [158, 159]. Due to the high-order behavior of the measured impedance around the self-resonant frequency, the simple RLC model is insufficient. However, the tuning of the major RLC branch based on the first-order representation of a decoupling capacitor provides a good initial guess for the optimization. The cost function is defined as the weighted summation of the mean values of the relative changes for both magnitude and phase of the complex impedance, which can be written as

$$f = \frac{1}{N} \sum_{i=1}^{N} \left( \alpha_A \frac{A_i - A_{0,i}}{A_{0,i}} + \alpha_\phi \frac{\phi_i - \phi_{0,i}}{\phi_{0,i}} \right)$$
(10.5)

where A and  $\phi$  represent the magnitude and phase of the impedance, respectively, and  $\alpha_A$  and  $\alpha_{\phi}$  are the corresponding weights in the summation. Gradient-based optimization technique from Keysight Advanced Design System (ADS) is used. The fitting results are shown in Fig. 10.11 for three sample capacitors from vendor A. The same topology shown in Fig. 10.10 is used in converting the measured impedance of the three capacitors into SPICE models. However, it can be seen from Fig. 10.11 that the capacitor of larger value requires more branches than smaller ones in order to achieve higher accuracy at lower frequencies.



Figure 10.10: Fixed circuit topology used for curve fitting.

# 10.4 System-Level Simulation and Measurement Correlation

Capacitors employed in a PDN are often provided by various vendors. Even though rated with the same capacitance, package size, temperature coefficient, capacitors from different vendors are most likely to have distinct impedance. As shown in Fig. 10.12, for the three capacitors from vendor A, B, and C with the same rated specifications, the measured impedance curves are quite different from each other especially the ESR and ESL. We have observed that among the three vendors, capacitors from vendor A have nominal values in terms of capacitance, ESR, and ESL. Therefore, the derating library based on capacitors from vendor A is chosen as the nominal case in the system-level simulation. The library is built upon individual operation condition including voltage and temperature. The library file can be repeatedly used by the power integrity simulation and optimization for different products and at different design cycles.

The derating library of capacitors is incorporated into Cadence tools for power integrity simulation. Measurement correlations are performed to verify the accuracy of the proposed simulation methodology with the constructed derating library. Faraday chokes shown in Fig. 10.13 are attached at the cables to suppress the floating current in order to achieve good accuracy at low frequencies. It can be seen from Fig. 10.14 that the simulated impedance



(a)



Figure 10.11: Fitting results including both the magnitude and phase for three different capacitors including (a) 0201 1 uF, (b) 0402 10 uF, and (c) 0603 22 uF.



Figure 10.11: (cont.)

with the derating library agrees very well with the measurement. In Fig. 10.14(c), at 10 kHz, the relative difference between the measured and simulated impedance with derating library is about 7%, whereas it is as large as 33% between the measured the simulated impedance with ideal capacitor models. Similarly, for the power domain in Fig. 10.14(d), the relative difference between the measurement and the simulation with the derating library is 3.4%, whereas it is around 34% between the measurement and the simulation with ideal capacitor models. The measurement correlations not only demonstrate the high accuracy of the proposed simulation methodology, but also prove the necessity and the importance of taking into account capacitor derating in the system-level simulation.

### 10.5 Conclusion

In this chapter, we propose the system-level power integrity simulations that incorporates the derating models of decoupling capacitors. The derating models of decoupling capacitors are constructed through the impedance mea-



Figure 10.12: Measured impedance of 0402 10 uF capacitors from three vendors.

surement and curve fitting method. The impedance measurement is based on the concept of two-port shunt-thru method considering the very low impedance of the high-frequency PDNs under investigation. Three methods of impedance measurement from Intel, vendor A, and a third-party consultant are compared and the most accurate one from vendor A is taken to perform the impedance measurement. Since circuit models are often preferred in the system-level simulation, a curve fitting method is employed to convert the measured impedance into circuit models. The fitting method is based on gradient-based optimization with the cost function built with both magnitude and phase of the measured impedance. The library file in the Cadence \*.amm format is constructed by taking into account the operation temperature and voltage. Information of capacitor vendors is also considered in building the derating library and capacitors from vendor A are chosen as the nominal case. The accuracy of the proposed simulation methodology is demonstrated through good correlations with measurement performed on the real products. The good agreement between the measurement and the simulation with derating library again demonstrates the importance of considering derating effects of decoupling capacitors in the power integrity simulation.



Figure 10.13: Faraday Chokes have been applied in the PDN measurement.



Figure 10.14: Measurement correlations of the proposed simulation methodology on six different power domains.



(c)



Figure 10.14: (cont.)



(e)



Figure 10.14: (cont.)

# CHAPTER 11

# CONCLUSION AND FUTURE WORK

### 11.1 Conclusion

In this dissertation, a multiphysics simulation is proposed to address the ever increasing couplings between electrical designs and thermal issues in electronic systems. The multiphysics simulation is implemented with the finite element method on full three dimension. The 3-D simulation achieves adequate accuracy in modeling the escalating complexities in the modern electronics in terms of both complicate geometries and non-uniform materials. In the multiphysics simulation, an electrical analysis and a thermal analysis are first integrated through an iterative scheme through temperature-dependent material properties. The electrical aspect of the multiphysics simulation consists of static and dynamic IR-drop analyses and full-wave electromagnetic analysis. The thermal aspect of the multiphysics simulation consists of both steady-state and transient thermal analyses. The particular type of and the combination of the electrical and thermal analyses depend on specific applications. The multiphysics simulation is further extended through incorporating fluid dynamics and structure mechanics to handle a broader range of problems. For example, it enables the analysis of an electronic system with integrated microchannel cooling. It also takes into account thermal stress such that the reliability concerns of interconnects arising from mechanical fatigue and failure can be addressed. The multiphysics simulation is also applied to characterized the thermal-mechanical systems such as MEMS and NEMS. Finally we construct multiphysics models of decoupling capacitors and incorporate them into the characterization of power distribution networks.

The capability of the multiphysics simulation is demonstrated through its applications in a variety of important problems, including the static and dynamic IR-drop analyses of power distribution networks, the thermalware high-frequency characterization of through-silicon-via structures, the full-wave electromagnetic analysis of high-power RF/microwave circuits, the modeling and analysis of three-dimensional ICs with integrated microchannel cooling, the characterization of micro- and nanoscale electrical-mechanical systems, and the modeling of decoupling capacitor derating in the power integrity simulations.

In the DC-IR drop analysis, the Joule heating effect is taken into account for an accurate prediction of voltage distribution and localized overheating in the integrated circuits. The transient electrical-thermal co-simulation devised for the characterization of PDNs not only detects the localized overheating but also depicts the large temperature gradient induced by dynamic variance in the supply voltage. It also characterizes the voltage drops with thermal effects accounted to insure the fluctuations are within design specifications. The impacts of different types of voltage pulses, power maps, and via pitches on the design of a PDN are also investigated.

Two types of TSV structures are examined based on the developed thermalaware full-wave electromagnetic analysis, namely, TSV arrays in the silicon interposer and TSV daisy chains. The signal transmission and couplings in the TSV structures are characterized by the insertion loss, return loss, and near- and far-end crosstalks. The impacts on the signal transmission and couplings exerted by the via-filling metal, the doping concentration of the silicon substrate, the via pitch, and the cooling efficiency are also investigated and discussed based on the two types of TSV structures. For the co-simulation of high-power RF/microwave circuits, the thermal impact on the characteristics of a matching network in high-power RF amplifiers is observed. The changes of the passing band of a SIW filter brought by high temperatures associated with high input power are also captured. As the simulated RF/microwave designs and many others are highly subject to temperature variance, the proposed multiphysics simulation provides an accurate approach for analyzing their electrical behaviors by including thermal impacts at the same time.

Through the multiphysics simulation on electronic systems with microchannels, the effectiveness of the integrated microchannel cooling in heat removal from both adjacent active device layers and localized over-heating are observed. At the same time, the thermal impacts on the high-frequency electromagnetic characteristics of the design have also been shown with the numerical example.

Reliability analysis of large-scale solder bump and bonding wire arrays is carried out the multiphysics simulation. Both types of structures are not only major forms of interconnects but also are attracting intensive concerns of reliability. The multiphysics simulation is able to provide detailed distributions of temperature, deformations, and stresses in a highly efficiently manner. Concentrated stresses of large amplitude associated several failure mechanisms have been successfully predicted by the proposed simulation. The multiphysics simulation successfully predicts the thermal expansion of a suspended bridge resistor, which agrees very well with the measurement correlation achieved through epi-DPM. The good agreement makes the multiphysics simulation a promising scheme in predicting and engineering the coupled thermal dynamics in many other micro- and nanoscale systems.

The aforementioned capability and the various applications is made possible through the efficient large-scale analysis in the multiphysics simulation. FETI, a domain decomposition method, is incorporated into the multiphysics simulation in order to handle large-scale problems efficiently. With FETI, the original large-scale problem is decomposed into smaller ones, which allows the possibility of parallel computing with multiple processors. By introducing Lagrange multipliers, FETI transforms the original problem into a reduced-order interface problem. The implementation of FETI consists of three steps: tearing, interconnecting, and recovering unknowns. All the three steps can be performed almost independently within individual subdomains, which makes FETI highly parallelizable in nature. The FETI-enabled efficient multiphysics simulation has demonstrated high parallel efficiency and the capability of significantly reducing the computation time by using multiple processors in parallel. The adaptive time-stepping incorporated into the multiphysics simulation is based on the algorithm of PID control. It improves the overall efficiency of the transient analysis by enabling the use of non-uniform steps for time marching and automatic detection of a steady state.



Figure 11.1: Multiphysics simulation in the design flow. Physics A, B, and C are represented with disks of different sizes to indicate the multiscale.

# 11.2 Future Work

From Fig. 11.1, the future work built upon the proposed multiphysics modeling and simulation can be classified into three categories: refinement on an individual event in the current multiphysics simulation; improvement through the coupling among the different events of different scales; and advancement from multiphysics simulation to sensitivity analysis, design of experiment, and optimization.

# 11.3 Refinement on a Single Event

Refinement on a single event/physics is considered as the most straightforward extension to this dissertation. Four examples in the following are used to illustrate how future work can be extended in this category.

The full-wave electromagnetic analysis in the multiphysics simulation is based on the assumption of electromagnetic steady sate. This assumption is valid in the situations when the time constants of electrical and thermal responses differ by a few orders of magnitude. However, there are situations when the time constant of the thermal response is comparable to that of the electrical response, for example, when a large transient electromagnetic pulse is injected into the system due to electrostatic discharge stress. The compatible time scale between electrical and thermal responses also exists in the applications of phase change memory (PCM) where the heating-up and cooling-down of a PCM device is in nano-second scale to allow fast reading and writing. As future work, the multiphysics simulation should include the transient full-wave electromagnetic analysis to handle the aforementioned issued and extend its capability.

In the current multiphysics simulation, steady-state thermal stress is calculated. The structure mechanics part can be further extended into the transient regime such that the process of crack propagation and wire debonding can be captured. It will greatly assist the design process if the crack propagation and wire debonding can be visualized in situ with the multiphysics simulation.

In this dissertation, the dissipated electrical energy is taken as heat sources. However, the heat energy may turn into electromagnetic waves in the form of thermal radiation or thermal noise. Thermal noise characterization in an electronic system becomes increasingly important. One possible approach of characterizing thermal noise is applying fluctuation-dissipation theorem.

Applying the current electrical-thermal co-simulation to characterize optical interconnects (OIs) is considered as another future work. The optical interconnects on silicon have been advocated as promising solutions to overcome the interconnect scaling problems [160, 161]. OIs do not have the resistive loss from the electrical wirings, which has been the driving force of introducing optics for system level or even on-chip level interconnections [161]. The other advantages of OIs over electrical interconnects include the high packing density, the low power consumption, and the enhanced signal integrity and timing [161]. A key component of OIs is the modulator at the transmitter end and the microring resonator based modulator has been an attractive technology with full CMOS compatibility. For the integrated microring resonators, the temperature increase due to the power loss alters the material refractive index and shifts the resonant wavelength. The thermally-induced resonance shift and the nonlinearity have been observed through experiments and reported in [162, 163, 164, 165]. A numerical scheme [165] based on the coupled-mode theory for optics and a lumped capacitance model for heat conduction is proposed to predict the self-heating dynamics in microring resonators. It is known that the temperature variance in the resonator is proportional to the input optical power and a bistable relations between input and output power can be observed for an resonator with intensity-dependent refractive index. As another future work, the multiphysics simulation will find its application in the design and analysis of optical interconnects, especially for the prediction of thermally induced resonance shift and bistability.

# 11.4 Improvement through the Couplings of Multiscale Events

Multiphysics and multiscale are tightly coupled. Improvement through the couplings of the multiscale events is considered as an important category of future work. A electronic system does not necessarily consists of similar structures but components of very different scales. For example, a 3-D integrated system may contain multiple dies, interposer, multilayer package, and board, which are further connected through vias, bumps, and traces. The dimension of a TSV in the 3-D system is often in the micrometer range, whereas the dimension of a package or board is in terms of centimeters, which leads to a scale ratio of at least  $10^4$ . As the heat source is often on the dies, the thermal analysis itself of the entire system is highly multiscale, not mentioning a multiphysics simulation. Another example is on the characterization of electromagnetic interference and compatibility (EMI/EMC) of electronic systems. A mobile device often contains multiple chips, antennas, and sensors on a single board. Due to the constraint of space, the multiple modules with fine features are often placed very close to each other. The EMI/EMC characterization thus requires considering the multiple components of different feature sizes simultaneously. The aforementioned two problems demonstrate the multiscale in space, whereas multiscale can also occur in time. For example, drift and diffusion occurring at molecular or atomic level in nanoscale device, material defects, and fluid flows may be coupled with the diffusion of heat or propagation of electromagnetic waves at macroscale, where the two events coupled together have very different time scales. Decomposing the multiscale system into individual events may be a good starting point in handling the multiscale problem, but stitching the individual models back to the original event governed by the same physical laws requires mathematical techniques and high performance computing algorithms. Meshing generation in the multiscale problem is another challenge and hence the possibility of future work belonging to this category.

## 11.5 Advancement from Simulation to Design

The simulation with a uniphysics is to reduce the cost and cycles for designers, whereas the multiphysics simulation itself can be considered as a system-level optimization governed by at least two different physical laws. If we are further aware how a multiphysics simulation should fit in the design flow shown in Fig. 11.1, we should regard the ambitious advancement from multiphysics simulation to design as another important category of future work. The advancement from simulation to design should include sensitivity analysis, uncertainty characterization, design of experiment, and optimization, which requires many repeated forward simulations. Through these advancements, new variables are defined in addition to the physical quantities involved in the governing equations describing the physical laws. These newly-defined variables may contain probability density functions, sensitivity gradients, and Lagrange multipliers. Computational complexity is one of the major bottlenecks in these advancements when many forward simulations are involved. The computational burden shall be relieved through novel mathematical techniques and high performance computing. Reduced-order modeling is one of the important techniques in resolving the bottleneck.

# REFERENCES

- N. Magen, A. Kolodny, U. Weiser, and N. Shamir, "Interconnect-power dissipation in a microprocessor," in *Proceedings of the 2004 International Workshop on System Level Interconnect Prediction*. ACM, 2004, pp. 7–13.
- [2] D. A. B. Miller, "Device requirements for optical interconnects to silicon chips," *Proceedings of the IEEE*, vol. 97, no. 7, pp. 1166–1185, 2009.
- [3] E.-P. Li, Electrical Modeling and Design for 3D Integration: 3D Integrated Circuits and Packaging, Signal Integrity, Power Integrity and EMC. Hoboken, NJ: Wiley-IEEE Press, 2012.
- [4] P. Garrou, C. Bower, and P. Ramm, Handbook of 3D Integration: Technology and Applications of 3D Integrated Circuits. New York: Wiley-VCH, 2008.
- [5] P. Ramm, A. Klumpp, J. Weber, N. Lietaer, M. Taklo, W. D. Raedt, T. Fritzsch, and P. Couderc, "3D integration technology: Status and application development," in *ESSCIRC*, 2010 Proceedings of the, 2010, pp. 9–16.
- [6] K. Banerjee, S. J. Souri, P. Kapur, and K. C. Saraswat, "3-D ICs: A novel chip design for improving deep-submicrometer interconnect performance and systems-on-chip integration," *Proceedings of the IEEE*, vol. 89, no. 5, pp. 602–633, 2001.
- [7] A. Chakraborty, K. Duraisami, A. Sathanur, P. Sithambaram, L. Benini, A. Macii, E. Macii, and M. Poncino, "Dynamic thermal clock skew compensation using tunable delay buffers," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 16, no. 6, pp. 639–649, 2008.
- [8] E. Pop, "Energy dissipation and transport in nanoscale devices," Nano Research, vol. 3, no. 3, pp. 147–169, 2010.

- [9] D. L. Brown, J. Bell, D. Estep, W. Gropp, B. Hendrickson, S. Keller-McNulty, D. Keyes, J. T. Oden, L. Petzold, and M. Wright, "Applied mathematics at the US department of energy: Past, present and a view to the future," Lawrence Livermore National Laboratory (LLNL), Livermore, CA, Tech. Rep., 2008.
- [10] J. M. Jin, The Finite Element Method in Electromagnetics, 3rd edition. Hoboken, NJ: Wiley, 2014.
- [11] J.-M. Koo, S. Im, L. Jiang, and K. E. Goodson, "Integrated microchannel cooling for three-dimensional electronic circuit architectures," *Journal of Heat Transfer*, vol. 127, no. 1, pp. 49–58, 2005.
- [12] D. Sekar, C. King, B. Dang, T. Spencer, H. Thacker, P. Joseph, M. Bakir, and J. Meindl, "A 3D-IC technology with integrated microchannel cooling," in *International Interconnect Technology Conference*, 2008, pp. 13–15.
- [13] J. Lau, *Thermal Stress and Strain in Microelectronics Packaging*. Springer Science & Business Media, 2012.
- [14] C. Farhat and F.-X. Roux, "A method of finite element tearing and interconnecting and its parallel solution algorithm," *International Jour*nal for Numerical Methods in Engineering, vol. 32, no. 6, 1991.
- [15] G. L. Sleijpen and D. R. Fokkema, "Bicgstab (l) for linear equations involving unsymmetric matrices with complex spectrum," *Electronic Transactions on Numerical Analysis*, vol. 1, no. 11, p. 2000, 1993.
- [16] D. J. Rixen and C. Farhat, "A simple and efficient extension of a class of substructure based preconditioners to heterogeneous structural mechanics problems," *International Journal for Numerical Methods in Engineering*, vol. 44, no. 4, pp. 489–516, 1999.
- [17] Y.-J. Li and J.-M. Jin, "Parallel implementation of the FETI-DPEM algorithm for general 3D EM simulations," *Journal of Computational Physics*, vol. 228, no. 9, pp. 3255–3267, May 2009.
- [18] C. W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations. Prentice Hall PTR, 1971.
- [19] E. Haier, S. Norsett, and G. Wanner, "Solving ordinary differential equations I, nonstiff problems," *Section III*, vol. 8, 1993.
- [20] K. Gustafsson, "Control theoretic techniques for stepsize selection in explicit Runge-Kutta methods," ACM Transactions on Mathematical Software (TOMS), vol. 17, no. 4, pp. 533–554, 1991.

- [21] K. Gustafsson, M. Lundh, and G. Söderlind, "A pi stepsize control for the numerical solution of ordinary differential equations," *BIT Numerical Mathematics*, vol. 28, no. 2, pp. 270–287, 1988.
- [22] L. Dubcova, P. Solin, J. Cerveny, and P. Kus, "Space and time adaptive two-mesh hp-finite element method for transient microwave heating problems," *Electromagnetics*, vol. 30, no. 1-2, pp. 23–40, 2010.
- [23] A. Valli, G. Carey, and A. Coutinho, "Control strategies for timestep selection in simulation of coupled viscous flow and heat transfer," *Communications in Numerical Methods in Engineering*, vol. 18, no. 2, pp. 131–139, 2002.
- [24] M. Swaminathan and E. Engin, Power Integrity Modeling and Design for Semiconductors and Systems. Pearson Education, 2007.
- [25] S. Lin and N. Chang, "Challenges in power-ground integrity," in Computer Aided Design, 2001. ICCAD 2001. IEEE/ACM International Conference on, 2001, iD: 1. pp. 651–654.
- [26] "International technology roadmap for semiconductors (ITRS) 2011."
  [Online]. Available: http://www.itrs2.net/2011-itrs.html
- [27] M. K. Gowan, L. L. Biro, and D. B. Jackson, "Power considerations in the design of the Alpha 21264 microprocessor," in *Design Automation Conference*, 1998. Proceedings, 1998, pp. 726–731.
- [28] M. Swaminathan, J. Kim, I. Novak, and J. P. Libous, "Power distribution networks for system-on-package: Status and challenges," *Advanced Packaging, IEEE Transactions on*, vol. 27, no. 2, pp. 286–300, 2004.
- [29] K. Banerjee and A. Mehrotra, "Global (interconnect) warming," Circuits and Devices Magazine, IEEE, vol. 17, no. 5, pp. 16–32, 2001.
- [30] M. Pedram and S. Nazarian, "Thermal modeling, analysis, and management in VLSI circuits: Principles and methods," *Proceedings of the IEEE*, vol. 94, no. 8, pp. 1487–1501, 2006.
- [31] J. Xie and M. Swaminathan, "Electrical-thermal co-simulation of 3D integrated systems with micro-fluidic cooling and Joule heating effects," *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 1, no. 2, pp. 234–246, Feb 2011.
- [32] D. Klokotov, J. Schutt-Aine, W. Beyene, D. Mullen, M. Li, R. Schmitt, and L. Yang, "Application of the latency insertion method to electrothermal circuit analysis," in *Electrical Performance of Electronic Pack*aging and Systems (EPEPS), 2011 IEEE 20th Conference on, 2011, pp. 263–266.

- [33] Y. Zhong and M. D. F. Wong, "Thermal-aware IR drop analysis in large power grid," in *Quality Electronic Design*, San Jose, CA, 2008, pp. 194–199.
- [34] C. Farhat and J. Mandel, "The two-level FETI method for static and dynamic plate problems part i: An optimal iterative solver for biharmonic systems," *Computer Methods in Applied Mechanics and Engineering*, vol. 155, no. 1, pp. 129–151, 1998.
- [35] Y. Li and J.-M. Jin, "A vector dual-primal finite element tearing and interconnecting method for solving 3-D large-scale electromagnetic problems," Antennas and Propagation, IEEE Transactions on, vol. 54, no. 10, pp. 3000–3009, 2006.
- [36] M.-F. Xue and J.-M. Jin, "Nonconformal FETI-DP methods for largescale electromagnetic simulation," Antennas and Propagation, IEEE Transactions on, vol. 60, no. 9, pp. 4291–4305, 2012.
- [37] W. Yao, J.-M. Jin, and P. T. Krein, "A dual-primal finite-element tearing and interconnecting method combined with tree-cotree splitting for modeling electromechanical devices," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, vol. 26, no. 2, pp. 151–163, 2013.
- [38] Q. K. Zhu, *Power Distribution Network Design for VLSI*. Hoboken, NJ: Wiley-Interscience, 2004.
- [39] G. Huang, M. Bakir, A. Naeemi, H. Chen, and J. D. Meindl, "Power delivery for 3D chip stacks: Physical modeling and design implication," in *Electrical Performance of Electronic Packaging*, 2007 IEEE, 2007, pp. 205–208.
- [40] Z. Xu, Q. Wu, H. He, and J. J. Q. Lu, "Electromagnetic-simulation program with integrated circuit emphasis modeling, analysis, and design of 3-D power delivery," *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 3, no. 4, pp. 641–652, 2013.
- [41] I. J. Chung and A. C. Cangellaris, "Hierarchical modeling and computationally-efficient transient simulation of on-chip power grid," in *Electronic Components and Technology Conference*, 2007. ECTC '07. Proceedings. 57th, 2007, pp. 1632–1637.
- [42] E. C. C. Yeh, W. J. Choi, K. N. Tu, P. Elenius, and H. Balkan, "Current-crowding-induced electromigration failure in flip chip solder joints," *Applied Physics Letters*, vol. 80, no. 4, pp. 580–582, 2002.

- [43] T. Lu and J.-M. Jin, "Electrical-thermal co-simulation for DC IR-drop analysis of large-scale power delivery," *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 4, no. 2, pp. 323–331, Feb 2014.
- [44] S.-H. Lee, K. Mao, and J.-M. Jin, "A complete finite-element analysis of multilayer anisotropic transmission lines from DC to terahertz frequencies," *Advanced Packaging, IEEE Transactions on*, vol. 31, no. 2, pp. 326–338, 2008.
- [45] X. Li and J.-M. Jin, "A comparative study of three finite element-based explicit numerical schemes for solving Maxwell's equations," *Antennas* and Propagation, IEEE Transactions on, vol. 60, no. 3, pp. 1450–1457, 2012.
- [46] X.-P. Wang, W.-Y. Yin, and S. He, "Multiphysics characterization of transient electrothermomechanical responses of through-silicon vias applied with a periodic voltage pulse," *Electron Devices, IEEE Transactions on*, vol. 57, no. 6, pp. 1382–1389, 2010.
- [47] Z. Ren, W.-Y. Yin, Y.-B. Shi, and Q.-H. Liu, "Thermal accumulation effects on the transient temperature responses in LDMOSFETs under the impact of a periodic electromagnetic pulse," *Electron Devices*, *IEEE Transactions on*, vol. 57, no. 1, pp. 345–352, 2010.
- [48] V. Kosel, R. Sleik, and M. Glavanovics, "Transient non-linear thermal FEM simulation of smart power switches and verification by measurements," in *Thermal Investigation of ICs and Systems, 2007. THER-MINIC 2007. 13th International Workshop on*, 2007, pp. 110–114.
- [49] C. Multiphysics, "COMSOL multiphysics user guide (version 4.3 a)," COMSOL, AB, 2012.
- [50] J. Cho, E. Song, K. Yoon, J. S. Pak, J. Kim, W. Lee, T. Song, K. Kim, J. Lee, H. Lee, K. Park, S. Yang, M. Suh, K. Byun, and J. Kim, "Modeling and analysis of through-silicon via (TSV) noise coupling and suppression using a guard ring," *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 1, no. 2, pp. 220–233, 2011.
- [51] H. Wang and J. Fan, "Modeling local via structures using innovative PEEC formulation based on cavity Green's functions with wave port excitation," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 61, no. 5, pp. 1748–1757, 2013.
- [52] S. Wu and J. Fan, "Analytical prediction of crosstalk among vias in multilayer printed circuit boards," *Electromagnetic Compatibility*, *IEEE Transactions on*, vol. 54, no. 2, pp. 413–420, 2012.

- [53] T. Harris, S. Priyadarshi, S. Melamed, C. Ortega, R. Manohar, S. Dooley, N. Kriplani, W. Davis, P. Franzon, and M. Steer, "A transient electrothermal analysis of three-dimensional integrated circuits," *Components, Packaging and Manufacturing Technology, IEEE Transactions* on, vol. 2, no. 4, pp. 660–667, April 2012.
- [54] S. Priyadarshi, T. Harris, S. Melamed, C. Otero, N. Kriplani, C. Christoffersen, R. Manohar, S. Dooley, W. Davis, P. Franzon, and M. Steer, "Dynamic electrothermal simulation of three-dimensional integrated circuits using standard cell macromodels," *Circuits, Devices Systems, IET*, vol. 6, no. 1, pp. 35–44, January 2012.
- [55] E.-P. Li, X.-C. Wei, A. C. Cangellaris, E.-X. Liu, Y.-J. Zhang, M. D'Amore, J. Kim, and T. Sudo, "Progress review of electromagnetic compatibility analysis technologies for packages, printed circuit boards, and novel interconnects," *Electromagnetic Compatibility, IEEE Transactions on*, vol. 52, no. 2, pp. 248–265, 2010.
- [56] C. Bermond, L. Cadix, A. Farcy, T. Lacrevaz, P. Leduc, and B. Flechet, "High frequency characterization and modeling of high density TSV in 3D integrated circuits," in *Signal Propagation on Interconnects*, 2009. SPI '09. IEEE Workshop on, 2009, pp. 1–4.
- [57] K. J. Han, M. Swaminathan, and T. Bandyopadhyay, "Electromagnetic modeling of through-silicon via (TSV) interconnections using cylindrical modal basis functions," *Advanced Packaging, IEEE Transactions* on, vol. 33, no. 4, pp. 804–817, 2010.
- [58] E.-X. Liu, E.-P. Li, Z.-Z. Oo, X. Wei, Y. Zhang, and R. Vahldieck, "Novel methods for modeling of multiple vias in multilayered parallelplate structures," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 57, no. 7, pp. 1724–1733, 2009.
- [59] B. Wu, X. Gu, L. Tsang, and M. B. Ritter, "Electromagnetic modeling of massively coupled through silicon vias for 3D interconnects," *Microwave and Optical Technology Letters*, vol. 53, no. 6, pp. 1204–1206, 2011.
- [60] G. V. der Plas, P. Limaye, I. Loi, A. Mercha, H. Oprins, C. Torregiani, S. Thijs, D. Linten, M. Stucchi, and G. Katti, "Design issues and considerations for low-cost 3-D TSV IC technology," *Solid-State Circuits*, *IEEE Journal of*, vol. 46, no. 1, pp. 293–307, 2011.
- [61] M. S. Tyagi, Introduction to Semiconductor Materials and Devices. New York: Wiley, 1991.

- [62] W.-S. Zhao, X.-P. Wang, and W.-Y. Yin, "Electrothermal effects in high density through silicon via (TSV) arrays," *Progress In Electro*magnetics Research, vol. 115, pp. 223–242, 2011.
- [63] W.-Y. Yin, K. Kang, and J.-F. Mao, "Electromagnetic-thermal characterization of on on-chip coupled (a)symmetrical interconnects," Advanced Packaging, IEEE Transactions on, vol. 30, no. 4, pp. 851–863, 2007.
- [64] G. Katti, M. Stucchi, K. D. Meyer, and W. Dehaene, "Electrical modeling and characterization of through silicon via for three-dimensional ics," *Electron Devices, IEEE Transactions on*, vol. 57, no. 1, pp. 256– 262, 2010.
- [65] D. M. Jang, C. Ryu, K.-Y. Lee, B.-H. Cho, J. Kim, T. S. Oh, W. J. Lee, and J. Yu, "Development and evaluation of 3-D SiP with vertically interconnected through silicon vias (TSV)," in *Electronic Components and Technology Conference*, 2007. ECTC '07. Proceedings. 57th, 2007, pp. 847–852.
- [66] P. Aaen, J. A. Plá, and J. Wood, Modeling and Characterization of RF and Microwave Power FETs. Cambridge University Press, 2007.
- [67] M. Celuch and P. Kopyt, "Modeling of microwave heating of foods," in M.W. Lorence and P.S. Pesheck, Eds., Development of Packaging and Products for Use in Microwave Ovens. Oxford: Woodhead Publishing, 2009, pp. 305–348.
- [68] P. Kopyt and M. Celuch, "Coupled electromagnetic-thermodynamic simulations of microwave heating problems using the FDTD algorithm," *Journal of Microwave Power and Electromagnetic Energy*, vol. 41, no. 4, pp. 18–29, 2007.
- [69] V. V. Yakovlev, S. M. Allan, M. L. Fall, and H. S. Shulman, "Computational study of thermal runaway in microwave processing of zirconia," *Microwave and RF Power Applications, J. Tao, Ed., Cépaduès Éditions*, pp. 303–306, 2011.
- [70] T. Lu and J.-M. Jin, "Transient electrical-thermal analysis of 3-D power distribution network with FETI-enabled parallel computing," *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 4, no. 10, pp. 1684–1695, 2014.
- [71] K. Kagawa, "High-frequency high-power transistor," U.S. Patent 5,371,405, Dec. 6 1994.
- [72] C. Blair, T. Ballard, and J. Curtis, "Output matched LDMOS power transistor device," U.S. Patent 6,177,834, Jan. 23 2001.

- [73] P. H. Aaen, J. A. Pla, and C. A. Balanis, "On the development of CAD techniques suitable for the design of high-power RF transistors," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 53, no. 10, pp. 3067–3074, 2005.
- [74] D. Deslandes and K. Wu, "Single-substrate integration technique of planar circuits and waveguide filters," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 51, no. 2, pp. 593–596, 2003.
- [75] Y. J. Cheng, K. Wu, and W. Hong, "Power handling capability of substrate integrated waveguide interconnects and related transmission line systems," *Advanced Packaging, IEEE Transactions on*, vol. 31, no. 4, pp. 900–909, 2008.
- [76] Y. Touloukian, R. Powell, C. Ho, and P. Klemens, *Thermophysical Properties of Matter*. New York: IFI/Plenum, 1970.
- [77] P. Auerkari, Mechanical and Physical Properties of Engineering Alumina Ceramics. Finland: Technical Research Centre of Finland, 1996.
- [78] G. G. Harman, Wire Bonding in Microelectronics. McGraw-Hill, 2010.
- [79] J. Lim, D. Kwon, J.-S. Rieh, S.-W. Kim, and S. Hwang, "RF characterization and modeling of various wire bond transitions," *Advanced Packaging, IEEE Transactions on*, vol. 28, no. 4, pp. 772–778, 2005.
- [80] Z.-C. Hao, W. Hong, J. X. Chen, X. P. Chen, and K. Wu, "Planar diplexer for microwave integrated circuits," *IEE Proceedings-Microwaves, Antennas and Propagation*, vol. 152, no. 6, pp. 455–459, 2005.
- [81] W. R. Davis, J. Wilson, S. Mick, J. Xu, H. Hua, C. Mineo, A. M. Sule, M. Steer, and P. D. Franzon, "Demystifying 3D ICs: The pros and cons of going vertical," *Design and Test of Computers, IEEE*, vol. 22, no. 6, pp. 498–510, 2005.
- [82] Y. Xie, G. H. Loh, B. Black, and K. Bernstein, "Design space exploration for 3D architectures," ACM Journal on Emerging Technologies in Computing Systems (JETC), vol. 2, no. 2, pp. 65–103, 2006.
- [83] K. Tu, "Reliability challenges in 3D IC packaging technology," Microelectronics Reliability, vol. 51, no. 3, pp. 517–523, 2011.
- [84] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S. Velusam, "Compact thermal modeling for temperature-aware design," in *Proceedings of the 41st Annual Design Automation Conference.* ACM, 2004, pp. 878–883.

- [85] X.-P. Wang, W.-Y. Yin, and S. He, "Multiphysics characterization of transient electrothermomechanical responses of through-silicon vias applied with a periodic voltage pulse," *Electron Devices, IEEE Transactions on*, vol. 57, no. 6, pp. 1382–1389, 2010.
- [86] T. Lu and J.-M. Jin, "Thermal-aware high-frequency characterization of large-scale through-silicon-via structures," *Components, Packaging* and Manufacturing Technology, IEEE Transactions on, vol. 4, no. 6, pp. 1015–1025, June 2014.
- [87] W. Chen, W.-Y. Yin, E. Li, M. Cheng, and J. Guo, "Electrothermal investigation on vertically aligned single-walled carbon nanotube contacted phase-change memory array for 3-D ICs," *Electron Devices*, *IEEE Transactions on*, vol. 62, no. 10, pp. 3258–3263, 2015.
- [88] T. Lu and J.-M. Jin, "Electrical-thermal co-simulation for analysis of high-power RF/microwave components," *IEEE Transactions on Elec*tromagnetic Compatibility, vol. PP, no. 99, pp. 1–10, 2016.
- [89] A. Sridhar, A. Vincenzi, M. Ruggiero, T. Brunschwiler, and D. Atienza, "3D-ICE: Fast compact transient thermal modeling for 3D ICs with inter-tier liquid cooling," in *Proceedings of the International Conference* on Computer-Aided Design. IEEE Press, 2010, pp. 463–470.
- [90] J. Xie and M. Swaminathan, "Electrical-thermal cosimulation with nonconformal domain decomposition method for multiscale 3-D integrated systems," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 4, no. 4, pp. 588–601, 2014.
- [91] Y. Shao, Z. Peng, and J.-F. Lee, "Thermal analysis of high-power integrated circuits and packages using nonconformal domain decomposition method," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 3, no. 8, pp. 1321–1331, 2013.
- [92] A. Corigliano, M. Dossi, and S. Mariani, "Domain decomposition and model order reduction methods applied to the simulation of multiphysics problems in MEMS," *Computers & Structures*, vol. 122, pp. 113–127, 2013.
- [93] F. M. White, Viscous Fluid Flow. McGraw-Hill New York, 2006.
- [94] J. Krupka, J. Breeze, A. Centeno, N. Alford, T. Claussen, and L. Jensen, "Measurements of permittivity, dielectric loss tangent, and resistivity of float-zone silicon at microwave frequencies," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 54, no. 11, pp. 3995–4001, 2006.

- [95] P.-S. Lee, S. V. Garimella, and D. Liu, "Investigation of heat transfer in rectangular microchannels," *International Journal of Heat and Mass Transfer*, vol. 48, no. 9, pp. 1688–1704, 2005.
- [96] H. Mizunuma, Y.-C. Lu, and C.-L. Yang, "Thermal modeling and analysis for 3-D ICs with integrated microchannel cooling," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions* on, vol. 30, no. 9, pp. 1293–1306, 2011.
- [97] D. Chen, E. Li, E. Rosenbaum, and S.-M. Kang, "Interconnect thermal modeling for accurate simulation of circuit timing and reliability," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, vol. 19, no. 2, pp. 197–205, 2000.
- [98] O. Schilling, M. Schäfer, K. Mainka, M. Thoben, and F. Sauerland, "Power cycling testing and FE modelling focussed on Al wire bond fatigue in high power IGBT modules," *Microelectronics Reliability*, vol. 52, no. 9, pp. 2347–2352, 2012.
- [99] D. Frear, J. Jang, J. Lin, and C. Zhang, "Pb-free solders for flip-chip interconnects," JOM Journal of the Minerals, Metals and Materials Society, vol. 53, no. 6, pp. 28–33, 2001.
- [100] J. Zhang, M. O. Bloomfield, J. q. Lu, R. J. Gutmann, and T. S. Cale, "Modeling thermal stresses in 3-D IC interwafer interconnects," *Semi*conductor Manufacturing, IEEE Transactions on, vol. 19, no. 4, pp. 437–448, 2006.
- [101] C. C. Jung, C. Silber, and J. Scheible, "Heat generation in bond wires," Components, Packaging and Manufacturing Technology, IEEE Transactions on, vol. 5, no. 10, pp. 1465–1476, 2015.
- [102] P. Antunes, A. Rocha, H. Lima, H. Varum, and P. Andre, "Thin bonding wires temperature measurement using optical fiber sensors," *Measurement*, vol. 44, no. 3, pp. 554–558, 2011.
- [103] F. Z. Kong, W. Y. Yin, J. F. Mao, and Q. H. Liu, "Electro-thermomechanical characterizations of various wire bonding interconnects illuminated by an electromagnetic pulse," *Advanced Packaging, IEEE Transactions on*, vol. 33, no. 3, pp. 729–737, 2010.
- [104] X. P. Wang, W. Y. Yin, and S. He, "Multiphysics characterization of transient electrothermomechanical responses of through-silicon vias applied with a periodic voltage pulse," *Electron Devices, IEEE Transactions on*, vol. 57, no. 6, pp. 1382–1389, 2010.
- [105] N. Li, J. Mao, W. S. Zhao, M. Tang, W. Chen, and W. Y. Yin, "Electrothermal cosimulation of 3-D carbon-based heterogeneous interconnects," *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, vol. 6, no. 4, pp. 518–526, April 2016.
- [106] K. Zhang and J.-M. Jin, "Parallel FETI-DP algorithm for efficient simulation of large-scale EM problems," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields*, 2016.
- [107] V. Košel, R. Sleik, and M. Glavanovics, "Transient non-linear thermal FEM simulation of smart power switches and verification by measurements," in *Thermal Investigation of ICs and Systems, 2007. THER-MINIC 2007. 13th International Workshop on.* IEEE, 2007, pp. 110– 114.
- [108] Z. Ren, W. Y. Yin, Y. B. Shi, and Q. H. Liu, "Thermal accumulation effects on the transient temperature responses in LDMOSFETs under the impact of a periodic electromagnetic pulse," *Electron Devices*, *IEEE Transactions on*, vol. 57, no. 1, pp. 345–352, 2010.
- [109] J. Liu, O. Salmela, J. Sarkka, J. E. Morris, P.-E. Tegehall, and C. Andersson, *Reliability of Microtechnology: Interconnects, Devices and Systems*. Springer Science & Business Media, 2011.
- [110] M. Huang, S. Mao, H. Feick, H. Yan, Y. Wu, H. Kind, E. Weber, R. Russo, and P. Yang, "Room-temperature ultraviolet nanowire nanolasers," *Science*, vol. 292, no. 5523, pp. 1897–1899, 2001.
- [111] Y.-M. Lin, S. B. Cronin, J. Y. Ying, M. S. Dresselhaus, and J. P. Heremans, "Transport properties of bi nanowire arrays," *Applied Physics Letters*, vol. 76, no. 26, pp. 3944–3946, 2000.
- [112] A. Dey, O. P. Bajpai, A. K. Sikder, S. Chattopadhyay, and M. A. S. Khan, "Recent advances in CNT/graphene based thermoelectric polymer nanocomposite: A proficient move towards waste energy harvesting," *Renewable and Sustainable Energy Reviews*, vol. 53, pp. 653–671, 2016.
- [113] M. Neek-Amal, P. Xu, J. Schoelz, M. Ackerman, S. Barber, P. Thibado, A. Sadeghi, and F. Peeters, "Thermal mirror buckling in freestanding graphene locally controlled by scanning tunnelling microscopy," *Nature Communications*, vol. 5, no. 4962, 2014.
- [114] E. Oesterschulze, M. Stopka, L. Ackermann, W. Scholz, and S. Werner, "Thermal imaging of thin films by scanning thermal microscope," J. Vac. Sci. Technol. B, vol. 14, no. 832, 1996.

- [115] X. Zheng, D. G. Cahill, R. Weaver, and C. Zhao, J, "Micron-scale measurements of the coefficient of thermal expansion by time-domain probe beam deflection," *Journal of Applied Physics*, vol. 104, no. 073509, 2008.
- [116] S.-W. Chung, J.-Y. Yu, and J. R. Heath, "Silicon nanowire devices," Applied Physics Letters, vol. 76, no. 15, pp. 2068–2070, 2000.
- [117] S. Piscanec, M. Cantoro, A. C. Ferrari, J. A. Zapien, Y. Lifshitz, S. T. Lee, S. Hofmann, and J. Robertson, "Raman spectroscopy of silicon nanowires," *Phys Rev B*, vol. 68, p. 241312, 2003.
- [118] G. Popescu, Quantitative Phase Imaging of Cells and Tissues, ser. McGraw-Hill biophotonics. McGraw-Hill Education, 2011.
- [119] G. Popescu, T. Ikeda, R. R. Dasari, and M. S. Feld, "Diffraction phase microscopy for quantifying cell structure and dynamics," *Opt. Lett.*, vol. 31, no. 6, pp. 775–777, Mar 2006.
- [120] C. Edwards, A. Arbabi, G. Popescu, and L. L. Goddard, "Optically monitoring and controlling nanoscale topography during semiconductor etching," *Light Sci Appl*, vol. 1, no. 9, p. e30, 2012.
- [121] R. Zhou, C. Edwards, A. Arbabi, G. Popescu, and L. L. Goddard, "Detecting 20 nm wide defects in large area nanopatterns using optical interferometric microscopy," *Nano Lett.*, vol. 13, no. 8, p. 37163721, 2013.
- [122] C. Edwards, R. Zhou, S.-W. Hwang, S. J. McKeown, K. Wang, B. Bhaduri, R. Ganti, P. J. Yunker, A. G. Yodh, J. A. Rogers, L. L. Goddard, and G. Popescu, "Diffraction phase microscopy: Monitoring nanoscale dynamics in materials science," *Appl Optics*, vol. 53, no. 27, pp. G33–G43, 2014.
- [123] C. Edwards, S. J. McKeown, J. Zhou, G. Popescu, and L. L. Goddard, "In situ measurements of the axial expansion of palladium microdisks during hydrogen exposure using diffraction phase microscopy," *Optical Materials Express*, vol. 4, no. 12, pp. 2559–2564, 2014.
- [124] B. Bhaduri, C. Edwards, H. Pham, R. Zhou, T. H. Nguyen, L. L. Goddard, and G. Popescu, "Diffraction phase microscopy: Principles and applications in materials and life sciences," Adv. Opt. Photon., vol. 6, no. 1, pp. 57–119, Mar 2014.
- [125] C. Edwards, A. Arbabi, B. Bhaduri, X. Wang, R. Ganti, P. J. Yunker, A. G. Yodh, G. Popescu, and L. L. Goddard, "Measuring the nonuniform evaporation dynamics of sprayed sessile microdroplets with quantitative phase imaging," *Langmuir*, vol. 31, no. 40, pp. 11020–11032, 2015.

- [126] D. Hohlfeld, M. Epmeier, and H. Zappe, "A thermally tunable, siliconbased optical filter," *Sensors and Actuators A: Physical*, vol. 103, no. 12, pp. 93–99, 2003.
- [127] Q. Fang, J. F. Song, T. Y. Liow, H. Cai, M. B. Yu, G. Q. Lo, and D. L. Kwong, "Ultralow power silicon photonics thermo-optic switch with suspended phase arms," *IEEE Photonics Technology Letters*, vol. 23, no. 8, pp. 525–527, April 2011.
- [128] T. Lalinský, E. Burian, M. Drzík, S. Hascík, Z. Mozolová, and J. Kuzmík, "Thermal actuation of a GaAs cantilever beam," *Journal* of Micromechanics and Microengineering, vol. 10, no. 2, pp. 293–298, 2000.
- [129] K. N. Lee, D. S. Lee, S. W. Jung, Y. H. Jang, Y. K. Kim, and W. K. Seong, "A high-temperature mems heater using suspended silicon structures," *Journal of Micromechanics and Microengineering*, vol. 19, no. 11, p. 115011, 2009.
- [130] C. C. Williams and H. K. Wickramasinghe, "Scanning Thermal Profiler," Appl. Phys. Lett., vol. 49, p. 15871589, 1986.
- [131] S. Goms, A. Assy, and P.-O. Chapuis, "Scanning thermal microscopy: A review," *Phys. Status Solidi A*, vol. 212, no. 3, p. 477494, 2015.
- [132] R. Huber, M. Koch, and J. Feldmann, "Laser-induced thermal expansion of a scanning tunneling microscope tip measured with an atomic force microscope cantilever," *Applied Physics Letters*, vol. 73, no. 17, pp. 2521–2523, 1998.
- [133] A. Laraoui, H. Aycock-Rizzo, Y. Gao, X. Lu, E. Riedo, and C. A. Meriles, "Imaging thermal conductivity with nanoscale resolution using a scanning spin probe," *Nature Communications*, vol. 6, no. 8954, 2015.
- [134] S. Li, K. Zhang, J.-M. Yang, L. Lin, and H. Yang, "Single quantum dots as local temperature markers," *Nano Letters*, vol. 7, no. 10, pp. 3102–3105, 2007.
- [135] P. Löw, B. Kim, N. Takama, and C. Bergaud, "High-spatial-resolution surface-temperature mapping using fluorescent thermometry," *Small*, vol. 4, no. 7, pp. 908–914, 2008.
- [136] I. Calizo, A. A. Balandin, W. Bao, F. Miao, and C. N. Lau, "Temperature dependence of the Raman spectra of graphene and graphene multilayers," *Nano Letters*, vol. 7, no. 9, pp. 2645–2649, 2007.
- [137] H. V. Pham, C. Edwards, L. L. Goddard, and G. Popescu, "Fast phase reconstruction in white light diffraction phase microscopy," *Appl. Opt.*, vol. 52, no. 1, pp. A97–A101, Jan 2013.

- [138] J. M. Huntley and H. Saldner, "Temporal phase-unwrapping algorithm for automated interferogram analysis," *Appl. Opt.*, vol. 32, no. 17, pp. 3047–3052, Jun 1993.
- [139] L. D. Smith, R. E. Anderson, D. W. Forehand, T. J. Pelc, and T. Roy, "Power distribution system design methodology and capacitor selection for modern CMOS technology," *IEEE Transactions on Advanced Packaging*, vol. 22, no. 3, pp. 284–291, 1999.
- [140] T.-L. Wu, H.-H. Chuang, and T.-K. Wang, "Overview of power integrity solutions on package and PCB: Decoupling and EBG isolation," *IEEE Transactions on Electromagnetic Compatibility*, vol. 52, no. 2, pp. 346–356, 2010.
- [141] M.-J. Pan and C. A. Randall, "A brief introduction to ceramic capacitors," *IEEE Electrical Insulation Magazine*, vol. 26, no. 3, pp. 44–50, 2010.
- [142] B. Brim, I. Novak, T. Michalka, W. Companioni, S. Tsubota, and S. Chitwood, "Needs and capabilities for modeling of capacitor derating," in *DesignCon*, 2016.
- [143] E.-P. Li, X.-C. Wei, A. C. Cangellaris, E.-X. Liu, Y.-J. Zhang, M. D'amore, J. Kim, and T. Sudo, "Progress review of electromagnetic compatibility analysis technologies for packages, printed circuit boards, and novel interconnects," *IEEE Transactions on Electromagnetic Compatibility*, vol. 52, no. 2, pp. 248–265, 2010.
- [144] C. Randall, R. Newnham, and L. Cross, "History of the first ferroelectric oxide, BaTiO3," Materials Research Institute, The Pennsylvania State University, University Park, Pa, USA, 2004.
- [145] B. Jaffe, *Piezoelectric Ceramics*. Elsevier, 2012.
- [146] M. B. Smith, K. Page, T. Siegrist, P. L. Redmond, E. C. Walter, R. Seshadri, L. E. Brus, and M. L. Steigerwald, "Crystal structure and the paraelectric-to-ferroelectric phase transition of nanoscale Ba-TiO3," *Journal of the American Chemical Society*, vol. 130, no. 22, pp. 6955–6963, 2008.
- [147] M. Vijatović, J. Bobić, and B. Stojanović, "History and challenges of barium titanate: Part II," *Science of Sintering*, vol. 40, no. 3, pp. 235–244, 2008.
- [148] Ceramic Dielectric Capacitors, Classes I, II, III and IV. EIA Standard RS-198, 1983.

- [149] J.-W. Liou and B.-S. Chiou, "Effect of direct-current biasing on the dielectric properties of barium strontium titanate," *Journal of the American Ceramic Society*, vol. 80, no. 12, pp. 3093–3099, 1997.
- [150] K. M. Johnson, "Variation of dielectric constant with voltage in ferroelectrics and its application to parametric devices," *Journal of Applied Physics*, vol. 33, no. 9, pp. 2826–2831, 1962.
- [151] J. Prymak, M. Randall, P. Blais, and B. Long, "Why that 47 uF capacitor drops to 37 uF, 30 uF, or lower," in *Proc. of the CARTS USA Conference*, 2008.
- [152] R. Bradt and G. Ansell, "Aging in tetragonal ferroelectric barium titanate," *Journal of the American Ceramic Society*, vol. 52, no. 4, pp. 192–198, 1969.
- [153] L. E. Wojewoda, M. J. Hill, K. Radhakrishnan, and N. Goyal, "Use condition characterization of MLCCs," *IEEE Transactions on Advanced Packaging*, vol. 32, no. 1, pp. 109–115, Feb 2009.
- [154] "PacketMicro, RF probing with calibration using Anritsu VNA." [Online]. Available: http://www.packetmicro.com
- [155] "Keysight Technologies. Evaluating DC-DC converters and PDN with the E5061B LF-RF network analyzer." [Online]. Available: http://www.keysight.com
- [156] D. M. Pozar, *Microwave Engineering*. John Wiley & Sons, 2009.
- [157] C. R. Paul, Inductance: Loop and Partial. John Wiley & Sons, 2011.
- [158] T. Roy, L. Smith, and J. Prymak, "ESR and ESL of ceramic capacitor applied to decoupling applications," in *Electrical Performance of Electronic Packaging*, 1998. IEEE 7th Topical Meeting on. IEEE, 1998, pp. 213–216.
- [159] M.-G. Kim, B. H. Lee, and T.-Y. Yun, "Equivalent-circuit model for high-capacitance MLCC based on transmission-line theory," *IEEE Transactions on Components, Packaging and Manufacturing Technol*ogy, vol. 2, no. 6, pp. 1012–1020, 2012.
- [160] A. Yariv, Y. Xu, R. K. Lee, and A. Scherer, "Coupled-resonator optical waveguide: A proposal and analysis," *Opt. Lett.*, vol. 24, no. 11, pp. 711–713, Jun 1999.
- [161] D. A. B. Miller, "Optical interconnects to electronic chips," Appl. Opt., vol. 49, no. 25, pp. F59–F70, Sep 2010.

- [162] P. Sun and R. M. Reano, "Low-power optical bistability in a freestanding silicon ring resonator," *Optics Letters*, vol. 35, no. 8, pp. 1124– 1126, 2010.
- [163] X. Zheng, Y. Luo, G. Li, I. Shubin, H. Thacker, J. Yao, K. Raj, J. E. Cunningham, and A. V. Krishnamoorthy, "Enhanced optical bistability from self-heating due to free carrier absorption in substrate removed silicon ring modulators," *Opt. Express*, vol. 20, no. 10, pp. 11478– 11486, May 2012.
- [164] M. Soltani, Q. Li, S. Yegnanarayanan, and A. Adibi, "Improvement of thermal properties of ultra-high q silicon microdisk resonators," *Optics Express*, vol. 15, no. 25, pp. 17305–17312, 2007.
- [165] A. Arbabi and L. Goddard, "Dynamics of self-heating in microring resonators," *Photonics Journal, IEEE*, vol. 4, no. 5, pp. 1702–1711, Oct 2012.