TRANSIENT JOULE HEATING IN NANO-SCALE EMBEDDED ON-CHIP INTERCONNECTS

A Dessertation
Presented to
The Academic Faculty

by

Banafsheh Barabadi

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
George W Woodruff School of Mechanical Engineering

Georgia Institute of Technology
May 2014

Copyright © May 2014 by Banafsheh Barabadi

TRANSIENT JOULE HEATING IN NANO-SCALE EMBEDDED ON-CHIP INTERCONNECTS

Approved by:

Dr. Yogendra K. Joshi, Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Satish Kumar, Advisor
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Samuel Graham
School of Mechanical Engineering
Georgia Institute of Technology

Dr. Muhannad Bakir
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Madhavan Swaminathan
School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Valeriy Sukharev
Design to Silicon Division
Mentor Graphics Corporation

Date Approved: December 16th, 2013

To my dearest family, Mahmood, Soori and Behnam, for their never-ending love.

ACKNOWLEDGEMENTS

I would like to take this opportunity to deeply thank all the people who have helped me throughout the course of my doctoral study at Georgia Institute of Technology. First and foremost, I wish to express my most sincere gratitude to my advisors, Dr. Yogendra Joshi and Dr. Satish Kumar, for their continuous support, guidance, and constructive critiques. I have been fortunate to present my work at various conferences, seminars, and workshops and to connect with other researchers in the field. I cannot thank my advisors enough for their confidence in me and for providing me with this great experience.

I am very thankful to my committee members, Dr. Muhannad Bakir, Dr. Samuel Graham, and Dr. Madhavan Swaminathan for their time and valuable feedback throughout my PhD proposal and defense process. I would also like to express my gratitude to Dr. Valeriy Sukharev not only for being on my PhD committee but also for being my mentor and constant support since January 2012, when I worked at Mentor Graphics under his supervision as an intern. His continuous guidance has greatly helped me gain insight into the current challenges of Integrated Circuits (IC) industry. The financial support for this work was provided in part by the Semiconductor Research Corporation (SRC) and Mentor Graphics.

I extend many thanks to my colleagues and alumni at the Microelectronics and Emerging Technologies Thermal Laboratory (METTL) and Micro Nano Devices & Systems Lab (MiNDS) for their valuable time, advice and suggestions that have helped in developing my research process. I would specially like to thank Dr. Emad Samadiani for

the insightful discussions on the theory of multi-scale modeling during the primary stages of my research.

I would particularly like to thank my friend, Dr. Amirhossein Atabaki, for his invaluable help and suggestions regarding the design and development of transient thermal measurement techniques to study embedded on-chip interconnects. His valuable recommendations have been a key influence on the advancement of my experimental work.

I am grateful to all the staff members at the Nanotechnology Research Center (NRC) and Microelectronics Research Center (MiRC) for their invaluable help and guidance with the micro/nano-fabrication tools and processes in the cleanroom. In particular, I would like to acknowledge Gary Spinner, Devin Brown, Viny Nguyen, Ben Hollerbach, and Mikkel Thomas who have assisted me with their expertise to resolve fabrication challenges.

Last but not the least, I wish my deepest gratitude to my dearests. I thank Ayin Mokrivala for his endless encouragement, patience, and understanding. His continuous support has kept me motivated in challenging situations. I am wholeheartedly thankful to my mother, Soori, my father, Mahmood, and my brother, Behnam, for their unconditional love and unwavering support, which have always helped me through difficult times. I believe beyond the shadow of doubt that I would not be in this position without their love or encouragement.

Thank you,

Banafsheh Barabadi

TABLE OF CONTENTS

Acknowledgements iv

List of Tables x

List of Figures xi

Nomenclature xx

Summary xxvi

Chapter 1: Introduction 1

1.1 Importance of Size Effects 2

1.2 Studies on Interconnect Joule Heating 3

1.3 Multi-scale Thermal Models for Microelectronics 5

1.3.1 Bottom-up Approaches 6

1.3.2 Top-down Approaches 8

1.3.3 Micro and Nanoscale Thermal Characterization Techniques 9

1.3.4 Measurement of Thin Metallic Film Properties 11

1.4 Overall Research Contributions 14

1.5 Thesis Outline 15

Chapter 2: Transient Heat Conduction in On-Chip Interconnects Using Proper Orthogonal Decomposition (POD) Technique 19

2.1 Objective 19

2.2 System Identification Methods 19

2.3 History of POD 21

2.4 Fundamentals of POD Method 21

2.4.1 Generating the Observation Matrix 22

2.4.2 Calculating Basis Functions (POD modes) 23

2.4.3 Calculating POD coefficients, bi, via Galerkin Projection method 25

2.4.4 Generating the POD temperature field: 27

2.5 Case Studies 28

2.6 Results and Comparison 31

2.6.1 Case 1 31

2.6.2 Case 2 35

2.7 Conclusion and Discussion 37

Chapter 3: Multi-scale Transient Analysis of Microelectronics Using Hybrid Scheme Approach 41

3.1 Introduction 41

3.2 Hybrid Scheme for Multi-scale Thermal Modeling 42

3.2.1 Fundamentals of POD Method 42

3.2.2 Progressive Zoom-in Approach 44

3.3 The Hybrid Scheme Procedure 45

3.4 Results and Discussion 46

3.5 Summary and Conclusion 58

Chapter 4: Rapid Multi-scale Transient Thermal Modeling of Packaged Microprocessors Using Hybrid Scheme 60

4.1 Objective 60

4.2 The Hybrid Scheme 61

4.3 Package and Chip Models 64

4.4 Transient Power Map 68

4.5 Results and Discussion 69

4.6 Summary and Conclusion 77

Chapter 5: Transient Thermal Characterization of Embedded Nanoscale Metallic Films 79

5.1 Objective 79

5.2 Micro and Nanoscale Thermal Characterization Techniques 79

5.2.1 Infrared (IR) Microscopy 80

5.2.2 Pump-Probe Transient Thermoreflectance (PPTTR) Technique 80

5.2.3 Network Identification by Deconvolution (NID) assisted PPTTR 81

5.2.4 Scanning Thermal Microscopy (SThM) 82

5.2.5 Scanning Joule Expansion Microscopy (SJEM) 82

5.2.6 Raman Thermometry 83

5.3 Proposed Approach: Sub-micron Resistance Thermometry 83

5.4 Design and Fabrication 84

5.4.1 Design of the layout 84

5.4.2 Fabrication process 85

5.5 Fabrication Challenges 89

5.5.1 Pt fabrication process 89
5.5.2 Determining the contact resistance 92
5.6 Device Setup and Characterization 95
5.7 Experimental Setup and Procedure 97
5.8 IR Imaging Setup and Procedure 99
5.9 Results and Comparison 100
5.9.1 Steady state measurements 100
5.9.2 Numerical model for verification 103
5.9.3 Transient measurements 104
5.10 Conclusion 110
Chapter 6: Thermal Conductivity Measurements of Nanoscale Embedded Metallic Thin Films 112
6.1 Importance of Size Effects 112
6.2 Objective 112
6.3 Measurement of Thin Metallic Film Properties 112
6.4 Proposed Approach: Steady State "Hourglass" Design 118
6.5 Design and Fabrication 124
6.6 Device Setup and Characterization 127
6.7 Experimental Setup and Procedure 130
6.7.1 Determining the initial electrical resistivity, 130
6.7.2 Determining the Temperature Coefficient of Resistance (TCR) 131
6.7.3 Determining the slope of R vs. P for each structure 131
6.8 Experimental Data 132
6.9 Electro-thermally Coupled Numerical Model 133
6.10 Results and Comparison 135
6.10.1 Numerical results 135
6.10.2 Numerical fit to measurements of κ as a function of thickness 136
6.10.3 Effect of constriction width, w_R on the model's sensitivity 140
6.10.4 Thermal conductivity as a function of film thickness 141
6.10.5 Study the dependence of thermal conductivity on temperature 144
6.10.6 Numerical fit to measurements of κ as a function of temperature 145
6.10.7 Thermal conductivity as a function of temperature 148
6.11 Conclusion 151

Chapter 7: Summary and Conclusion 154
7.1 Conclusion 154
7.2 Future Recommendations 160
Appendix A: Analytical Proof for POD Model 162
References 165

LIST OF TABLES

Table 3-1 Material properties and dimensions of the package 51

Table 3-2 Properties and dimensions of the function blocks for chip level simulation.. 56

Table 4-1 Material properties, thermal properties, and dimensions of the package. * minimum reported value for thermal conductivity of copper released by Intel [103]... 68

Table 4-2 Material properties, thermal properties, and dimensions of each layer of the die for chip level simulation.. 69

Table 5-1 Summary of micro/nanoscale thermal measurement techniques. 84

Table 5-2 Modeling Parameters 106

Table 6-1 Fabricated devices' IDs and parameters 126

Table 6-2 Modeling Parameters 136

Table 6-3 Experiment specification and results for fabricated devices including thermal conductivities. 146

Table 6-4 Experimental specification for determining β_k including the initial thermal conductivities 152

LIST OF FIGURES

Figure 1-1 Schematic of interconnect stack with metal lines and vias [2]. 2

Figure 1-2 The overall flowchart of numerical and experimental techniques developed in this work to address transient Joule heating in nanoscale embedded on-chip interconnects. 17

Figure 2-1 Schematic of the computational domain with a cross sectional area of 1.44 μm x 720 nm. It consists of a set of 360 nm x 360 nm interconnects that are evenly spaced and embedded in the dielectric. The mesh used in the POD and FE models is shown. In this study: $H_{int} = H_{de}$= 360 nm and P = 4H_{int} = 1.44 μm. 29

Figure 2-2 Different types of heat sources used in this study. Case 1- only step function (solid line) & Case 2-sinusoidal and step function (dashed line). 32

Figure 2-3 Eigenvalues or energy percentage in log form vs. number of the POD modes. In order to build a reliable reduced order model, the number of basis functions used for the projection was chosen such that the cumulative correlation energy of the modes are greater or equal to 99.99%. The first two modes capture over 98% of the energy. 33

Figure 2-4 First five POD modes, or basis functions, plotted in 2D contours. The POD modes are normalized by the total sum of the modes chosen for each study. 34

Figure 2-5 First five b-coefficients vs. time using the Galerkin Projection Technique for Case 1. 35

Figure 2-6 Spatial variation of temperature rise after 20 μs for FE (top) and POD (bottom) models using 5 basis functions for Case 1. 36

Figure 2-7 Comparison of temporal dependence of temperature rise in the left-most node of the top edge (node 1 in the interconnect), x=0 using 17, 9, 5, and 2 observations for Case 1. 37

Figure 2-8 Spatial variation of temperature rise after 18 μs in x direction for the upper edge of the structure (a), y direction for the left edge of the structure (c), and along

the diagonal (e). Spatial variation of temperature after 19.2 μs in x-direction for the upper edge of the structure (b), y-direction for the left edge of the structure (d), and along the diagonal (f). FE results are plotted in solid lines and POD results using 5 basis functions are plotted in circular markers. The results are for Case 2. 38

Figure 2-9 Comparison of temporal dependence of temperature rise at nodes 1-5 along the diagonal for Case 2. FE results are plotted in circular markers and POD results are shown by dashed lines... 39

Figure 3-1 Flowchart of the hybrid scheme for multi-scale thermal modeling. ... 46

Figure 3-2 (a) Schematic of a simplified FCBGA for package level modeling, (b) Zoomed-in schematic of the die layer used in chip level modeling. 48

Figure 3-3 Spatial distribution of temperature rise extracted from FE method after 1 s for (a) FCBGA package and (b) Chip. 49

Figure 3-4 Two dimensional contour plots of the first 4 POD modes at z = 1.331 mm from the bottom of the package; this plane crosses the center of die. 50

Figure 3-5 Spatial distribution of temperature rise at 1 s extracted from the POD model (a) and FE simulation (c). The domain is sliced vertically along XZ plane. The right-most slice is the A-A cross section (b).. 53

Figure 3-6 Comparison of temporal dependence of temperature rise between FE (markers) and POD (solid lines) models at 4 different points (a) and the corresponding randomly generated total chip power (b). 55

Figure 3-7 Dynamic randomly generated power profile for function blocks 1, 2, 3 (a-c) and randomly generated total chip power (d).. 57

Figure 3-8 Transient temperature distribution at the interface of Device and Interconnect/Dielectric layers. 58

Figure 4-1 Flowchart description of the methodology used in this study for multi-scale transient thermal modeling of a representative Flip Chip Ball Grid Array (FCBGA) microprocessor package. 64

Figure 4-2 (a) Photo of the Intel Core 2 Duo processor code-named Penryn die including major blocks [110], (b) Spatial distribution of power of the die (Power map). 65

Figure 4-3 (a) Schematic of a simplified FCBGA package used for package level modeling, (b) zoomed-in schematic of the die layer used in chip level modeling. 67

Figure 4-4 Spatial distribution of temperature rise extracted from FE method after 10 s with uniform heating for (a) FCBGA package & (b) chip. 70

Figure 4-5 (a) Volume-averaged temporal distribution of temperature until steady state condition is reached. The arrow indicates the time where power disturbance is applied. (b) Oscillatory power disturbance applied after steady state is reached. (c) Impulsive power disturbance used to demonstrate the POD prediction capabilities........ 71

Figure 4-6 Two dimensional contour plots of the first three POD Modes at z = 1.33 mm which is the plane crossing the center of die horizontally (a, b, c) and at y=18.5 mm which is the plane crossing the center of the die vertically (d, e, f). 73

Figure 4-7 Spatial distribution of temperature for the surge-type power after 3 s extracted from the POD model (a) and FE simulation (c). The domain is sliced vertically along the A-A cross section (b) four times in the die area... 74

Figure 4-8 Comparison of Temporal dependence of temperature between FE (markers) and POD (solid lines) models four representative points in the domain marked on the plot. 75

Figure 4-9 Transient temperature rise (K) at the interface of Device and Si Layer (a and b) and at the bottom surface of the die , i.e. bottom of the Interconnect/Dielectric Layer (c and) at t=1.5 s (a & c) and t=3 s (b & d). 76

Figure 5-1 (a) Image of the infrared (IR) microscope used for steady state measurements. (b) Schematic diagram of the IR microscopy procedure. 81

Figure 5-2 (a) Layout of Interconnect and RTD architecture (b) Zoomed-in view of serpentine lines of RTD (c) Cross-sectional view of structure (not to scale)............. 85

Figure 5-3 (a) Optical Image of Interconnect layer. (b) SEM image of the serpentine section of RTDs fabricated on top of Interconnect layer. (c) Zoomed-in SEM image of the serpentine section of an RTD. 88

Figure 5-4 (a) SEM image of final device. (b) Zoomed-in SEM image of the structure which shows interconnects fabricated beneath the RTDs, Serpentine section of the RTDs, and the extension of the RTDs fabricated on top. 90

Figure 5-5 (a) SEM image of a representative serpentine section of an RTD with defects. (b) Zoomed-in SEM image of the serpentine lines. (c) and (d) are two examples of faulty fabricated serpentine lines. 91

Figure 5-6 (a) SEM image of three adjacent serpentine lines for energy dispersive X-ray spectroscopy (EDX). EDX spectra of the points marked in the SEM image. 92

Figure 5-7 Schematic of the sample and the designed setup mounted inside the electron beam evaporation chamber. 93

Figure 5-8 SEM image of the junction area between Pt (serpentine part of RTDs) and Au (connecting extensions of RTDs). There is a contact resistance associated with the junction of Pt and Au. 94

Figure 5-9 SEM image of a typical RTD with (a) resistance R1 and (b) resistance R2. $R_1 = 2R_2 + ECR_0$ where R_0 is the electrical contact resistance for the Pt-Au junction. 95

Figure 5-10 SEM images of structure showing cracks created on surface of Pt film during deposition. 96

Figure 5-11 (a) Fabricated device mounted and wire-bonded to a printed circuit-board (PCB). The connecting soldered wires are shown. (b) Schematic of experimental setup. 97

Figure 5-12 Representative calibration measurement for determining the temperature coefficient of the electrical resistance (*TCR*) for RTDs. 98

Figure 5-13 (a) Optical image of a representative fabricated device specifying the location were the IR measurements were obtained. (b) A representative temperature map

attained by steady state IR thermometry. (c) Plot of temperature distribution along the arrow shown on the thermal map. 101

Figure 5-14 Comparison of temperature measurements of the middle interconnect using RTDs, IR Microscopy, and the Interconnect as a resistance thermometer. 102

Figure 5-15 (a) A representative quasi-steady spatial temperature distribution at the cross section of structure and (c) along the width of interconnect. (b) SEM image of the placement of a typical RTD over an interconnect line. 104

Figure 5-16 Experimental (diamond shaped black markers) and numerical (circular red markers) steady state temperature of the middle interconnect. 105

Figure 5-17 Experimental (solid black line) and numerical (dashed red line) normalized step response of the middle interconnect subjected to a 10 kHz square pulse (rising edge). Thermal time constant on the rising edge: Model: 11 μs; Experiment: 9 μs 107

Figure 5-18 Frequency response of the device. Black markers represent the experimentally obtained data and the solid red line is a first order model fitted to the data. The f_{3dB} of the system is 95 kHz. 108

Figure 5-19 Experimental (solid black line) and numerical (dashed red line) normalized transient thermal response of the middle interconnect subjected to a 10 kHz square pulse (dotted blue line). The primary vertical axis is for the input power and the secondary vertical axis is associated with the experimental and numerical temperature data. 109

Figure 5-20 Experimental (solid black line) and numerical (dashed red line) normalized transient thermal response of the middle interconnect subjected to a 10 kHz sinusoidal input power (dotted blue line). The primary vertical axis is for the input power and the secondary vertical axis is associated with the experimental and numerical temperature data. 110

Figure 6-1 Layout of Nath and Chopra's experimental structure used to measure thermal conductivity [53]. 113

Figure 6-2 Layout of the (a) suspended structure and (b) a cross section of the metal structure used by Liu et al. [57] to measure thermal conductivity. 115

Figure 6-3 Schematic diagram of the experimental setup for the scanning Joule expansion microscopy (SJEM) adopted from [46]. 117

Figure 6-4 (a) Layout of the proposed structure used to investigate the lateral thermal conductivity of embedded thin metal films, (b) Top view of the metal design referred to as "hourglass". As shown, w_R is the width of the constriction, r_R is the radius of the constricted area and r_i is the radius of the half circles used to create the outer part of the hourglass geometry. 121

Figure 6-5 Spatial distribution of (a) current density and (b) temperature in the proposed structure. For this model w_R = 300 nm, r_R = 1 μm, r_i = 5 μm, and thermal conductivity K=240 W/mK. 122

Figure 6-6 Top view of temperature distribution in the proposed structure for thermal conductivity varying from 140-240 W/mK. For this model w_R = 300 nm, r_R = 1 μm, r_i = 5 μm. To achieve a better visual comparison when changing thermal conductivity, the input current was set at a high value of 0.1 A. 123

Figure 6-7 FE results of resistance vs. input power for different values of thermal conductivity for three combination of w_r , r_R, r_i. (a) w_R = 300 nm, r_R = 3 μm, r_i = 10 μm (b) w_R = 200 nm, r_R = 3 μm, r_i = 10 μm (c) w_R = 200 nm, r_R = 5 μm, r_i = 10 μm. 125

Figure 6-8 Optical images of fabricated structures with (a) 10, (b) 20, and (c) 30 unit pattern "hourglass" per structure. (d) SEM image of one unit pattern in the structure. (e) zoomed-in SEM image of the constriction within the structure. 128

Figure 6-9 (a) Fabricated device mounted on Cu block with thermally conductive epoxy and attached to the circuit-board. (b) The wires are soldered to the board and the board is attached to a heat sink with thermal grease. 129

Figure 6-10 Calibration measurements for determining the temperature coefficient of the electrical resistance (TCR) for ten devices with 112 nm Cu film (diamond shaped markers) and eight devices with 60 nm Cu film (circular markers). Average TCR for 112

nm and 60 nm devices are 0.00241 K^{-1} and 0.00187 K^{-1} respectively. The bulk value of TCR for Cu is 0.003862 K^{-1}.. 130

Figure 6-11 Flowchart of the steady state experimental procedure in determining the thermal conductivity of embedded metal thin films. 133

Figure 6-12 Representative plot of resistance vs. input power measurements for (a) 112 nm thick Cu (device 112-5) and (b) 60 nm thick Cu (device 60-4). The solid red line is a linear fit to the experimental data. 135

Figure 6-13 (a) Layout of the FE model representing the unit pattern of the fabricated structure. (b) Zoomed-in section of the structure where the input current is applied. (c) Cross sectional view of the structure. The thermal boundary condition on top and bottom are natural convection and constant temperature respectively. In longitudinal direction (x) the walls are considered to be thermally insulated due to symmetry......... 137

Figure 6-14 (a) spatial distribution of current density in the unit pattern of the structure for Device 60-4 when subjected to 10 mA of current. (b) Top view of current density in the “hourglass” part of the structure. (c) Zoomed-in view of current density distribution in the constricted area within “hourglass”. 138

Figure 6-15 (a) Steady state temperature distribution in the unit pattern of the structure for Device 60-4 when subjected to 10 mA of current. (b) Top view of temperature distribution in the “hourglass” part of the structure. (c) Zoomed-in view of temperature distribution in the constricted area within “hourglass”. Thermal Conductivity for Device 60-4 is determined to be 151.62 W/mK. 139

Figure 6-16 Representative plot of numerically generated normalized thermal conductivity versus the slope of (*R* vs. *P*) for 112 nm Cu layers. Results are for device 112-9. The experimentally measured slope value is 0.1289. The corresponding thermal conductivity is determined to be 238.19 W/mK. The solid red line is a quadratic fit to the numerical data. 140

Figure 6-17 Representative plot of numerically generated normalized thermal conductivity versus the slope of (*R* vs. *P*) for 60 nm Cu layers. Results are for device 60-

4. The experimentally measured slope value is 0.688. The corresponding thermal conductivity is determined to be 151.62 W/mK. The solid red line is a quadratic fit to the numerical data. 141

Figure 6-18 Normalized thermal conductivity of 112 nm thick Cu layer versus the slope of (R vs. P) for different w_R. Representative results are for devices 112-4 (w_R=100 nm), 112-5 (w_R=200 nm), and 112-10 (w_R=300 nm) . The solid lines are quadratic fits to the numerical data. 142

Figure 6-19 Normalized thermal conductivity of 60 nm thick Cu films versus the slope of (R vs. P) for different w_R. Representative results are for devices 60-2 (w_R=100 nm), 60-4 (w_R=200 nm), and 60-8 (w_R=300 nm). The solid lines are quadratic fits to the numerical data. 143

Figure 6-20 Experimental data for the thermal conductivity of 112 and 60 nm Cu layers. There are ten and eight data points for 112 and 60 nm Cu layers respectively. Each data point represents one device. 144

Figure 6-21 Representative plot of numerically generated values for β_k versus the slope of (R vs. P) for 112 nm Cu layers at high power inputs. Results are for device 112-7. The experimentally measured slope value is 0.1945. The corresponding β_k is determined to be 0.103 W/mK^2. The initial thermal conductivity used in the model is 233.38 W/mK. The solid red line is a quadratic fit to the numerical data. 147

Figure 6-22 Representative plot of numerically generated values for β_k versus the slope of (R vs. P) for 60 nm Cu layers at high power inputs. Results are for device 60-5. The experimentally measured slope value is 0.6396. The corresponding β_k is determined to be 0.31 W/mK^2. The initial thermal conductivity used in the model is 146.57 W/mK. The solid red line is a quadratic fit to the numerical data. 148

Figure 6-23 Experimental data for the temperature dependent thermal conductivity of 112 and 60 nm Cu films. There are five and four data points for 112 and 60 nm Cu layers respectively. Each data point represents one device. 149

Figure 6-24 Representative plot of experimental data (black markers) and numerical results (dashed red line) for resistance vs. input power. The model is developed using temperature dependent conductivity for the full range of power spectrum. Results are for device 60-5. The initial value of thermal conductivity is κ_0=146.57 W/mK and β_k=0.31 W/mK^2. 150

NOMENCLATURE

Nomenclature for Chapter 2 and Appendix A

A	heat wave amplitude (W/m^3)
A_{ij}	coefficient in Eq. (2.9)
B_{ij}	coefficient in Eq. (2.9)
E_m	cumulative correlation energy
H	height (m)
J	current density (MA/cm^2)
Kn	Knudsen number
L	smallest length scale in the structure (m)
T	temperature (K)
X	arbitrary function of x
Y	arbitrary function of y
b	POD coefficient (K)
c_i	coefficient in Eq. (2.9)
c_p	specific heat (J/kgK)
f	constant spatial function
q_i	coefficient in Eq. (2.9)
$q'''(t)$	volumetric heat generation (W/m^3)
m	number of POD modes used
n	number of observations
$\langle .,. \rangle$	inner product of two functions
$(\dot{\ })$	time derivative

$(\ )''$	second spatial derivative
∇^2	Laplace operator

Greek symbols

Θ	the basis functions
Γ	time dependent coefficients
Λ	molecular mean free path (nm)
Ω	two-dimensional domain
Π	time dependent coefficients, Eq. (A.19)
α	thermal diffusivity (m^2/s)
β	weight coefficients
δ	arbitrary constant
φ	POD mode
γ	arbitrary constant
κ	thermal conductivity (W/mK)
λ	eigenvalue
μ	arbitrary constant
ρ	density (kg/m^3)
ρ_r	electrical resistivity (μΩ-cm)
$\Re$	covariance matrix
τ	heat wave period (μs)

Subscript

int	interconnect
de	dielectric

obs	observation

Superscript

C	mean centered
o	initial value

Nomenclature for Chapter 3

A	cross sectional area (m^2)
b_i	*i*-th POD coefficient (K)
dt	time step (s)
E_m	cumulative correlation energy
h	heat transfer coefficient (W/m^2)
K	thermal conductivity (W/mK)
Q	total Chip Power (W)
T	temperature (K)
T_0	time average of temperature (K)
z	height (mm)

Greek symbols

λ_i	*i*-th eigenvalue
φ_i	*i*-th POD mode

Subscripts

amb	room/ ambient
h_eff	effective horizontal value
hS	solder bumps (horizontal)

hU	underfill (horizontal)
S	solder bumps
tot	total
U	underfill
v_eff	effective vertical value

Nomenclature for Chapter 4

A_f	surface area of a single fin (m^2)
A_t	surface area associated with both the fins and the exposed portion of the heat sink (m^2)
A	cross sectional area (m^2)
b_i	*i*-th POD coefficient (K)
dt	time step (s)
E_m	cumulative correlation energy
h	heat transfer coefficient (W/m^2)
K	thermal conductivity (W/mK)
N	total number of the fins
Q	total Chip Power (W)
T	temperature (K)
T_0	time average of temperature (K)
z	height (mm)

Greek symbols

η_f	fin efficiency of a single fin

λ_i *i*-th eigenvalue

φ_i *i*-th POD mode

Subscripts

amb room/ ambient

eff effective

Nomenclature for Chapter 5

ECR electrical contact resistance (Ω)

f_{3dB} 3dB frequency (kHz)

LTI first-order linear time invariant system

R resistance of the serpentine lines (Ω)

R^2 coefficient of determination

T temperature (K)

TCR temperature coefficient of the electrical resistance (K^{-1})

Subscript

o initial value

Nomenclature for Chapter 6

A cross sectional area of the resistance (m^2)

L length of the resistance (m)

number number of data points

P input power (mW)

R	electrical resistance (Ω)
r_{i}	radius of the half circles used to create the outer part of the hourglass geometry (μm)
$RMSE$	root-mean-square error
r_{R}	radius of the constricted area (μm)
T	temperature (K)
TCR	temperature coefficient of the electrical resistance (K^{-1})
y_i^{exp}	experimental data points
y_i^{mod}	numerical data points
w_{r}	width of the constriction (μm)

Greek symbols

β_k	slope of ΔT vs. $\Delta\kappa$, Eq. (6.3) (W/mK^2)
κ	thermal conductivity (W/mK)
ρ^E	Electrical Resistivity (Ω·m)

Subscripts

o	initial value

SUMMARY

Major challenges in maintaining quality and reliability in today's microelectronics devices come from the ever increasing level of integration in the device fabrication, as well as the high level of current densities that are carried through the microchip during operation. In order to have a framework for design and reliability assessment, it is imperative to develop a predictive capability for the thermal response of micro-electronic components. A computationally efficient and accurate multi-scale transient thermal methodology was developed using a combination of two different approaches: "Progressive Zoom-in" method and "Proper Orthogonal Decomposition (POD)" technique. The proposed technique has the capability of handling several decades of length scale from tens of millimeter at "package" level to several nanometers at "interconnects" level at a considerably lower computational cost, while maintaining satisfactory accuracy. This ability also applies for time scales from seconds to microseconds corresponding to various transient thermal events. The proposed method also provides the ability to rapidly predict thermal responses under different power input patterns, based on only a few representative detailed simulations, without compromising the desired spatial and temporal resolutions. It is demonstrated that utilizing the proposed model, the computational time is reduced by at least two orders of magnitude at every step of modeling.

Additionally, a novel experimental platform was developed to evaluate rapid transient Joule heating in embedded nanoscale metallic films representing buried on-chip interconnects that are not directly accessible. Utilizing the state-of-the-art sub-micron embedded resistance thermometry the effect of rapid transient power input profiles with

different amplitudes and frequencies were studied. It is also demonstrated that a spatial resolution of 6 μm and thermal time constant of below 1 μs can be achieved using this technique. Ultimately, the size effects on the thermal and material properties of embedded metallic films were studied. A state-of-the-art technique to extract thermal conductivity of embedded nanoscale interconnects was developed. The proposed structure is the first device that has enabled the conductivity measurement of embedded metallic films on a substrate. It accounts for the effect of the substrate and interface without compromising the sensitivity of the device to the thermal conductivity of the metallic film. Another advantage of the proposed technique is that it can be integrated within the structure and be used for measurements of embedded or buried structures such as nanoscale on chip interconnects, without requiring extensive micro-fabrication. The dependence of the thermal conductivity on temperature was also investigated. The experimentally measured values for thermal conductivity and its dependence on temperature agree well with previous studies on free-standing nanoscale metallic bridges.

CHAPTER 1: INTRODUCTION

Since the 1970s, when the microprocessor became a commercially available and pervasive product, its clock rate has increased by approximately one million times. The key to this unprecedented advancement is the scaling of the interconnect wiring and transistor dimensions. In fact, a major concern in the design of microprocessors is the quality and reliability of on-chip interconnects [1]. These interconnects are usually Al-Cu- or Cu-based submicron lines deposited on an insulation layer. Figure 1-1 displays a schematic cross-section of an interconnect stack in microprocessors adapted from [2]. Because of the increasing level of integration in microprocessors, interconnects are subjected to high current density and, hence, to high temperature increase under operating conditions. Moreover, because of the large thermal expansion mismatch between the metallic line and the underlying dielectric layer, high thermomechanical stresses develop. Several experimental and computational studies suggest that these factors are primarily responsible for morphological changes in the lines that result in open-circuit and short-circuit failures in interconnects and as a consequence, limiting the quality and reliability of the entire circuit. The basic elements of thermomechanical fatigue behavior of microelectronic interconnect structures, such as lines and vias, based on accelerated test results have been studied and FE analysis has been developed [3].

Interconnect lines always contain a variety of pre-existing defects such as voids and cracks [4]. Local hot spots, which originate from these defects, often have a major role in controlling the micro-mechanisms of interconnect line failures. Such failure processes are governed by the kinetics of inhomogeneous diffusions and/or reactions. For

example, the effect of electromigration, as well as the variation of diffusion rates, will accelerate void growth and translation, and their accompanied stress buildups, leading to a final failure of the interconnect line. The transient heat transfer in the system greatly influences the morphology of failure and the pattern of damage evolution, which depend strongly on the electric current loading rate.

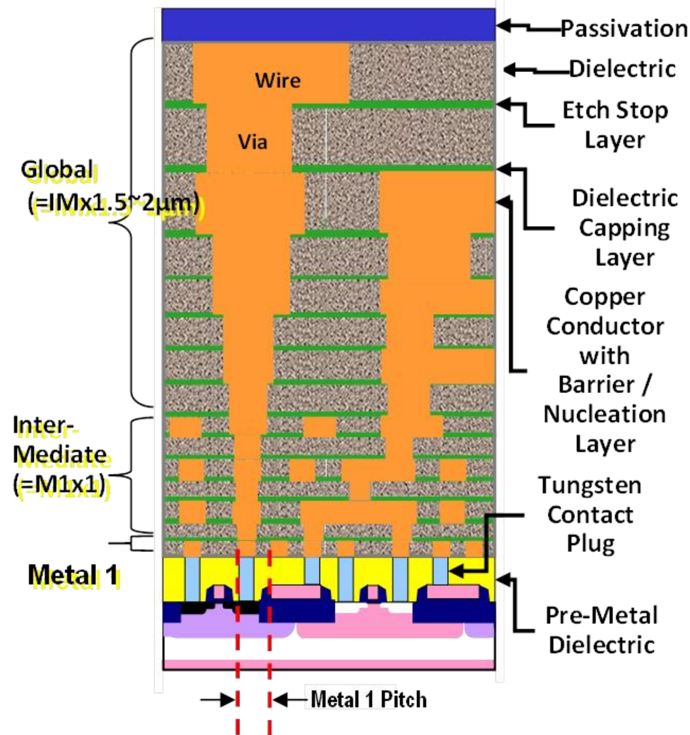

Figure 1-1 Schematic of interconnect stack with metal lines and vias [2].

Therefore, to precisely predict the performance and reliability of Interconnects, it is essential to determine the temperature distribution in the structure for realistic power inputs.

1.1 Importance of Size Effects

Another influential factor in the performance and reliability design of interconnects is the knowledge of their thermal and electrical properties. As the size of

interconnects reduces to, or even smaller than the electron mean free path (~40 nm in Cu at room temperature), electron transport becomes dominated by scattering at the metal-dielectric interface, and at grain boundaries. This scattering can decrease the electrical and thermal conductivity to less than half that of the bulk value [5-10], which confirms the need to also include the size effects in the thermal analysis of interconnects.

1.2 Studies on Interconnect Joule Heating

Joule heating in interconnects has been studied by various methods using different levels of approximation. These methods include analytical (2D) models, finite difference (FD) models, and finite element (FE) models [11-14]. A set of long uniformly spaced interconnects is often modeled as two-dimensional thermal spreading from a localized heat source, as is shown in Figure 1-1 adapted from [2]. Bilotti et al. [11] derived an analytical expressions for the heat spreading factor using double Schwarz-Christoffel conformational transformation. Chiang et al. [15] modeled the finite length interconnects incorporating a fin type equation. Resistance to heat conduction through the dielectric was assumed to be constant in this 2D model, with consideration given to heat conduction along the interconnect itself. This type of fin equation can be solved for any specified temperature condition near the ends. By adding average contributions to temperature rise from other interconnect levels, Chiang et al. [15] were able to include multi-level interconnects in their 2D modeling.

FE modeling is exemplified in the work of Chen et al. [14]. Interconnect temperature rise was computed using a commercially available finite element solver to achieve more realistic structures. A combination of analytically and numerically fitted solutions provided expressions for temperature rise in interconnect structures. Banerjee et

al. [16] studied the thermal break down of metal interconnects under conditions of short-time high Joule heating under electrostatic discharge (ESD) and electrical overstress (EOS) conditions. These particular failure types play an important role in interconnect/dielectric reliability. Banerjee, et al. developed a model that included the heating of the layered metal layers and the surrounding oxide. Their model related the maximum allowable current density to the signal pulse width. A methodology to quantify the role of electro-migration (EM) reliability and interconnect performance in determining optimal interconnect design in low-k/Cu systems was presented in [17]. The issue of reliability of interconnects has also been addressed [18, 19]. Chiang et al [19] also showed that heating effects can severely degrade reliability and speed performance.

It is, however, important to recognize that interconnects are not isolated entities. They are embedded in a stack of metal lines, vias, and dielectric layers as a part of a chip that is ultimately enclosed in a package. The majority of heat is generated in the transistors and metallic interconnects, with characteristic length scale of tens of nanometers ($\sim 10^{-8}$ m). This heat is then conducted through the passive/active components, substrate, and heat sink, prior to being convectively removed by the ambient environment. At this level, the characteristic length scale is on the order of tens of centimeters (10^{-1} m). Therefore, the effect of the entire structure should be accounted to acquire an accurate and realistic thermal solution for interconnects . In addition, transient thermal events in the interconnect structure can vary from tens of seconds (10^{2}) at the package level to microseconds (10^{-6}) at the interconnect level. In other words, the heat diffusion problem in interconnects is a multi-scale problem both in length and time scale.

The traditional FD and FE methods, discussed earlier, require a large number of computational nodes for such problems. As a result, computational times are long, even for a unit cell (micro-models), and hence unfeasible. Any method that is less complex, mainly analytical, will not be able to capture all the physics of the thermal problem. Additionally, previous models have been primarily concerned with the steady-state Joule heating in interconnects, while the pulse nature of the electric currents necessitates studying the effects of transient heat conduction in the interconnects as well. This confirms the need for the development of high fidelity transient multi-scale thermal models that can address the aforementioned issues. These multi-scale models should also be incorporated with nanoscale thermal characterization of interconnects where size effects are eminent. At nanoscale, thermal and electrical properties of metal lines can be determined through modeling which includes detailed electron and phonon transport simulations. Another approach at the nanoscale level is to conducting experiments to determine effective properties of interconnects. In any case, high resolution transient thermal measurements at interconnect level are essential to characterize the thermal transport in the structure and to validate the simulations.

1.3 Multi-scale Thermal Models for Microelectronics

Various reduced or compact thermal approaches have been proposed in the literature to address multi-scale thermal transport in chips and packages which trade resolution and accuracy for shorter computational time to carry out parametric design studies. [20, 21]. These methods can be classified in two main categories based on their computational procedures.

1.3.1 Bottom-up Approaches

Among these methodologies, the traditional bottom-up approaches are extensively used for transient thermal modeling. One of these methods is known as resistance-capacitance network.

In resistance-capacitance network based models, the transient model is mainly constructed using thermal impedances and by establishing the corresponding networks with thermal resistors and capacitors [22, 23]. Nonetheless, the accuracy of the models decreases for complex geometries, complex boundary conditions, and nonlinearity in the heat conduction equation [24].

Another common approach is recognized as compact model. Compact models can be finite volume or finite element based. In a traditional finite element (FE) or finite volume (FV) analysis, the domain is discretized into elements in a way that the area inside an element is homogeneous. It can, however, have anisotropic thermal conductivity. Compact models do not require conventional bilinear rectangular or homogeneous elements and can have elements comprising both metal and dielectric region.

Some of the first compact modeling work was done by Kreuger and Bar-Cohen in 1992 [25]. They modeled a chip package with a simplified resistor network. They showed that with a reduction in mesh size, shorter simulation times can be achieved. Compact thermal models for on-chip interconnect heating analysis have also been developed [26, 27]. The three-dimensional compact finite element modeling of interconnects with vias by Gurrum et al. [27] is briefly explained. The network topography of compact models becomes complex with the increase in model size,

potentially also compromising the accuracy of the model [28]. Another limitation of compact model is that it will be difficult to handle fluid/solid interactions.

Additionally, both of these models have primarily addressed the steady-state Joule heating in interconnects. However, pulsed currents and the resulting transient heat conduction in interconnect arrays remains a key concern in the design for reliability for the next generation high performance chips.

An alternative bottom-up based method for solving the heat conduction equation is the Transmission Line Matrix (TLM) approach [29, 30]. The TLM formulation is based on a resistance and capacitance network that represents the thermal system and can address transient events. The advantages of this formulation over traditional FE and FD methods are that TLM allows for temperature-dependent and inhomogeneous material parameters, non-uniform meshing, and non-uniform time stepping. Transient Joule heating in copper interconnects embedded in silicon dioxide with constant current density using the TLM and FE methods has been analyzed in [31]. Also investigated were the effects of the duration and amplitude of rapid square-wave source current pulses [32]. The stability of the results has been shown to be a limitation to this method [31, 33]. Furthermore, as the complexity of the structure increases, the simulation times increase, requiring a combination of TLM with multi-scale model reduction methods[34].

Another multi-scale approach is called Proper Orthogonal Decomposition (POD) that can effectively reduce the order of the dynamics of a continuous system. POD is a robust method of data analysis that provides low-dimensional but accurate descriptions of a high-dimensional system. It expands a set of data on empirically determined basis functions for modal decomposition and can be used to numerically predict the

temperature distribution more rapidly than full-field simulations. POD was first introduced by Lumley [35] in the field of turbulence; Holmes et al. [36] provided a thorough summary for applications of POD in various fields. In this work, POD was partially utilized for multi-scale modeling. As shown in Chapter 2, POD method is capable of predicting transient temperature distribution regardless of the temporal or spatial dependence of the applied heat source [37] . This feature provides the ability to predict temperature distributions for arbitrary heat inputs, by using a smaller sample set of applied heat sources and power maps, resulting in considerably decreased simulation time. One of the limitations of POD is that much time is spent in generating the initial observations for the model. Also, the accuracy of POD depends on the accuracy of these observations. A detailed review on the history and fundamentals of POD is provided in Chapter 2.

1.3.2 Top-down Approaches

Top-down based approaches are another category of multi-scale thermal modeling is microelectronic industry.

A more recent top-down approach is recognized as behavioral thermal modeling method; which is a combination of the generalized pencil-of-function (GPOF) [38, 39] and sub-space methods. [40, 41]. GPOF was developed in the communication community to estimate poles of an electromagnetic system by solving a generalized eigenvalue problem. These methods are mainly used for high-performance multicore micro-processor design. Nonetheless, sub-space methodologies in general potentially suffer from a lack of predictability problems.

Progressive zoom-in method is another top-down based approach to model transient thermal problems. It integrates package and chip level analyses, acquiring the advantages of each (e.g. [42]). It also has the capability of covering several orders of magnitude in length and time scale [43]. It is however, boundary condition dependent. Progressive zoom-in approach was also utilized in this work, Chapter 3 and 4. In this technique, a simplified numerical model is developed for the entire package, which computes the transient temperature distributions and heat fluxes on the top and bottom surfaces of the embedded chip. Then, by applying the calculated top and bottom wall temperature and heat fluxes distributions at each time step as boundary conditions for the chip, a full numerical study can be performed to determine the detailed temperature fields inside the chip. A detailed description of this method is provided in Chapter 3 and 4.

1.3.3 Micro and Nanoscale Thermal Characterization Techniques

Various techniques have been developed and reported in literature for steady state and transient thermal characterization of the modern microelectronic devices. These methods have different application based on their spatial resolutions, thermal time constants, suitability for embedded structures, and ability to be integrated with the device. Table 5-1 provides a summary of the common techniques utilized for high resolution thermal measurement in microelectronics. A detailed review of thermometry and thermal transport in micro and nanoscale devices is provided in Chapter 5. A shortened review of these techniques is provided a below:

One of the most regularly used techniques for non-contact thermal imaging is infrared (IR) microscopy. IR microscopy operated based on the IR emission of the

surface of the sample under test. Photons emitted from the surface are focused on the quantum detector of the IR microscope and ultimately create an electrical signal in the detector. Based on the fact that infrared radiation is temperature dependent, the electrical signal in the detector is processed to determine temperature. The spatial resolution of IR microscope can go up to 3 μm for steady state and 30 μm for transient measurements. The thermal time constant of transient thermal measurements can be as small as 1 μs [44].

Another category of thermal characterization in nanoscale is Atomic Force Microscope (AFM) based approaches. Scanning Thermal Microscopy (SThM), [45] and Scanning Joule Expansion Microscopy (SJEM) [46] are amongst these techniques. In SThM method, a very small thermocouple is fabricated on the tip of an atomic force microscope (AFM) [47-50] that can provide a maximum resolution of 50 nm. SThM has a high spatial resolution, on smooth objects, and high bandwidth. However, the heat transfer between the probe and sample depends on the tip contact with the sample, which can vary with sample hardness, wear, or contact force [51]. In addition, the interface quality can significantly affect the thermal transport and hence precise calibration on similar surfaces is required.

SJEM is a powerful technique to extract in-plane thermal conductivity of thin metallic films whose thickness is comparable to the electron mean free path. Figure 6-3, adapted from [46], displays the schematic diagram for an experimental setup used for SJEM. This method measures the periodic thermal expansion amplitude at the sample surface, which corresponds to the periodic temperature at the surface. The main disadvantage of this method is that it can only measure an AC temperature change.

Moreover, the amplitude of expansion is highly dependent on the heating frequency, the underlying layers dimensions, and thermal properties. Therefore, a minor mismatch in the coefficient of thermal expansion (CTE) can introduce significant error in measurements.

Other common optical thermometry techniques are known as Raman and micro-Raman thermal imaging method. Raman thermometry operates based on the phenomena that the energy of the scattered photons from a structure is different than that of the incident photons due to the inelastic scattering and the exchange of energy with lattice vibrations in the material itself. As the temperature increases, the number of phonons in the excitation mode escalates which will increase the ratio between the anti-Stokes and Stokes peaks. This ratio is used in Raman thermometry to determine temperature [52]. In addition, the shift in the Raman frequency as a function of temperature can be utilized to calculate temperature.

As stated above, Chapter 5 provides a detailed review of thermometry and thermal transport techniques in micro and nanoscale.

1.3.4 Measurement of Thin Metallic Film Properties

Several techniques have also been utilized for the measurement of thin film thermal properties, more specifically, thermal conductivity for thin metal layers. Some of the well-established measurement techniques are described briefly below. A thorough review on various micro and nanoscale thermal characterization techniques is provided in Chapter 6.

There have been several reported techniques under steady state condition. Nath and Chopra [53] originally introduced techniques to measure thermal conductivity of metallic thin films. The outline of their setup is presented in Figure 6-1 adapted from

[53]. They used both steady state and transient measurements. They utilized a U-shape copper block and a heating element representing a heat source. The lead sheet at the other end of the structure is considered to be the heat sink with thermocouples to measure the temperature rise. They modeled the metal film and substrate with a one dimensional heat diffusion equation. Thermal conductivity was determined by using a bare substrate and a metal deposited substrate. Similarly, Pompe and Schmidt [54] established a steady state measurement technique for thermal conductivity measurements with different heat source and heat sink.

Later on, Shojaei-Zadeh et al. [55], Zhang et al. [56], and Liu et al. [57] utilized a suspended micro-fabricated metal bridge for thermal conductivity measurements. As shown in Figure 6-2 from [57], Liu et al., for instance, measured the lateral thermal conductivity of thin copper layers of thicknesses 50 and 144 nm at temperatures between 40 and 400 K, caused by Joule heating, using electrical-resistance thermometry.

Nevertheless, there are limitations to the discussed steady state approaches. Almost all of these methods require a suspended metal bridge or a combination of a metal and low-K material bridge that is constructed through significant micro-fabrication. Another drawback is that by using a suspended metal bridge these structures do not account for the interfacial region. In addition, using fabrication techniques, such as etching, to undercut the substrate can potentially change the quality of the interface.

Transient experiments have also been reported in literature. A transient approach to measure the thermal conductivity of the metallic thin films was introduced by Kelemen [58]. In their proposed technique, a pulsed heat source was applied at one end of the film

and the temperature was measured at two points along the film. Using a one dimensional heat diffusion model, they were able to determine the thermal conductivity.

Amongst the established technique to measure thermal conductivity of bulk substrates and thin films is the 3ω method originally developed by Cahill [59]. Lu et al. [60] implemented the 3ω technique to measure specific heat and thermal conductivity of suspended thin platinum wires. Yang and Asheghi [61] further extend this method such that the suspension of the wire was no longer required. Nevertheless, to reduce the heat conduction to the substrate, the underlying silicon dioxide had to be etched away from the sides and yet an extensive three dimensional numerical analysis was performed to account for the remaining substrate.

The main disadvantage of the aforementioned transient techniques is their inability to measure thermal conductivity of sub 100 nm embedded metallic lines with the exception of Yang and Asheghi [61]. Their technique also has some limitations. Yang and Asheghi [61] reported a very low sensitivity due to the effect of substrate in spite of their extensive three dimensional numerical analysis that is required for the model. To obtain the best sensitivity, the length of the interconnect needs to be only a few microns. Similarly, the interfacial defects along the interconnect were not taken into account.

The final category is high spatial resolution temperature measurement approaches that are mainly used to characterize thermal transport in nanoscale interconnects. Hence, they are commonly shared with the previous section “1.3.4 Measurements of Thin Metallic Film Properties”. Among these are the Pump-Probe Transient Thermoreflectance (PPTTR) and Joule Expansion Microscopy (SJEM) explained earlier. The (PPTTR) technique was proposed by Paddock and Eesley [62]. PPTTR has the

ability to differentiate between the thermal conductivity of thin films and their interface thermal resistance [63, 64]. It is important to recognize that both PPTTR and SJEM are stand-alone measurement tools that require extensive setting up procedures.

The limitations associated with these methodologies justify the need for a thermal conductivity measurement technique that can be integrated within the structure and accounts for the interfacial effects. As mentioned above, a thorough review of the micro and nanoscale thermal characterization techniques is provided in Chapter 6.

1.4 Overall Research Contributions

In this work, numerical and experimental methods were developed to address transient Joule heating in nanoscale embedded on-chip interconnects which can be summarized as:

- A multi-scale reduced order transient thermal methodology was developed which has the capability of handling several decades of length scale from tens of millimeter at "package" level to several nanometers at "interconnects" level. This ability also applies for time scales from seconds to microseconds corresponding to various transient thermal events, at lower computational cost, while maintaining satisfactory accuracy. The proposed method also provides the ability to rapidly predict thermal responses under different power input patterns, based on only a few representative detailed simulations, without compromising the desired spatial and temporal resolutions. It is demonstrated that utilizing the proposed model, the computational time is reduced by at least two orders of magnitude at every step of modeling.

- Additionally, a new platform was developed to evaluate rapid transient Joule heating in embedded nanoscale metallic films such as on-chip interconnects that are not directly accessible. The effect of rapid transient power input profiles with different amplitudes and frequencies in Cu interconnects were studied using sub-micron resistance thermometry technique. Utilizing this technique, a spatial resolution of 6 μm was achieved. It was demonstrated that the transient thermal measurement can be obtained with good accuracy input power fluctuations up to 95 kHz.
- Ultimately, a state of the art technique to extract thermal conductivity of embedded nanoscale interconnects were developed. The proposed structure is the first device and technique that has enabled the conductivity measurement of embedded metallic films on a substrate. It accounts for the effect of the substrate and interface while having a satisfactory sensitivity to the thermal conductivity of the metallic film. The proposed technique can be integrated within the structure and be used for measurements of embedded or buried structures such as nanoscale on chip interconnects. And lastly, it doesn't require extensive micro-fabrication.

1.5 Thesis Outline

The overall description of these contributions is provided below. Figure 1-2 also shows the flowchart of the techniques developed in this work. Chapters associated with each step are also displayed in the flowchart.

In Chapter 2, a two dimensional (2-D) reduced order modeling approach based on Proper Orthogonal Decomposition (POD) implementing the Galerkin projection

technique was developed to study the transient Joule heating in interconnects. In this model, the boundary conditions were assumed to be thermally insulated to represent embedded interconnects that are buried in the bulk of a microelectronic device and receive no cooling. The effect of different types of current pulses, pulse duration, and pulse amplitude were investigated utilizing the proposed model. It was observed that the POD model can predict the transient temperature distribution, regardless of the temporal dependence of the heat source. To validate this distinctive capability, an analytical proof in 2-D was developed and is described in Appendix A. This feature of the proposed model provides a predictive capability based on a smaller set of data, which can significantly decrease the computational cost for various transient forcing functions.

In Chapter 3, to improve the capabilities of the developed model and to further reduce the computational cost, a multi-scale reduced order transient thermal methodology called *hybrid scheme* was developed which incorporates three dimensional POD technique into another multi-scale modeling approach called *"Progressive Zoom-in"* method [65]. This integration resulted in a model that has the capability of handling several decades of length scale from tens of millimeter at "package" level to several nanometers at "interconnects" level. This ability also applies for time scales from seconds to microseconds corresponding to various transient thermal events, at a significantly lower computational cost, without compromising the desired accuracy. This model was developed for low power portable systems, where heat sinks and forced cooling are not employed owing to the compact form factor.

In Chapter 4, the previously proposed hybrid scheme method was further advanced to address the transient thermal problem in a packaged high power

microprocessor where, in fact, force convection plays a key role in the thermal transport of the structure. Another extension to the case studied in Chapter 3 is the implementation of a realistic highly spatially resolved power map used for Intel Core 2 Duo Penryn processor. In addition, to resemble realistic power variations of microprocessors in function, two types of transients were chosen. The thermal time constants associated with these two thermal scenarios differ by two orders of magnitude.

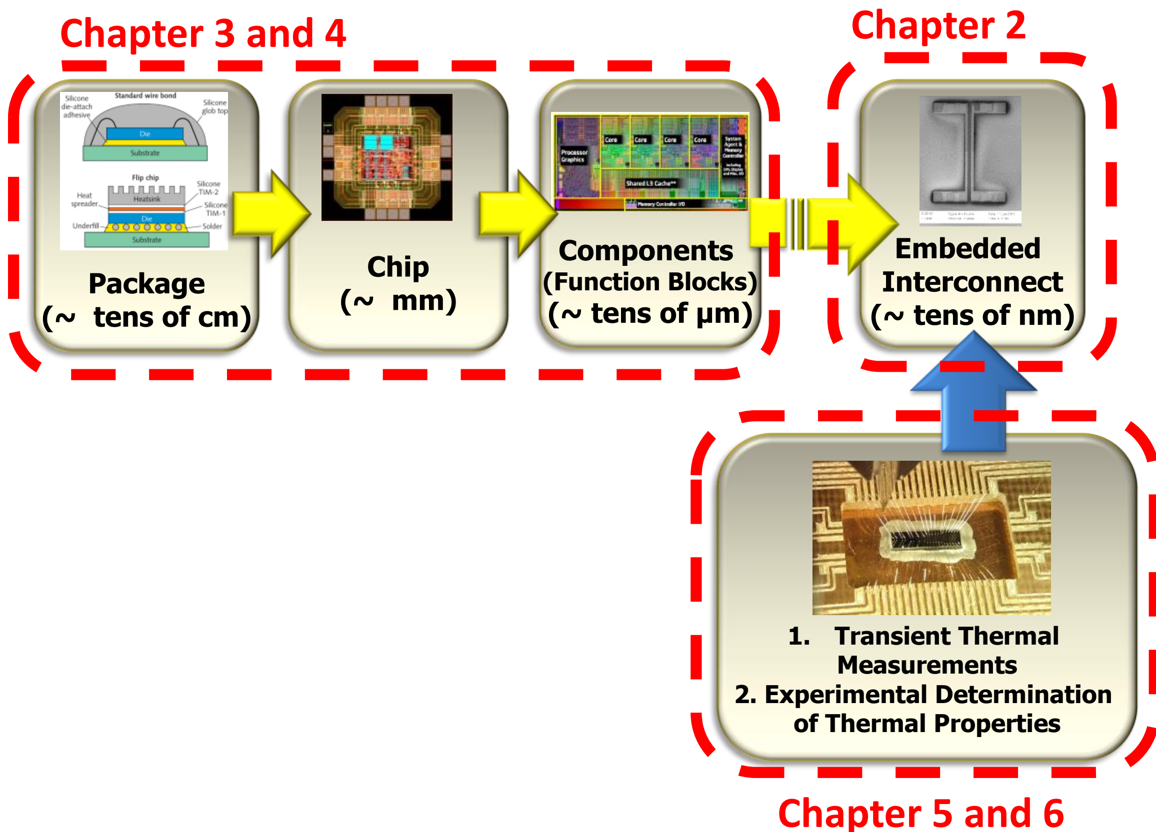

Figure 1-2 The overall flowchart of numerical and experimental techniques developed in this work to address transient Joule heating in nanoscale embedded on-chip interconnects.

In Chapter 5, an experimental platform was developed to study rapid transient Joule heating in embedded nanoscale metallic films representing buried on-chip interconnects that are not directly accessible. Utilizing sub-micron resistance

thermometry (RTD) technique the effect of rapid transient power input profiles with different amplitudes and frequencies were studied. It is also demonstrated in Chapter 5 that a spatial resolution of 6 μm and thermal time constant of 9 μs were achieved using this technique. By obtaining the frequency response of the system, it was shown that the RTDs can follow input power fluctuations up to 95 kHz and therefore the transient thermal measurement can be obtained with good accuracy for this range of input frequencies. The measurements were validated against other measurement techniques such as IR microscopy (only at steady state condition) and numerical simulations.

In Chapter 6, a new concept was proposed to determine the thermal conductivity of embedded nanoscale interconnects. This novel approach induces strong sensitivity of the heating structure to the thermal conductivity of the metallic layer. This technique utilizes a laterally varying resistor structure to produce lateral heat gradient and to induce lateral heat diffusion in the plane of the metallic layer. Ultimately, through steady-state Joule heating and electrical resistance thermometry the thermal conductivity of the embedded metallic films can be identified. Another advantage of the proposed technique is that it does not require extensive micro-fabrication. The size effect on thermal conductivities of 112 nm and 60 nm embedded Cu films were studied. The average values of thermal conductivity at room temperature were found to be 216.14 W/mK and 160.50 W/mK for the 112 nm and 60 nm films respectively. Additionally, the dependence of the thermal conductivity on temperature was also investigated. The experimentally measured values for thermal conductivity and its dependence on temperature agree well with previous studies on free-standing Cu bridges.

CHAPTER 2: TRANSIENT HEAT CONDUCTION IN ON-CHIP INTERCONNECTS USING PROPER ORTHOGONAL DECOMPOSITION (POD) TECHNIQUE

2.1 Objective

In this chapter, a reduced order modeling approach based on Proper Orthogonal Decomposition (POD) implementing the Galerkin projection technique was developed to investigate the transient Joule heating in interconnects in a two dimensional (2-D) inhomogeneous system. This study considers the cases with insulated boundary conditions corresponding to the regions embedded in the bulk of a microelectronic device. The effect of different types of current pulses, pulse duration, and pulse amplitude were investigated. The developed POD model can predict the transient temperature distribution, regardless of the temporal dependence of the heat source. This feature of the proposed model provides a predictive capability based on a smaller set of POD modes, which can significantly decrease the computational cost for various transient forcing functions. To validate this unique capability, an analytical proof in 2-D was developed and is described in Appendix A.

2.2 System Identification Methods

The identification of linear and nonlinear systems has been broadly used in various fields of research. System identification (SI) method has the ability to build dynamic models from the input and output data sets which can be acquired either experimentally or numerically [66, 67]. It is to be noted that once the model is

constructed and its parameters are determined, the model can be used to predict the behavior of the system for various thermal scenarios. In general, SI methods can be constructed based on the following steps [66-69]:

- First step is to choose a proper input signal (i.e. power input in the field of heat transfer) that covers a wide range of frequencies.
- Secondly, the output error models are utilized. In the third step, the number of inputs and outputs that are used in the model will be determined based on different criteria such as the maximum frequency to which the structure is subjected.
- In the next step, nonlinear optimization techniques are used to minimize the responses of the time series data and the SI results.
- Finally to validate the SI model, the results are compared against the output of the system for input signals that are not utilized in construction the model.

In general, SI methods are either parametric or nonparametric. In the parametric category, the model is constructed using differential equations and the goal is to determine the mathematical parameters of the SI model. The step response and the frequency response of the system, on the other hand, are among the commonly used nonparametric SI approaches. It is shown that POD can be considered as an efficient non-parametric system identification method and for instance, in the field of vibration, it can be utilized to diagnose and monitor the performance of vibrating structural assemblies [70-72].

2.3 History of POD

The history of POD goes back over 100 years [73], when it was used as a means for processing statistical data. Since that time, it has been applied in many engineering fields, including fluid flow and turbulence [74-76], structural vibrations [77, 78], and control theory [79]. More relevant to the theme of this research, POD has been used to analyze microelectromechanical systems (MEMS) and electronic packaging [80, 81]. More recently, it has been applied to transient heat conduction problems [82-84]. Bleris and Kothare [85] studied Microsystems using empirical eigenfunctions obtained from the POD technique to address the problem of thermal transient regulation. A boundary condition independent POD-Galerkin Methodology for 1D heat conduction was studied by Raghupathy et al. [86]. Berkooz et al. [76] provided a thorough summary for applications of POD in various fields. Based on the application, POD can be referred to as Principal Components Analysis (PCA) [87], Singular Value Decomposition (SVD) [88], Karhunen–Lo`eve (KL) decomposition [89], or Hotelling transformation [90]. A summary of the equivalence of these three POD methods and the connections among them have been demonstrated by Liang, et al. [91].

2.4 Fundamentals of POD Method

POD provides an optimal set of empirical basis functions (also known as POD modes) from an ensemble of observations, obtained either experimentally or numerically. They characterize and capture the overall behavior and complexity of a physical system using a reduced number of degrees of freedom. While the determination of the optimal basis requires some computational effort, the overall cost of the simulations is much lower than full-field simulations. POD offers the most efficient method of capturing the

dominant components of an infinite-dimensional process with a finite number of modes [76, 92].

In this technique, data sets are expanded for modal decomposition on empirically determined basis functions, which minimize the least squares error between the true solution and the truncated representation of the POD model. The temperature distribution can be determined from the expansion into the POD modes as:

$$T(x, y, t) = T_0(x, y) + \sum_{i=1}^{m} b_i(t)\varphi_i(x, y) \tag{2.1}$$

where T_0 is the time average of temperature (i.e. the mean vector of the observation matrix), $\varphi_i(x, y)$ is the i-th POD mode, and $b_i(t)$ is the i-th POD coefficient, explained later. The procedure to generate a POD based reduced order model is described below:

2.4.1 Generating the Observation Matrix

The initial step in generating the POD observation matrix, T_{Obs}, is to collect a series of observations (a.k.a snapshots) of temperature distribution at different time instants. The matrix is formed by collecting the temperature values at n instances of time in the entire domain using either a numerical or experimental approach.

Having the ability to utilize experimentally obtained data as the initial observations makes POD a strong candidate to characterize a potentially complex system without generating any numerical model. These initial experimental data can be acquired using thermal sensors within the structure of interest. The spatial accuracy of the POD model will be dependent on the number of sensors placed in the domain. Therefore, based on the desired spatial resolution, proper number of thermal sensors should be utilized.

In this study an FE based model was used to obtain the initial observation matrix. T_0 in Eq. (2.1) is the average of all observed data for any point in the domain. As expected, the accuracy of the POD method depends on the accuracy of the observations. Hence, it is of great importance to perform a grid independence analysis. The other critical factor is to remain above the lowest limit of the number of observations, n, which is problem dependent.

2.4.2 Calculating Basis Functions (POD modes)

Once the observation matrix is produced, the POD modes can be calculated. In Eq. (2.1), m is the number of POD modes used in the decomposition, which can range from 1 to n-1, where n is the number of observations. To determine the POD modes, the method of snapshots is used, where each POD mode is expressed as a linear combination of the linearly independent observations [76].

$$\varphi_i(x, y) = \sum_{k=1}^{n} \beta_{ki}\left(T_{Obs,k} - T_0\right) \tag{2.2}$$

where $T_{\mathrm{Obs,k}}$ is the k-th column of the observation matrix T_{Obs}, corresponding to the full temperature field at the k-th instant of time. As described in [76, 93], each eigenvector of the solution of Eq. (2.3) consists of the weight coefficients β_{i}:

$$\sum \Re\beta = \lambda\beta \tag{2.3}$$

where λ is the matrix of eigenvalues and $\Re \in R^{n\times n}$ is the covariance matrix defined as:

$$\Re = \frac{1}{n}\left(T^{C}{}_{Obs}\right)^{T}\left(T^{C}{}_{Obs}\right) \tag{2.4}$$

where $T^{C}{}_{Obs}$ is the *mean-centered observation matrix* obtained by subtracting its mean vector (here T_0), in order to have a zero-mean for the new matrix.

Having calculated the weight coefficients, β, the n POD modes can be determined from Eq. (2.2). The energy captured by the *i*-th basis function in the problem is relative to its corresponding eigenvalue, λ_i from Eq. (2.3). Sorting these eigenvalues in a descending order results in an ordering of the corresponding POD modes. Therefore, the first POD mode calculated from Eq. (2.2) captures the largest portion of energy relative to the other basis functions.

To determine the truncation degree of the POD method, the *cumulative correlation energy*, E_m, captured by the first *m* POD modes is defined by Bizon et al. [94]:

$$E_m = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{2.5}$$

The number of retained POD modes is quite critical in capturing the physics of the problem. It is shown that an insufficient number of the POD modes can cause significant phenomena not to be detected [95]. On the contrary, taking too many POD modes can produce unexpected behavior, or make the model unstable [96]. To be able to generate a reliable POD model, in the present study the number of POD modes is determined in such a way that the cumulative energy of the modes, calculated from Eq. (2.5), is larger than 99.99%.

In general, the required number of POD modes is different for every problem and varies based on the complexity of the structure. For linear problems, if the method of Galerkin Projection is incorporated, the number of the required POD modes, and consequently the number of terms kept in Eq. (2.1), does not depend on the transient power profile. In other words, since the POD modes for a linear problem are independent

of the power input (Analytical proof in Appendix A), the number of time stamps needed to generate the observation matrix is independent of the thermal time constant of power input and it only depends on the geometry, material, and thermal properties of the structure.

2.4.3 Calculating POD coefficients, bi, via Galerkin Projection method

There are multiple techniques to calculate the POD coefficients b_i for a new test case such as: 1. Direct Interpolation Method [97, 98], 2. Flux Matching Process [99, 100], and 3. Galerkin Projection Method [101]. The Galerkin Projection method is more accurate compared to the other two methods in addressing a time dependent heat source, since it solves the energy equation for the entire time domain. Therefore, this method is used in this study.

The Galerkin Projection method projects the governing equations onto the POD spanned space. When POD modes are used in a Galerkin Projection method, they create a finite-dimensional dynamic system with the smallest possible degrees of freedom (DOF). In this study, the inhomogeneous transient heat conduction equation is a partial differential equation (PDE). This technique converts this PDE to a set of m coupled ordinary differential equations (ODE). The key step in model reduction is to solve a discrete number of coupled ODEs instead of solving a discretized PDE. To further describe the method, we start with the transient heat conduction equation:

$$\frac{\partial T}{\partial t} - \frac{1}{\rho c_p} q'''(t) - \alpha \nabla^2 T = 0 \tag{2.6}$$

where, ρ, c_p, and α are the material density, specific heat capacity, and respectively. $q'''(t)$ is the domain time dependent volumetric heat generation. Eq. (2.6) is then projected onto the space spanned by POD modes:

$$\left\langle \varphi_j, \frac{\partial T}{\partial t} - \frac{1}{\rho c_p} q'''(t) - \alpha \nabla^2 T \right\rangle = 0,\ j = 1,2,\ldots,m \tag{2.7}$$

where $\langle .,. \rangle$ denoted the inner products, also referred to as the projection of a vector to one another. Using the temperature field from Eq. (2.1) and integrating Eq. (2.7) over the entire two-dimensional domain (Ω), we have:

$$\int_\Omega (\varphi_j, \frac{\partial T}{\partial t} - \frac{1}{\rho c_p} q'''(t) - \alpha \nabla^2 T) d\Omega = 0 \tag{2.8}$$

Descritizing Eq. (2.8) and using Eq. (2.1) result in a set of coupled ODEs for the POD coefficients that can be written in a matrix form as:

$$A_{ij}\dot{b}_j(t) - B_{ij} b_j(t) - (c+q)_i = 0, i, j = 1,2,\ldots,m \tag{2.9}$$

where ($\dot{}$) denotes the derivative with respect to time. Coefficients A_{ij}, B_{ij}, c_i, and q_i in Eq. (2.9) are:

$$A_{ij} = \int_\Omega \varphi_j \cdot \varphi_i d\Omega \tag{2.10a}$$

$$B_{ij} = \int_\Omega \alpha \varphi_j \cdot \nabla^2 \varphi_i d\Omega = -\int_\Omega \alpha (\frac{\partial \varphi_j}{\partial x} \cdot \frac{\partial \varphi_i}{\partial x} + \frac{\partial \varphi_j}{\partial y} \cdot \frac{\partial \varphi_i}{\partial y}) d\Omega$$

$$+ \int_y \left(\alpha \varphi_j \cdot \frac{\partial \varphi_i}{\partial x} \right]_{x = x_{\min}}^{x = x_{\max}} dy + \int_x \left(\alpha \varphi_j \cdot \frac{\partial \varphi_i}{\partial y} \right]_{y = y_{\min}}^{y = y_{\max}} dx \tag{2.10b}$$

$$c_j = \int_\Omega \alpha \varphi_j \cdot \nabla^2 T_o d\Omega = -\int_\Omega \alpha (\frac{\partial \varphi_j}{\partial x} \cdot \frac{\partial T_o}{\partial x} + \frac{\partial \varphi_j}{\partial y} \cdot \frac{\partial T_o}{\partial y}) d\Omega$$

$$+\int_y \left(\alpha\varphi_j \cdot \frac{\partial T_o}{\partial x} \right]_{x=x_{\min}}^{x=x_{\max}} dy + \int_x \left(\alpha\varphi_j \cdot \frac{\partial T_o}{\partial y} \right]_{y=y_{\min}}^{y=y_{\max}} dx \qquad (2.10c)$$

$$q_j = \int_\Omega \frac{1}{\rho c_p} \varphi_j \cdot q'''(t) d\Omega \qquad (2.10d)$$

The last two terms on the right hand side of Eq. (2.10b and 2.10c) are the boundary terms. If the boundary conditions are homogeneous or insulation, these are eliminated and B_{ij} and c_i are simplified to:

$$B_{ij} = -\int_\Omega \alpha\left(\frac{\partial\varphi_j}{\partial x} \cdot \frac{\partial\varphi_i}{\partial x} + \frac{\partial\varphi_j}{\partial y} \cdot \frac{\partial\varphi_i}{\partial y}\right) d\Omega \qquad (2.11b)$$

$$c_j = -\int_\Omega \alpha\left(\frac{\partial\varphi_j}{\partial x} \cdot \frac{\partial T_o}{\partial x} + \frac{\partial\varphi_j}{\partial y} \cdot \frac{\partial T_o}{\partial y}\right) d\Omega \qquad (2.11a)$$

The concept of orthogonality was subsequently applied. Having calculated all the coefficients and substituted those into Eq. (2.9), these coupled ODEs can be solved using the sixth-order Runge-Kutta method. Notably, the initial conditions for Eq. (2.9) can be determined by the projection of the POD modes on the initial value for the temperature as:

$$b_j^{\,o} = b_j(t=0) = \left\langle \varphi_j, T(x, y, t=0) \right\rangle = 0,\ j = 1,2,\ldots,m \qquad (2.12)$$

2.4.4 Generating the POD temperature field:

Calculating a sufficient number of POD modes and POD coefficients (b_j and $b_j^{\,o}$) will provide the temperature field from Eq. (2.1) anywhere in the domain and at any instant of time.

2.5 Case Studies

The geometry and topology of on chip interconnects in micro-electronic devices can be quite complex. This study focuses on a simplified but realistic 2D model domain. A single interconnect located at the center of a large array of metal lines is considered as shown in Figure 2-1. By restricting the problem domain to one corner of this interconnect and the surrounding dielectric material, symmetry arguments can be employed to justify insulated boundary conditions on all sides. This idealized configuration represents the worst thermal scenario, in which the interconnect effectively receives no cooling. This structure approximates long and uniformly spaced interconnects. The interconnect has equal width and height; i.e. the geometrical aspect ratio is 1. The dielectric thickness, H_{de}, is equal to the interconnect height ($H_{de} = H_{int} = 360$ nm). Interconnect pitch P is usually a variable and, in this study, was taken to be $P = 4H_{int} = 1.44$ µm. The initial condition was assumed to be room temperature $T_o = 300$ K. The continuum assumption was verified by calculating the Knudsen number, Kn:

$$Kn = \frac{\Lambda}{L} \qquad (2.13)$$

where Λ is molecular mean free path and L is the smallest length scale in the structure [102]. For this structure, the Knudsen number based on the mean free path of electrons in copper at 300 K ($\Lambda = 39$ nm) and the width of interconnects ($H_{int} = 360$ nm) is calculated to be $Kn = 0.108$. Therefore, the continuum approach is valid. The material properties of the metal and dielectric, representing copper and silicon, were: specific heat capacity $C_p = 380$ and 1000 J/kgK, thermal conductivity $\kappa = 400$ and 0.17 W/mK, and density $\rho = 8933$ and 2200 kg/m^3, respectively.

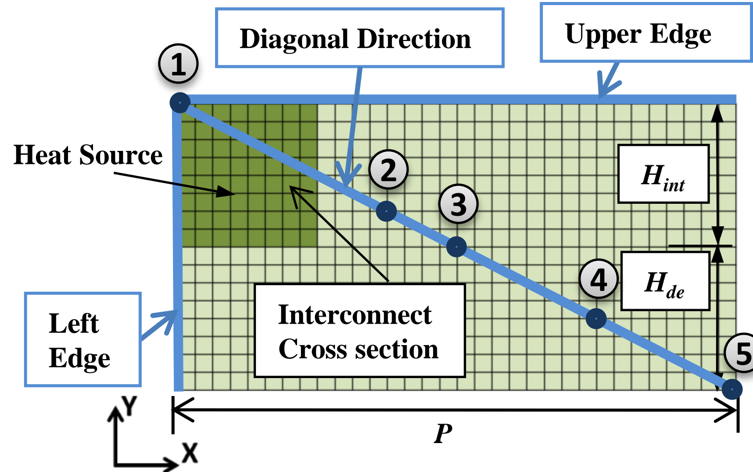

Figure 2-1 Schematic of the computational domain with a cross sectional area of 1.44 µm x 720 nm. It consists of a set of 360 nm x 360 nm interconnects that are evenly spaced and embedded in the dielectric. The mesh used in the POD and FE models is shown. In this study: H_{int} = H_{de}= 360 nm and P = 4H_{int} = 1.44 µm.

Given that no analytical solution exists for this problem, a detailed FE model was developed in the commercial code LS-DYNA and the results were used as a basis for the evaluation of the POD modeling approach. The Crank-Nicholson time integration scheme and conjugate gradient iterative solver are chosen for the transient thermal simulations. The results presented here are for Δt = 10.0 ns. The convergence of the FE model was verified with respect to the solver type, time step, and time integration method. The FE model consists of 561 nodes (17× 33) and 512 elements. The grid size was determined based on the mesh independence analysis. For consistency between the FE and the POD models, the same number of nodes at the same position was chosen for the POD model. The top edge, the left edge, and the diagonal of the structure are used in this study for the spatial thermal analysis (Figure 2-1); nodes 1-5 on the diagonal are chosen for further temporal thermal analysis of the results. The selection was made based on the directions and positions with the largest value of temperature and temperature gradient.

The effect of current pulse type, pulse duration, and pulse amplitude were investigated in this study. By using a representative step function as the heat source POD modes are generated. Using just a few POD modes, the model predicted the exact transient thermal behavior of the system for all other cases with different temporal dependence of the heat source, and without generating any new observations. Furthermore, the result of the POD model was compared with a finite element (FE) model and good agreement was found, with a maximum difference of 2%.

To assess the predictive capability of the POD model, two cases corresponding to different thermal scenarios with different time dependent heat sources were considered. For a better comparison between these cases, a constant value for volumetric heat generation, q_0''', is calculated based on the current density J = 10 MA/cm^2 and the electrical resistivity of ρ_r = 2.2 μΩ-cm:

$$q_0''' = \rho_r J^2 = 2\times10^{13}\ (W/m^3) \tag{2.14}$$

Subsequently, the time dependent heat sources in (W/m^3) for the cases (exhibited in Figure 2-2) are:

Case 1: $q_1''' = q_0'''$ (Step function with the magnitude of 2×10^{13} W/m^3)

Case 2: $q_2''' = 10\times q_0'''\times\left[\sin(\frac{\pi}{2}10^6 t)\ +1\right]$ (Combination of a step function with the magnitude of $q_2''' = 10\times q_0'''$ and a continuous sinusoidal function, with the period of τ_2 = 4 μs and amplitude of $A_2 = 10\times q_0''' = 2\times10^{14}$ (W/m^3)

An unwanted surge in the electrical current is represented by Case 1, while Case 2 represents a condition under which a combination of a sinusoidal noise and a current surge abruptly occur in the interconnect line. The amplitude of the noise in Case 2 is ten

times higher than, and the frequencies different from, those of Case 1. It is of great importance to notice that observations are only generated for Case 1, and the results of the transient thermal behavior of the second case are determined based on the results obtained from Case 1 without generating new observations. This ability of the POD method in predicting thermal behavior of other cases based on a smaller set of modes can significantly decrease the computational cost for the transient analyses. Since the basis functions (~ POD modes) are only dependent on the geometry, the POD modes for any other scenario are the same as for Case 1 as long as the governing equations are linear. This distinctive characteristic of POD is proved in the "Appendix A" and is numerically confirmed next.

2.6 Results and Comparison

2.6.1 Case 1

For $q_1''' = q_0''' = 2\times10^{13}$ (W/m^3), 17 observations of the transient temperature using FE simulation were taken in the first 20 μs and were used to calculate the POD modes. The energy percentage for each POD mode is plotted against the mode number in Figure 2-3. The magnitude of the eigenvalue and the energy captured by each mode reduces with the index of POD modes. By keeping the first five POD modes, the *cumulative correlation energy*, E_m, was greater than 99.99%. The first two modes, alone, capture over 98% of the energy. For the present study, at least 5 modes are needed to capture the desired accuracy of 99.99% in the result. If the number of observations, n, were chosen to be less than the minimum required POD modes (here 5), the results will not have the required accuracy. This is demonstrated in Figure 2-7 for the temperature rise time history at the

top-left-most node (node 1) in Case 1 by using 17, 11, 9, 5, and 2 observations. It can be inferred that as long as the number of the observations are more than 5, the results are independent of the number of observations, while for the case of 2 observations (dash-dotted line) the results are not acceptable.

Figure 2-4 shows the two-dimensional contours of the first five POD modes, normalized based on the total sum of the modes. Using Galerkin Projection, described earlier, the POD coefficients were calculated as functions of time. Figure 2-5 shows the time dependence of the first five *b*-coefficients. It is apparent that the first coefficients vary in the smoothest way and the coefficients with large indices have large fluctuations during the initial stage. It can also be inferred from Figure 2-5 that the value of POD coefficients successively decreases by about an order of magnitude. This shows that only the first few terms in Eq. (2.1) are dominant and need to be included in the calculations.

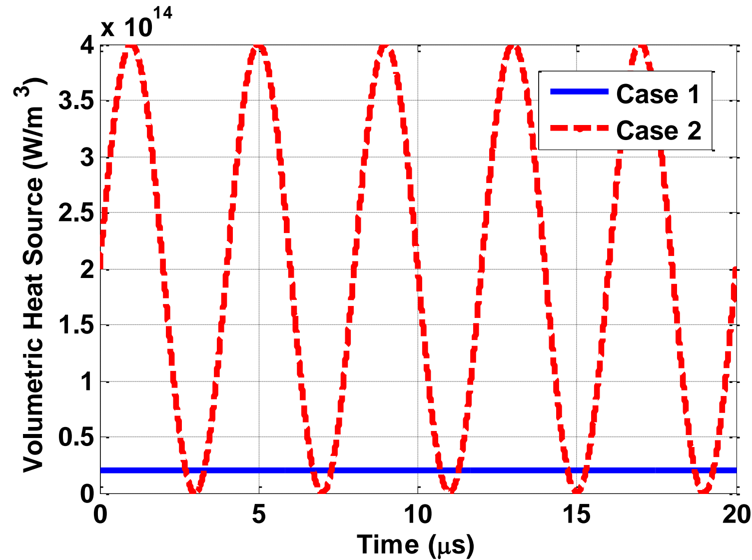

Figure 2-2 Different types of heat sources used in this study. Case 1- only step function (solid line) & Case 2-sinusoidal and step function (dashed line).

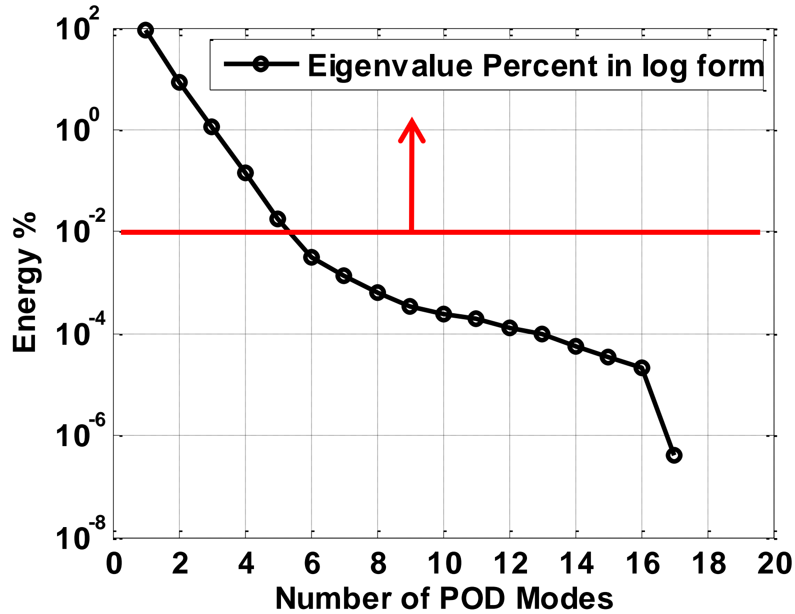

Figure 2-3 Eigenvalues or energy percentage in log form vs. number of the POD modes. In order to build a reliable reduced order model, the number of basis functions used for the projection was chosen such that the cumulative correlation energy of the modes are greater or equal to 99.99%. The first two modes capture over 98% of the energy.

Figure 2-6 shows the contours of the spatial distribution of temperature rise from FE method (top figure) and POD method (bottom picture) after 20 μs for Case 1. Over time, heat diffuses from top left, where the source is located, throughout the entire structure. Considering adiabatic boundary condition at all boundaries, the temperature continuously increases in the domain with time and temperature contours are perpendicular to all the edges. On account of relatively high thermal conductivity and low heat capacity of the copper, temperature gradient within the interconnect cross-section is relatively negligible.

To make a more detailed comparison between the POD and FE results, the temperature rise with time at the top-left-most node (node 1), marked in Figure 2-1, is considered. Figure 2-7 demonstrates this comparison for FE (solid line) and POD results

using 5 modes (dashed line) for Case 1 in the first 20 μs. As shown in Figure 2-7, temperature increases rapidly as the current source is applied. The maximum error between the two models is less than 1%.

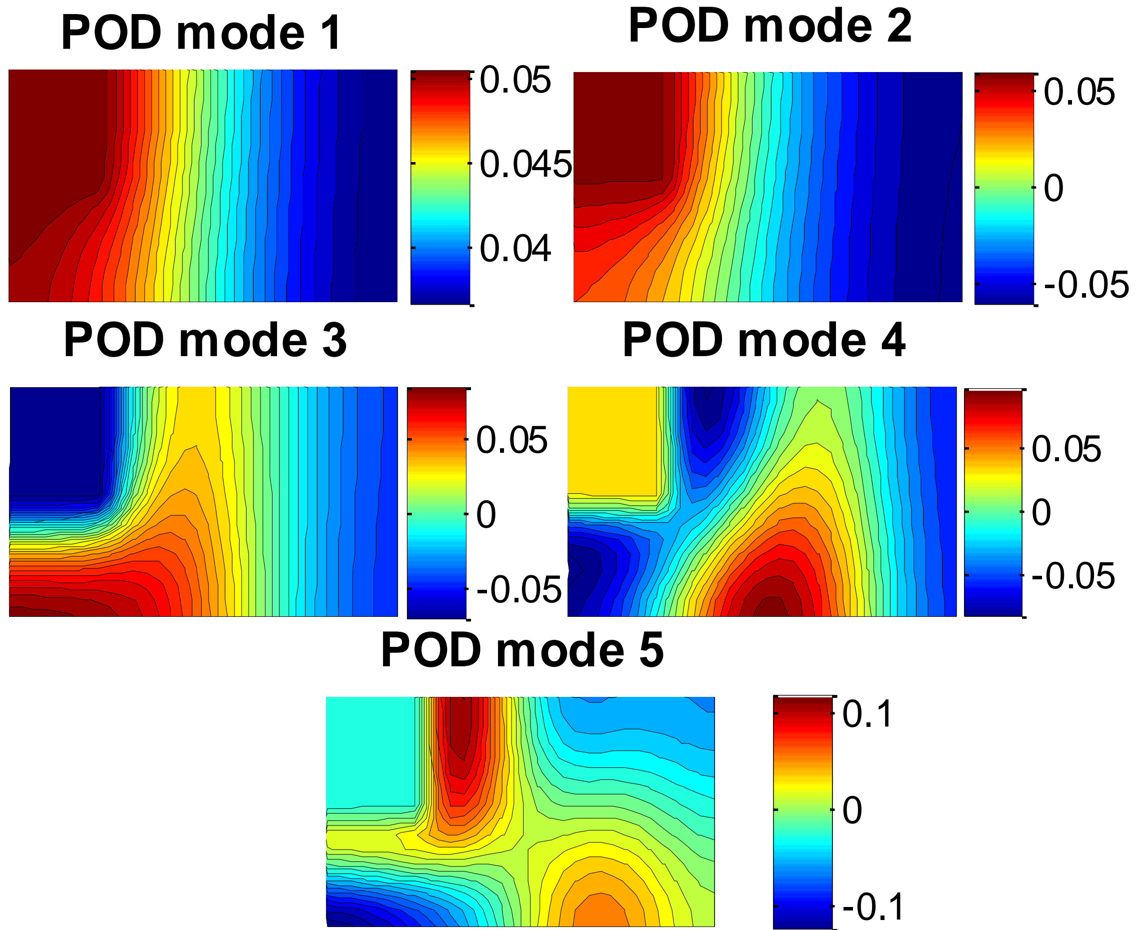

Figure 2-4 First five POD modes, or basis functions, plotted in 2D contours. The POD modes are normalized by the total sum of the modes chosen for each study.

As previously mentioned, to show the distinctive capability of the POD model in the prediction of the transient temperature distribution, two cases were considered; Case 1 required a new set of observations while the second case didn't need any new observations.

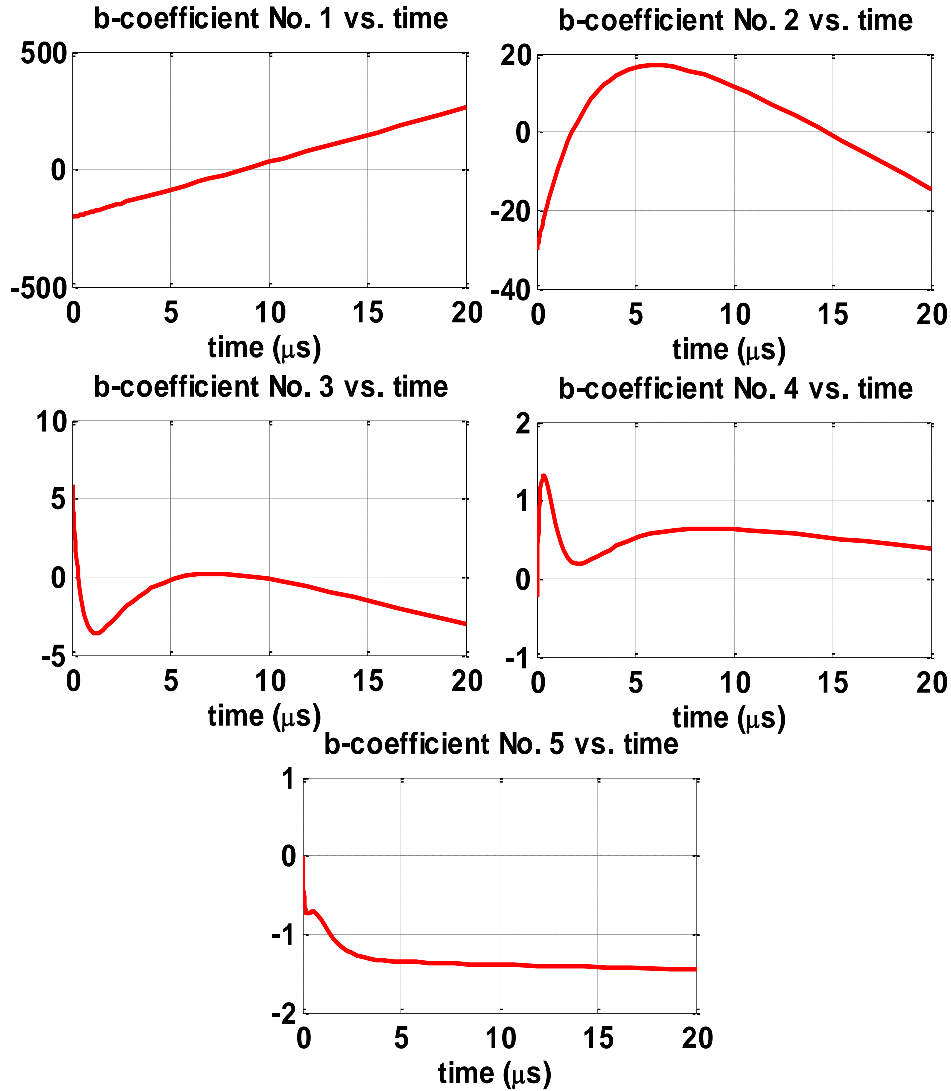

Figure 2-5 First five b-coefficients vs. time using the Galerkin Projection Technique for Case 1.

2.6.2 Case 2

In the second case, the capability of the proposed model in predicting the temperature field for a combination of a transient heat pulse and a continuous oscillating source of noise is investigated. The sinusoidal function representing a volumetric heat source, plotted as dashed line in Figure 2-2, has a period of τ_2 = 4 μs and amplitude of A_2 $= 10 \times q_o''' = 2 \times 10^{14}$(W/m^3). There are no new generated observations for this case. Figure 2-8 shows the temperature rise at several locations for Case 2. FE results and POD results

using 5 basis functions are shown at two different times. Figures 2-8 (a), (c), and (e) for time=18 μs correspond to a time slightly before the first crest in the fifth cycle of the heat source, while Figures 2-8 (b), (d), and (f), at 19.2 μs, simulate a time slightly after the bottom of the sinusoidal curve in the fifth cycle. Figures 2-8 (a-f), also, demonstrate that the difference between the two models decreases in both x and y directions, as distance from the heat source increases (for lower values of x and for higher values of y). The regions on the graphs with negligible temperature gradient correspond to the location of the interconnect. A maximum truncation error of less than 1.5% exists between the two models.

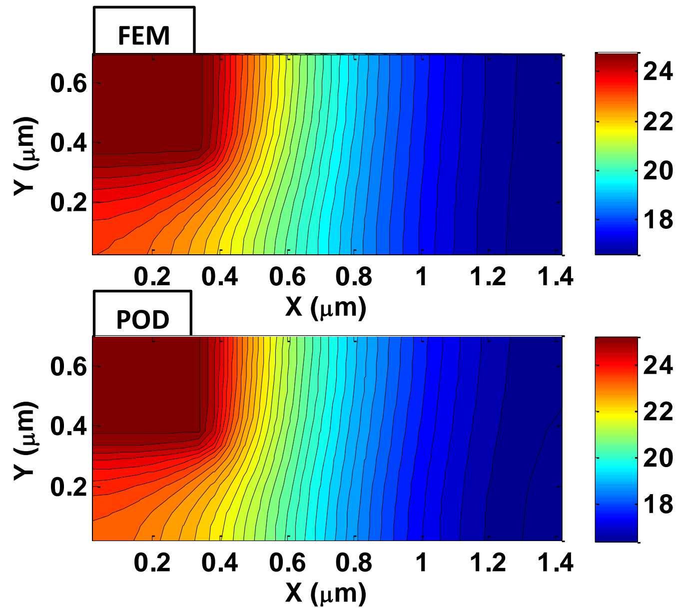

Figure 2-6 Spatial variation of temperature rise after 20 μs for FE (top) and POD (bottom) models using 5 basis functions for Case 1.

A comparison of temperature rise between FE (circular markers) and POD (solid lines) models of nodes 1-5 along the diagonal (noted in Figure 2-1) is provided for Case 2 in Figure 2-9. Based on the results presented for Case 2, it can be interpreted that our POD model using Galerkin projection technique can predict the transient thermal

behavior for a single sinusoidal heat wave. These results further confirm the ability of the POD with Galerkin projection technique to predict the transient thermal behavior of this structure for any temporal dependent heat source, based on a single available observation matrix. Other types of heat sources such as step functions and sinusoidal functions with different frequencies and amplitude were also investigated and verified [103]. This capability of POD is further confirmed through an analytical proof of a 2D transient problem provided in the Appendix A.

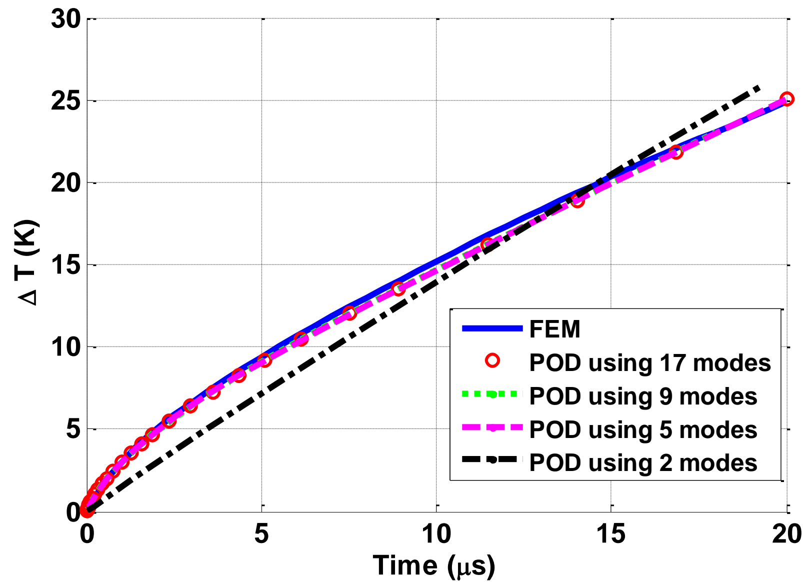

Figure 2-7 Comparison of temporal dependence of temperature rise in the left-most node of the top edge (node 1 in the interconnect), x=0 using 17, 9, 5, and 2 observations for Case 1.

2.7 Conclusion and Discussion

In this chapter, the 2-D Proper Orthogonal Decomposition (POD) method and Galerkin projection technique were implemented to address the transient Joule heating in a two dimensional inhomogeneous arrangement of interconnects embedded in a dielectric material. The POD method characterizes and captures the overall behavior and complexity of a physical system by using a reduced number of degrees of freedom. This results in a much lower computational cost than a full-field simulation method. The most

remarkable characteristic of the POD is its optimality. Data sets are expanded for modal decomposition on empirically determined basis functions in a way that minimizes the least square error between the true solution and the truncated representation of the POD model. This makes the POD method one of the most efficient methods of capturing the dominant components of a large-dimensional system with a finite number of modes.

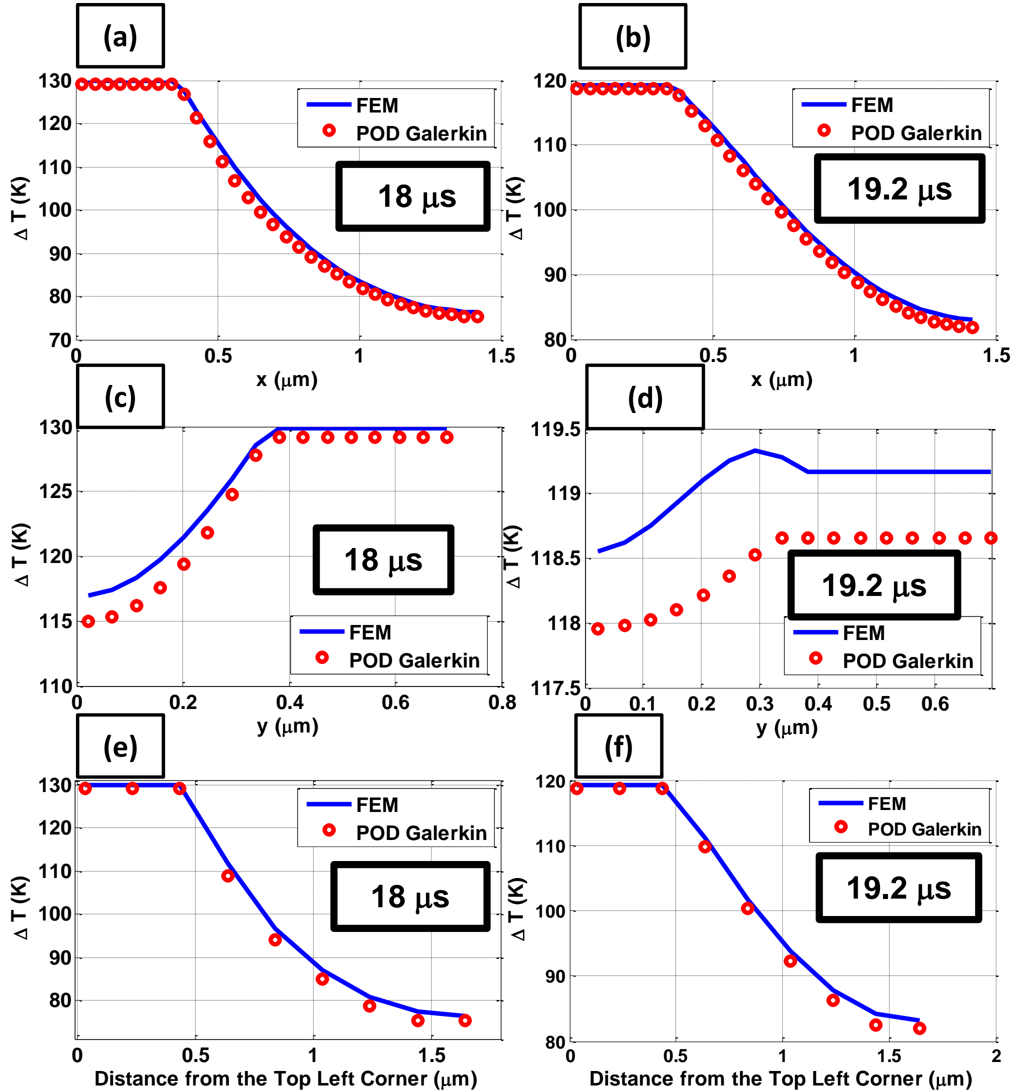

Figure 2-8 Spatial variation of temperature rise after 18 μs in x direction for the upper edge of the structure (a), y direction for the left edge of the structure (c), and along the diagonal (e). Spatial variation of temperature after 19.2 μs in x-direction for the upper edge of the structure (b), y-direction for the left edge of the structure (d), and along the diagonal (f). FE results are plotted in solid lines and POD results using 5 basis functions are plotted in circular markers. The results are for Case 2.

Adiabatic boundary conditions were considered for the 2-D structure, representing the worst thermal scenario in regions of microelectronic devices where the interconnects in fact receive no cooling. A sufficient number of POD modes can be easily estimated which will contain the most energy. The POD modes are obtained at the system level using the observations from finite element (FE) model. The number of POD modes kept in the analysis is determined in such a way that the cumulative energy of the modes was larger than 99.99% of the total energy. The POD coefficients were subsequently calculated using the method of Galerkin projection.

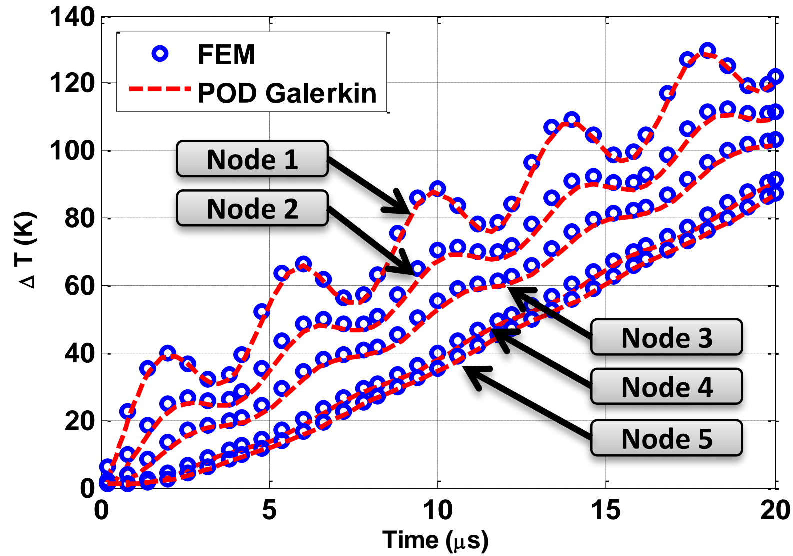

Figure 2-9 Comparison of temporal dependence of temperature rise at nodes 1-5 along the diagonal for Case 2. FE results are plotted in circular markers and POD results are shown by dashed lines.

To assess the POD predictions, two time dependent heat source conditions were considered. In both cases, the POD model predictions were in good agreement with the corresponding FE models. The truncation errors calculated based on the difference of the POD and FE model were found to be less than 2%. The results show that the truncation error does not increase as the amplitude of the heat source increases. The POD modes are

not sensitive to the temporal dependence of the heat source. This important feature can drastically decrease computational cost, making POD a fast and robust method for reduced order modeling of transient heat conduction in microelectronic devices.

CHAPTER 3: MULTI-SCALE TRANSIENT ANALYSIS OF MICROELECTRONICS USING HYBRID SCHEME APPROACH

3.1 Introduction

In the preceding chapter, a two-dimensional (2D) transient heat conduction framework was developed to analyze inhomogeneous domains, using a reduced order modeling approach based on Proper Orthogonal Decomposition (POD) and Galerkin projection. As discussed, POD modes are generated by using a representative step function as the heat source. It was also verified that model rapidly predicted the transient thermal behavior of the system for several cases, without generating any new observations, and using just a few POD modes.

In order to further improve this technique and to reduce the computational cost even more, a multi-scale reduced order transient thermal methodology called *hybrid scheme* was developed which incorporates three dimensional POD technique into another multi-scale modeling approach called *"Progressive Zoom-in"* approach [65]. The proposed model has the capability of handling several decades of length scale from tens of millimeter at "package" level to several nanometers at "interconnects" level. This ability also applies for time scales from seconds to microseconds corresponding to various transient thermal events, at a considerably lower computational cost, while maintaining satisfactory accuracy. Hybrid scheme also provides the ability to rapidly predict thermal responses under different power input patterns, based on only a few

representative detailed simulations, without compromising the desired spatial and temporal resolutions.

In this chapter, an FCBGA package with an embedded die was considered for thermal modeling using a hybrid scheme. Random dynamic power distributions were considered for the total chip power, as well as for the function blocks that compose the entire chip to demonstrate the capability of this method in predicting different thermal scenarios. To validate this methodology, the results were compared with a finite element (FE) model developed in COMSOL®. The results of the model developed through hybrid scheme were in good agreement with the corresponding FE models.

3.2 Hybrid Scheme for Multi-scale Thermal Modeling

A hybrid scheme is developed in this chapter which combines the implementation of POD technique and progressive zoom-in approach, as summarized below:

3.2.1 Fundamentals of POD Method

As explained in the previous chapter, POD is a methodology that offers an optimal set of basis functions, also known as POD modes, which are empirically determined from an ensemble of observations. These observations are obtained either experimentally or from numerical simulation (this study). In this technique, the three dimensional (3-D) temperature distribution is determined from the expansion:

$$T(x,y,z,t) = T_0(x,y,z) + \sum_{i=1}^{m} b_i(t)\varphi_i(x,y,z) \tag{3.1}$$

where T_0 is the time average of temperature (i.e., the mean vector of the observation matrix), $\varphi_i(x, y, z)$ is the i-th POD mode, and $b_i(t)$ is the i-th POD coefficient [37]. A detailed procedure to generate a 2-D POD based reduced order model is provided

in the Chapter 2. Therefore, only the primary steps to generate a POD based reduced order model are outlined below including the POD coefficients derived for a 3-D analysis.

- **Generating the observation matrix**

This step is explained in details in Chapter 2, section 2.3.1.

- **Calculating basis functions (POD modes)**

This step is also explained thoroughly in Chapter 2, section 2.3.2.

- **Calculating POD coefficients, b_i**

As demonstrated in Chapter 2, the POD coefficients, b_i, can be determined by solving the descritized matrix of coupled ordinary differential equations (ODEs) (Eq. (2.9)) using the sixth-order Runge-Kutta method shown below:

$$A_{ij}\dot{b}_j(t) - B_{ij}b_j(t) - (c+q)_i = 0, i, j = 1,2,...,m \tag{2.9}$$

Coefficients A_{ij}, B_{ij}, c_i, and q_i in Eq. (2.9) were derived and presented for 2-D POD model in the previous chapter. For this study, coefficients in Eq. (2.9) were determined for 3-D analysis as:

$$A_{ij} = \int_{\Omega} \varphi_j \cdot \varphi_i d\Omega \tag{3.2a}$$

$$B_{ij} = \int_{\Omega} \alpha\varphi_j \cdot \nabla^2\varphi_i d\Omega = -\int_{\Omega} \alpha\left(\frac{\partial\varphi_j}{\partial x}\cdot\frac{\partial\varphi_i}{\partial x} + \frac{\partial\varphi_j}{\partial y}\cdot\frac{\partial\varphi_i}{\partial y} + \frac{\partial\varphi_j}{\partial z}\cdot\frac{\partial\varphi_i}{\partial z}\right)d\Omega$$

$$+\int_x \left(\left[\alpha\varphi_j \cdot \frac{\partial\varphi_i}{\partial y}\right]_{y=y_{\min}}^{y=y_{\max}} dx + \int_y \left[\alpha\varphi_j \cdot \frac{\partial\varphi_i}{\partial x}\right]_{x=x_{\min}}^{x=x_{\max}} dy + \int_z \left(\alpha\varphi_j \cdot \frac{\partial\varphi_i}{\partial z}\right]_{z=z_{\min}}^{z=z_{\max}} dx \tag{3.2b}$$

$$c_j = \int_{\Omega} \alpha\varphi_j \cdot \nabla^2 T_o d\Omega = -\int_{\Omega} \alpha\left(\frac{\partial\varphi_j}{\partial x}\cdot\frac{\partial T_o}{\partial x} + \frac{\partial\varphi_j}{\partial y}\cdot\frac{\partial T_o}{\partial y} + \frac{\partial\varphi_j}{\partial z}\cdot\frac{\partial T_o}{\partial z}\right)d\Omega$$

$$+\int_x \left(\alpha\varphi_j \cdot \frac{\partial T_o}{\partial y} \right]_{y=y_{\min}}^{y=y_{\max}} dx + \int_y \left(\alpha\varphi_j \cdot \frac{\partial T_o}{\partial x} \right]_{x=x_{\min}}^{x=x_{\max}} dy + \int_z \left(\alpha\varphi_j \cdot \frac{\partial T_o}{\partial z} \right]_{z=z_{\min}}^{z=z_{\max}} dz \tag{3.2c}$$

$$q_j = \int_\Omega \frac{1}{\rho c_p} \varphi_j \cdot q'''(t) d\Omega \tag{3.2d}$$

The last two terms on the right hand side of Eq. (3.2b and 3.2c) are the boundary terms. If the boundary conditions are homogeneous or insulation, these are eliminated and B_{ij} and c_i are simplified to:

$$B_{ij} = -\int_\Omega \alpha(\frac{\partial \varphi_j}{\partial x} \cdot \frac{\partial \varphi_i}{\partial x} + \frac{\partial \varphi_j}{\partial y} \cdot \frac{\partial \varphi_i}{\partial y} + \frac{\partial \varphi_j}{\partial z} \cdot \frac{\partial \varphi_i}{\partial z}) d\Omega \tag{3.3a}$$

$$c_j = -\int_\Omega \alpha(\frac{\partial \varphi_j}{\partial x} \cdot \frac{\partial T_o}{\partial x} + \frac{\partial \varphi_j}{\partial y} \cdot \frac{\partial T_o}{\partial y} + \frac{\partial \varphi_j}{\partial z} \cdot \frac{\partial T_o}{\partial z}) d\Omega \tag{3.3b}$$

- **Generating the POD temperature field**

A sufficient number of POD modes and POD coefficients need to be calculated, which can then be used in Eq. (3.1) for the determination of the temperature field anywhere in the domain and at any instant of time. To determine the truncation degree of the POD method, the *cumulative correlation energy*, E_m, captured by the first m POD modes is defined by Bizon et al. [94] and explained in Chapter 2, Eq. (2.2):

$$E_m = \frac{\sum_{i=1}^{m} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \tag{2.2}$$

3.2.2 Progressive Zoom-in Approach

The Progressive zoom-in method integrates package and chip level analyses, acquiring the advantages of both (e.g. [42]). Figure 3-1 shows a flowchart of the

approach used in this study for multi-scale transient thermal modeling of a representative FCBGA package. The overall hybrid approach is outlined below:

3.3 The Hybrid Scheme Procedure

- **Thermal simulation at the package level**

The first step is to model the entire structure, i.e., the package, including the surrounding mold, underfill, solder bumps, substrate, etc. This simulation is performed in COMSOL. It is important to note that at this level, the chip is modeled as a solid block with effective material and thermal properties without considering internal details.

- **Applying POD technique to package level**

Once the temperature distribution at package level was determined, a POD model was developed. The POD model provides the ability to predict dynamic temperature distribution for different power maps and types of power sources without developing any further full finite element models which can significantly decrease computational cost and potentially be used to define a criterion for the optimal distribution of the current density in the domain.

- **Transferring the solution from package level to the chip level**

Once the temperature distribution at the package level was obtained, a combination of temperature and heat flux at the top and bottom walls of the chip was extracted and interpolated for higher spatial resolution. These data were then applied as boundary conditions for the chip level simulation.

- **Chip level thermal simulation**

At this level, chip is no longer treated as a solid block. It was divided into subdomains called function blocks. Each block represents a specific component with

unique functionality on the chip and consists of three sub-layers: 1. Top Si layer, 2. Middle device layer, and 3. Interconnect/Dielectric multilayer. Function blocks were simulated based on the assigned power generation and calculated effective material/thermal properties for each layer within that block. At this level of thermal simulation, the spatial resolution is limited to the sub-layers. Once the chip is divided into subdomains, the power map needs to be determined at any instance of time for each individual function block.

- **Continue to the desired resolution on the chip**

This method can be continued to multiple levels such that the desired spatial resolution on the chip is achieved. Only representative results for two steps (package and chip level) are presented in this chapter.

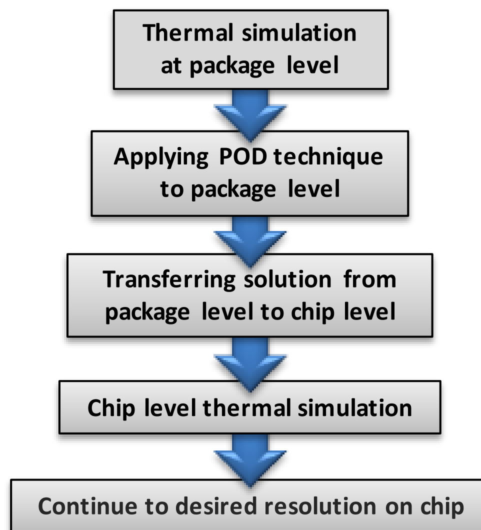

Figure 3-1 Flowchart of the hybrid scheme for multi-scale thermal modeling.

3.4 Results and Discussion

Figure 3-2 (a) shows the schematic of the simplified Flip Chip Ball Grid Array (FCBGA) package used in this study for the package level modeling. This model is for low power portable systems, where heat sinks and forced cooling are not employed due

to the compact form factor. As described in Section 2, the first step is to model the package for which the material properties and dimensions are required. Table 3-1 lists these for the die, solder bumps, underfill, mold, and substrate. These values were mainly provided by Mentor Graphics Corporation and the rest were chosen based on Ref. [104]. Ref. [105] is used as a guideline for the dimensions of the FCBGA package. Due to the fact that solder bumps are embedded in the underfill layer, effective density and specific heat are calculated based on volume averaging. It is assumed that 60% of the surface area between the die and substrate is covered with underfill and 40% is solder bumps. The effective vertical ($K_{v_{eff}}$) and horizontal ($K_{h_{eff}}$) thermal conductivity values are calculated based on thermal resistor network formulation:

$$K_{h_{eff}} = \frac{1}{\left(\frac{\forall_U}{\forall_{tot}}\frac{1}{K_U}+\frac{\forall_S}{\forall_{tot}}\frac{1}{K_S}\right)} \tag{3a}$$

$$K_{v_{eff}} = \left(K_U\frac{A_{hU}}{A_h} + K_S\frac{A_{hS}}{A_h}\right) \tag{3b}$$

where $\forall_{tot}$ is the entire volume, $\forall_U$ and $\forall_S$ are volumes of underfill and solder bumps, respectively. Similarly, A_{hU} and A_{hS} are the cumulative horizontal cross sectional areas of the underfill and solder bumps. K_U and K_S are the thermal conductivities of the underfill and solder bumps respectively. Considering that the solder bumps are vertically orientated, it is expected that the vertical effective thermal conductivity of the effective underfill layer will be significantly higher than its horizontal value. The computed values are 20.2 and 1.47 W/mK respectively. The package dimensions, also listed in Table 3-1, were provided by Mentor Graphics Corporation.

Natural convection boundary condition was imposed on the top surface and vertical boundaries of the package with a heat transfer coefficient of h=15 W/m^{2}K,

chosen to be in the typical range for air [42]. A constant temperature boundary condition was applied to the bottom surface. The initial temperature and the surrounding temperature were assumed to be equal to the room temperature T_{amb}= 300 K.

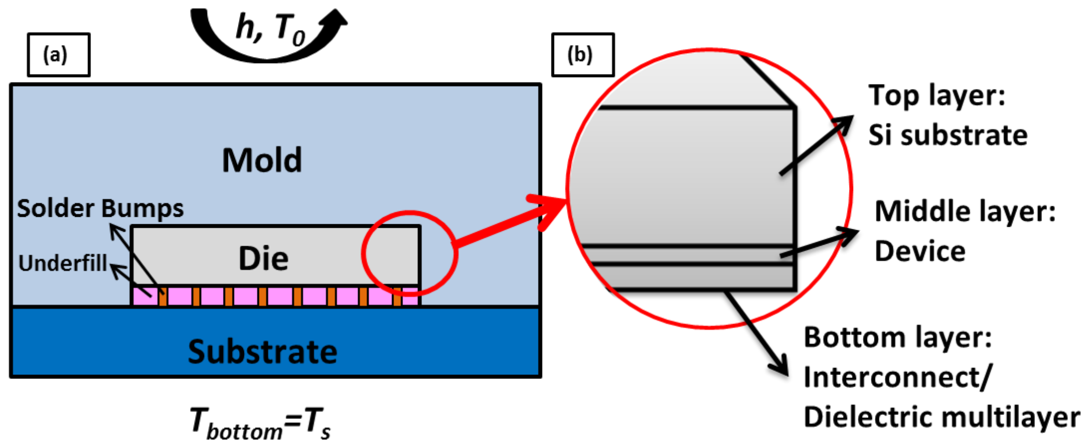

Figure 3-2 (a) Schematic of a simplified FCBGA for package level modeling, (b) Zoomed-in schematic of the die layer used in chip level modeling.

A detailed finite element (FE) model was developed in COMSOL using a time step of dt = 0.05 s. The convergence of the FE model was verified with respect to the solver type, time step, and time integration method. The FE model of the package consists of 75,919 elements, of which 343 are for the chip (die). This grid size was determined after performing mesh independence analysis. For the grid independence study, the mesh resolution of the model was continuously refined until there was less than 1% difference in the computed temperatures. This analysis indicated that the grid size of 75,919 elements is sufficient. Total chip power is $Q = 3\sin 2\pi t + 3$ (W), which is applied for 1 s. The temperature rise in the simulation domain is represented by ΔT (K) throughout the paper. Figure 3-3 (a) shows the spatial distribution of the temperature rise

in the FCBGA package extracted from the FE model after 1 s. The temperature rise of the chip is plotted separately in Figure 3-3 (b).

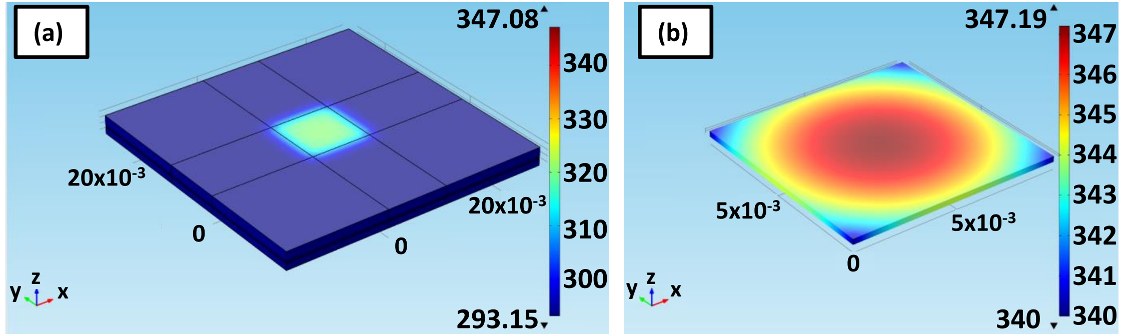

Figure 3-3 Spatial distribution of temperature rise extracted from FE method after 1 s for (a) FCBGA package and (b) Chip.

After obtaining the transient temperature field at the package level, the POD model is developed using the algorithm demonstrated in Chapter 2, section 2.3. 26 observations of the transient temperature solution were taken in the first 0.5 s using the package level FE model. These observations correspond to the temperature solutions obtained at different time instants using total chip power of $Q = 3\sin 2\pi t + 3$ (W). It is important to note that the observations are generated only for this case and results for any different power dissipation are calculated without any new observations. In fact, the POD solutions of these transient thermal scenarios are independent of the initial observations. Essentially, for any linear system, once the solution to a sample case of chip total power is obtained, there is no need to generate new observations or full field FE simulations. The ability of the POD method to predict other cases based on a smaller sample set can significantly decrease computational cost. After the observations have been generated, the POD basis functions (POD modes) are calculated. In order to build a reliable but fast reduced order model, only four POD modes are used in the present model. This was chosen such that the *cumulative correlation energy*, E_m, Eq. (2.2), was greater than

99.9%. The first two modes alone capture over 96% of the energy. The results will not have the desired accuracy if the number of initial observations, *n*, is less than the minimum required POD modes (~ 4 in the present case).

Since the POD modes are three-dimensional, for better visualization, two-dimensional contours of the first four POD modes at height z=1.33 mm, across the center of the die, are illustrated in Figure 3-4. This height was chosen because it has the highest temperature gradient, due to material inhomogeneity and the application of power source only to the die. The POD modes are normalized with the total sum of the modes for a more accurate comparison.

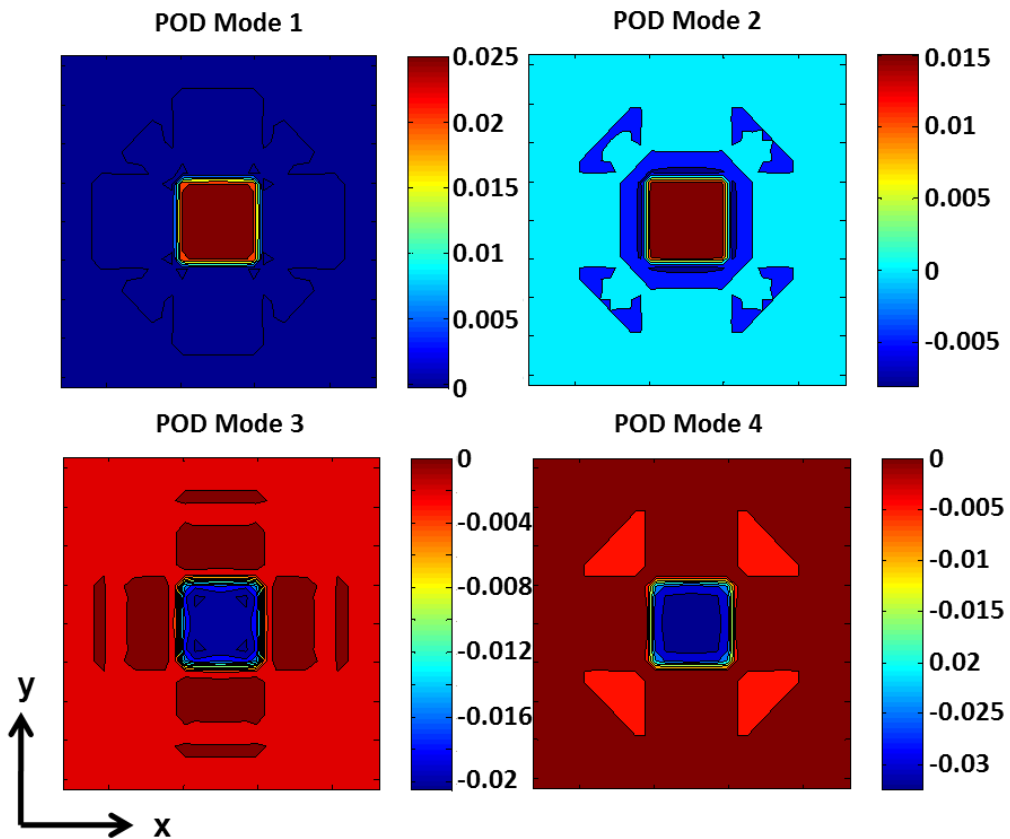

Figure 3-4 Two dimensional contour plots of the first 4 POD modes at z = 1.331 mm from the bottom of the package; this plane crosses the center of die.

Table 3-1 Material properties and dimensions of the package

	Thermal Conductivity (W/mK)	Density (Kg/m³)	Specific Heat Capacity (J/KgK)	Dimension (mm³)
Die	98.4	2300	721	10x10x0.266
Underfill	0.6	1820	236	10x10x0.1
Solder Bumps	50	8510	183	40 % of the die surface area
Effective underfill	1.47 (horizontal) 20.24 (vertical)	4496	214.8	60 % of the die surface area
Mold	0.5	1820	236	37x37x1.9
Substrate	0.7	1700	920	37x37x1.1

To have a realistic and accurate thermal simulation, a detailed dynamic power map of the embedded chip is required. However, one of the major challenges in microelectronics is the determination of the dynamic power dissipation in the chip, since power values and temperature distribution are coupled in an electro-thermal loop. A randomly generated function was assumed for the dynamic chip power in this study to illustrate the application of the POD formulation. Figure 3-7 (d) shows the randomly generated power distribution for the chip for the first 1 s. The minimum and maximum allowed values for the power were chosen to be 3 W and 18 W, respectively. In essence, there are three changes in the nature of the previously used power source ($Q= 3\sin 2\pi t + 3$ (W)) and the current random chip power:

- The first case is only applied for 0.5 second, whereas the second case used for POD approach models the entire 1s.
- The magnitude of the maximum value for the second case is 18 W vs. 6 W for the initial FE simulation.

- The temporal behavior of the power has changed from a well-defined sinusoidal function to a randomly generated step function.

The benefit of using the POD model to predict the transient thermal profile for a different power source than the original one is that no new observation or full field simulation is required. The POD coefficients were calculated as functions of time using the method of Galerkin Projection [37].

Once the POD modes and the *b*-coefficients are calculated, the transient temperature field can be determined using Eq. (3.1). Figure 3-5 (a) displays the three dimensional spatial distribution of temperature extracted from the POD model at 1 s. For higher precision, the domain is sliced vertically along the XZ plane and four of these slices are presented. The right-most slice is the A-A cross section across the center of the die (Figure 3-5 (b)). To validate the results of the POD model, a full field FE model with a time step of 0.05 s was developed in COMSOL using the same grid points and elements used in the POD model. The results are shown in Figure 3-5 (c). It can be inferred that the POD model closely predicts the transient thermal behavior of the system not only for the given time domain but also for projected future time (> 0.5 s) using just a few POD modes. The mean error between the POD and FE model is 7.2 % over the entire space and time domain. Required computation time for the full field FE simulation is 23.7 min vs. 40 s for the POD simulations. The first POD simulation run-time is 40 s, while additional simulations with different power sources take 15 s each. The computations are performed on a workstation using an Intel(R) Core (TM) i7 @ 2.20 GHz with 8 GB RAM.

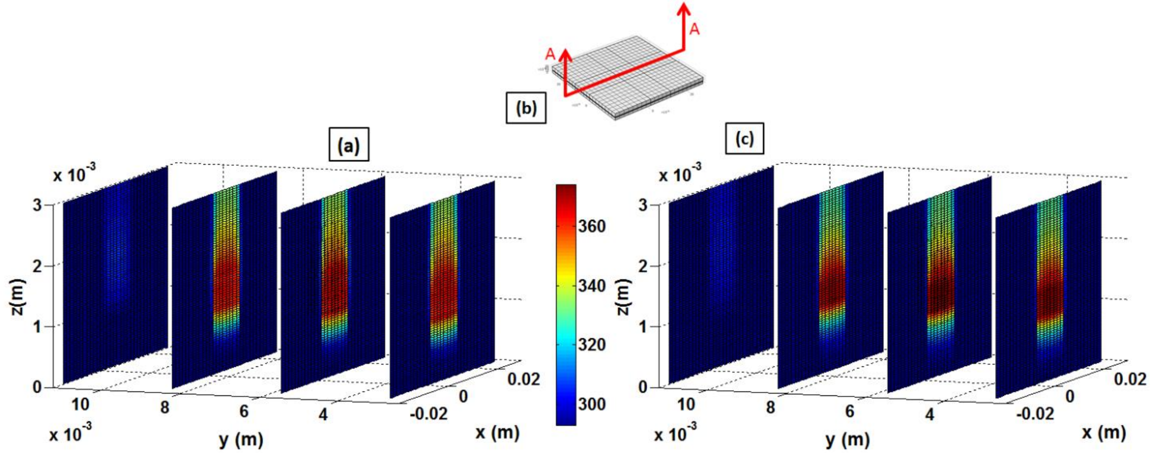

Figure 3-5 Spatial distribution of temperature rise at 1 s extracted from the POD model (a) and FE simulation (c). The domain is sliced vertically along XZ plane. The right-most slice is the A-A cross section (b).

For a more comprehensive comparison between POD and FE results, the time-dependent temperature rise at four different points in the FCBGA package (center of the mold, die, underfill, and substrate) was considered (Figure 3-6 (a)). The maximum error occurs at the center of the die between t=0.833 and t=1 s. As illustrated in Figure 3-6 (b), this is the time period when the maximum jump in the total chip power occurs. The dotted arrow in Figure 3-6 points to the time of this maximum jump in the temperature plot. This error can be attributed to the fact that only four POD modes are kept for this simulation and the rest of them are neglected. The error between POD model and FE results will be reduced by considering a larger number of modes because of the fact that they are excited during an abrupt change in the input power profile and partially compensate for this difference. It is to be noted that the error generated at the package level will propagate through the simulation to the chip level analysis. In fact, the error generated at each level will be passed on to the following level of modeling. In some cases the accumulated error can potentially cause a large variation of the POD solution

from FE results. One approach to control and limit the growth of inaccuracy is to use the hybrid scheme iteratively by feeding the solution from one level to the previous level until the solution converges with the desired accuracy.

After obtaining the transient thermal solution at the package level and with the POD model, the next step in the Hybrid Scheme is to transfer the solution to the chip with the higher spatial resolution in the form of boundary conditions. Due to the transient nature of this analysis, a combination of temperature and heat flux is extracted on the top and bottom surface of the die, respectively, between 0 and 1s at 10 different time intervals (every 0.1 s) and applied as temporal boundary conditions at the chip level model. The four side walls of the die are assumed to be adiabatic after considering the high aspect ratio of the die. The solution is interpolated on a much higher spatial resolution at this level (268,033 elements to model the chip at this level vs. 343 elements to model the chip at the package level).

At the chip level simulation, the die is no longer treated as a solid block. It is segmented into 10 subdomains called function blocks. In practical applications, each block represents a specific component with unique functionality on the chip. In this study, the blocks were artificially created for illustration of the proposed methodology [106]. As demonstrated in Figure 3-2(b), each block has three layers: 1. Top Si layer with the thickness of 0.249 mm, 2. Middle layer which is a 5 µm-thick device layer, and 3. Interconnect/Dielectric multilayer at the bottom with the thickness of 16.72 µm. The third layer consists of 21 sub-layers including 10 metal layers.

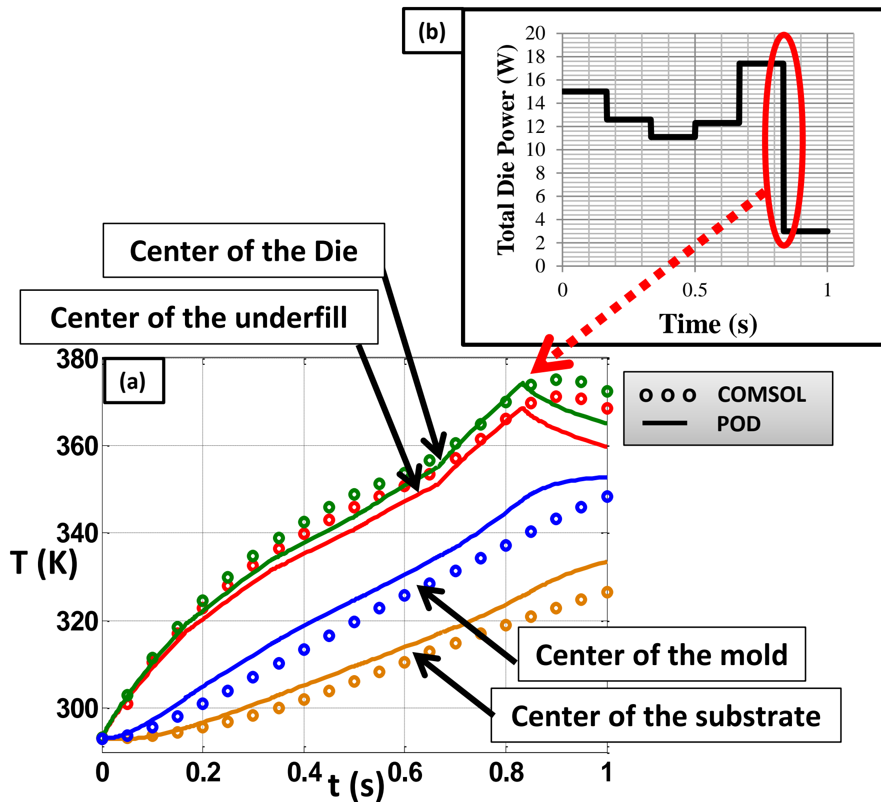

Figure 3-6 Comparison of temporal dependence of temperature rise between FE (markers) and POD (solid lines) models at 4 different points (a) and the corresponding randomly generated total chip power (b).

Due to the high level of geometrical complexity, a combination of directional volume and surface -averaging methods were used to determine the effective properties of the functional blocks. Table 3-2 indicates the calculated material properties of the blocks at the chip level simulations. Density and specific heat are calculated using the volume averaging method. In-plane thermal conductivity is determined based on the ratio of the volume of the interconnects to the total volume, due to the fact that the in-plane thermal transport is governed mainly by the interconnects. On the other hand, vias are the dominant paths of through-plane heat transfer in each block. Therefore, for the vertical thermal conductivity, the values are calculated based on the ratio of the volume of the

vias to the entire volume of each block. At this stage, the spatial resolution is limited to the sub-layers of the blocks.

Table 3-2 Properties and dimensions of the function blocks for chip level simulation.

		Vertical Thermal Conductivity (W/mK)	Horizontal Thermal Conductivity (W/mK)	Density (Kg/m^3)	Specific heat capacity (J/KgK)
Interconnect/ Dielectric Layer	Block 1	0.480	3.531	1512.17	742.01
	Block 2	0.481	3.530	1512.16	742.01
	Block 3	0.487	3.558	1512.73	741.99
	Block 4	0.476	3.494	1511.44	742.05
	Block 5	0.494	3.655	1514.61	741.90
	Block 6	0.466	3.417	1509.90	742.12
	Block 7	0.487	3.584	1513.21	741.96
	Block 8	0.492	3.634	1514.20	741.91
	Block 9	0.481	3.493	1511.46	742.05
	Block 10	0.485	3.564	1512.82	741.98
Device Layer		34	34	2320	678
Si Layer		130	130	2329	700

Once the chip is divided into subdomains, the dynamic power grid needs to be assigned to individual function blocks. For this study, the Joule heating produced in the third layer is neglected and the only powered layer is the device layer. Using the same method as described earlier, 10 random power sources with minimum and maximum values of 0 and 3 W were generated between 0 and 1s. The power sources for blocks 1, 2, and 10 are presented in Figures 3-7 (a-c) as representatives. The block power sources are generated in such way that their sum will equal the total chip power used for the package level simulation as shown in Figure 3-7 (d).

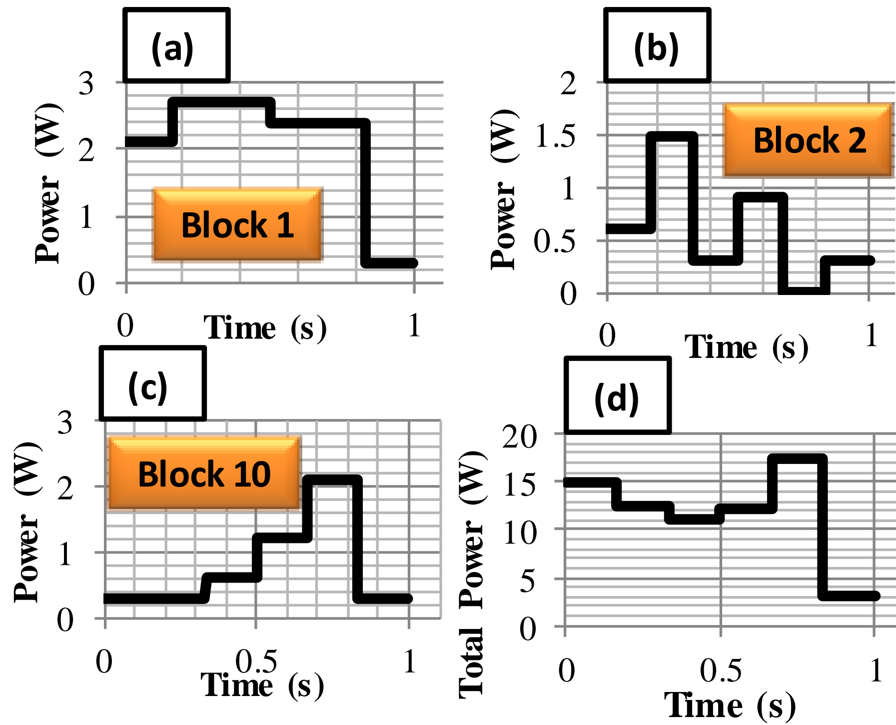

Figure 3-7 Dynamic randomly generated power profile for function blocks 1, 2, 3 (a-c) and randomly generated total chip power (d).

After allocating the power sources to the function blocks, an FE model is developed using the time step of dt=0.05 s for the final step of the hybrid scheme. As mentioned, the model consists of 268,033 elements. The computational time to run the transient simulation for 1 s is 26.27 min. Figure 3-8 displays the 2D spatial distribution of temperature rise extracted from the FE solution at various times between 0 and 1s at height z=16.72 μm, which is the plane between the device layer and Interconnect/Dielectric multilayer (~ plane between layer 2 and 3). This plane is chosen for the demonstration since it has the highest through-plane temperature gradient.

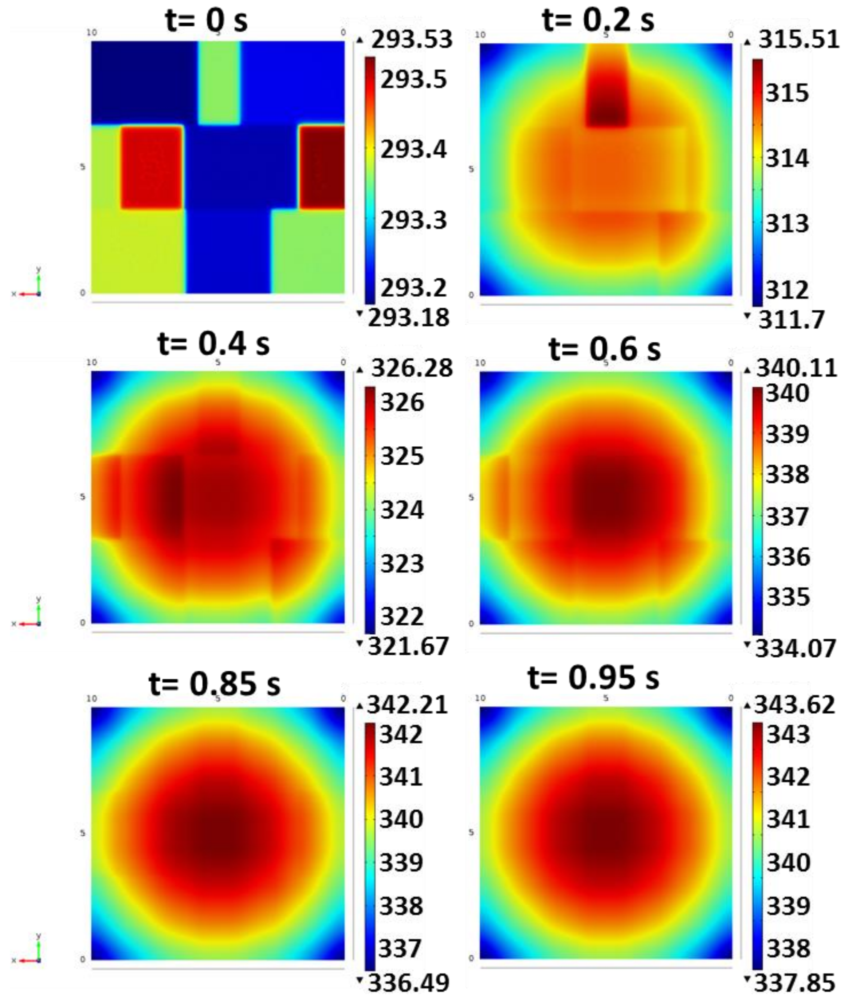

Figure 3-8 Transient temperature distribution at the interface of Device and Interconnect/Dielectric layers.

3.5 Summary and Conclusion

In this chapter, a computationally efficient and accurate multi-scale reduced order transient thermal model was developed which has the capability of modeling several decades of length scale from "package" to potentially "chip components" at a considerably lower computational cost, while maintaining satisfactory accuracy. Another distinct benefit of the proposed method is that, for any linear system, the POD solution is independent of the transient power profile. In other words, once the solution to a sample power input is obtained, there is no need to generate new observations or full field FE

simulations. This important feature can drastically decrease computational cost for parametric numerical simulations, making POD a fast and robust method for reduced order model of transient heat conduction in microelectronic devices. It was also shown that the POD model accurately predicts the transient thermal behavior of the system for not only the time domain considered for the initial observations but also for time outside the specified initial domain. The mean error between the POD and FE model was 7.2 % over the entire space and time domain.

The hybrid scheme proposed in this study is not limited to the two levels considered in the present study. One of the strengths of this method is that the algorithm can be scaled to multiple levels and can be used to simulate more detailed structures on the chip while taking advantage of the capabilities of POD method to avoid any further full field simulation. In essence, without losing the desired resolution, the hybrid scheme proposes a new approach to further decrease the computational cost by orders of magnitude.

CHAPTER 4: RAPID MULTI-SCALE TRANSIENT THERMAL MODELING OF PACKAGED MICROPROCESSORS USING HYBRID SCHEME

The preceding chapter described the development of a computationally efficient and accurate multi-scale thermal methodology for chip and package, using a hybrid scheme comprising of two different multi-scale approaches: *Progressive Zoom-in* and *Proper Orthogonal Decomposition (*POD*)*. A Flip Chip Ball Grid Array (FCBGA) package was considered for low power chips where heat sink and forced cooling are not employed due to the compact form factor and lower power. Dynamic power distributions were applied for the total chip power, as well as for the function blocks on the chip. To validate this methodology, the results were compared with a finite element (FE) model developed in COMSOL®. It was shown that the suggested method has the ability of modeling several decades of length scale from package to chip components at a considerably lower computational cost than the methods available in the literature.

4.1 Objective

In this chapter, the previously proposed hybrid scheme method was further enhanced to account for the transient thermal analysis of a packaged high power microprocessor where, in fact, force convection plays a major role in the thermal transport of the structure. In addition to the case investigated in the Chapter 3, a realistic highly spatially resolved power map for the microprocessor was considered. Moreover, two transients were investigated:

Scenario A: The initial application of the power map to the chip as a step function.

Scenario B: A sudden impulsive/oscillatory change to the existing power map of the chip.

These two types of transient were chosen to resemble realistic power variations of microprocessors in function. It is important to recognize that the thermal time constants associated with these two thermal scenarios can differ by orders of magnitude. Depending on the characteristics of the package, reaching the steady state condition in Scenario A can take up to several hundreds of seconds whereas the quasi-steady condition in Scenario B can be reached within seconds.

4.2 The Hybrid Scheme

The hybrid scheme combines the implementation of POD technique and progressive zoom-in approach, as explained in Chapter 3. Figure 4-1 shows a flowchart of the hybrid scheme used in this study for multi-scale transient thermal modeling of a representative Flip Chip Ball Grid Array (FCBGA) microprocessor package. Key steps are labeled to make it easier to follow the procedure throughout the chapter. The overall approach is outlined below [107]:

- **Obtaining transient thermal solution for an applied volume averaged heat generation at package level (a1)**

The first step is to model the entire structure, i.e., the package, including the heat sink, die, underfill, solder bumps, and substrate. The volume averaged power is then applied uniformly as a step function, and the transient thermal solution is obtained for 10

seconds **(thermal scenario A).** At this level of simulation, the chip is modeled as a solid block with effective material and thermal properties without considering internal details.

- **Attaining steady state solution (a2)**

After determining the transient temperature distribution at package level for a pulsed power input **(thermal scenario A)**, the simulation was carried out until steady state.

- **Determining the transient thermal solution with a sudden oscillatory power input (a3)**

Once the steady state condition has reached, the second transient scenario was explored. A sudden impulsive power input was applied to the existing power of the chip and the transient thermal solution was computed for the first 1.5 seconds (**thermal scenario B**).

- **Predicting the POD-based transient temperature field with an impulsive power disturbance (a4)**

The transient temperature distribution acquired from the previous step is used to build a POD model for the package level simulation. The procedure to build a POD-based model was explained in Chapter 2. The POD model provides the ability to predict dynamic temperature distribution for various spatial and temporal power maps without the need to develop any additional full field models (shown in Chapter 2 and 3 for other cases). In order to demonstrate this capability of POD in this study, a POD-based thermal model was computed for an abrupt impulsive power **(thermal scenario B)** application and the results were compared with a full field FE based model developed in COMSOL. This provides the capability of predicting other heating scenarios which can significantly

decrease the computational cost and can be potentially used to define a criterion for the optimal distribution of the current density in the domain.

- **Zooming into Chip Level as transient boundary conditions (Temperature and heat flux) (a5)**

Once the temperature distribution at the package level was predicted from POD, a combination of temperature and heat flux at the top and bottom surfaces of the chip was extracted at every second from the package level solution. These transient boundary conditions were then applied for the chip level simulation on a much higher resolution.

- **Chip Level transient thermal simulation with the spatially resolved power map (a6)**

At this level the chip is no longer treated as a solid block, but as a three layer stack: 1. Top Si layer, 2. Middle device layer, and 3. Interconnect/Dielectric multilayer. The effective material and thermal properties of these layers were calculated. A spatially detailed power map is assigned to the device layer in a pulsed form.

- **Obtaining full field transient solution (a7)**

The spatially resolved transient temperature profile is determined inside the chip for 10 seconds. This zoom-in method can be continued to additional levels, such that the desired spatial resolution on the chip is achieved. In this chapter, we restrict the simulations to two steps, package and chip level.

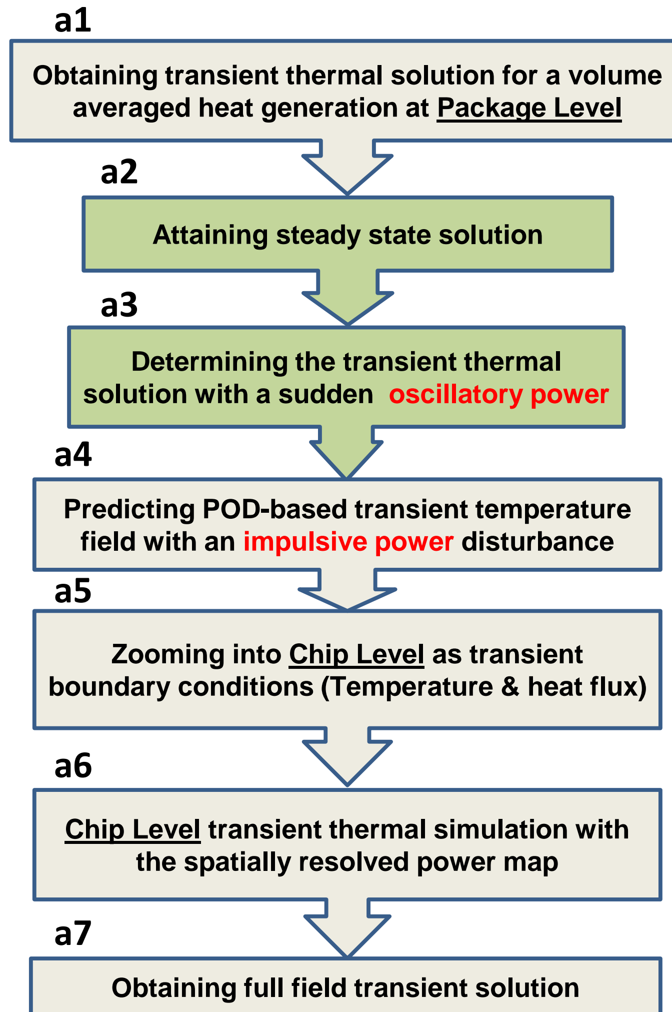

Figure 4-1 Flowchart description of the methodology used in this study for multi-scale transient thermal modeling of a representative Flip Chip Ball Grid Array (FCBGA) microprocessor package.

4.3 Package and Chip Models

The chip modeled here is the Intel Core 2 Duo processor code-named Penryn [108], shown in Figure 4-2 (a) including its major blocks. Penryn has two cores that are mirror images of each other right below the L2 cache. The die size with the 6MB L2 cache is 107 mm^2 with power envelope of 5–150 W. In the present computations, the die is modified to be a 100 mm^2 square. Penryn is derived from the 65 nm based Intel® Core™ micro-architecture (Merom) [109]. A simplified Flip Chip Ball Grid Array (FCBGA) package is considered in this paper, as shown in Figure 4-3 (a). The majority

of the package specifications, including the heat sink, are based upon the thermal design guide released by Intel [110]. For the remaining properties, reference [111] was used as a guideline. Table 4-1 lists the dimensions as well as the properties of various materials. The thermal conductivity of the heat sink is chosen to be the minimum value for copper [110]. For the underfill layer, the effective density and specific heat are calculated based on volume averaging; assuming 40 % of the surface area between the die and substrate is covered with 40 % solder bumps. In the calculation of the underfill thermal conductivity, the effective vertical and horizontal values are formulated based on the dominant paths of heat propagation in the layer [65]. As the solder bumps are vertically orientated, it is expected that the vertical effective thermal conductivity of the effective underfill layer to be significantly higher than the effective horizontal value. It is important to note that for the package level simulation (step a1 in Figure 4-1), the chip is treated as a solid block with internal material homogeneity and power uniformity.

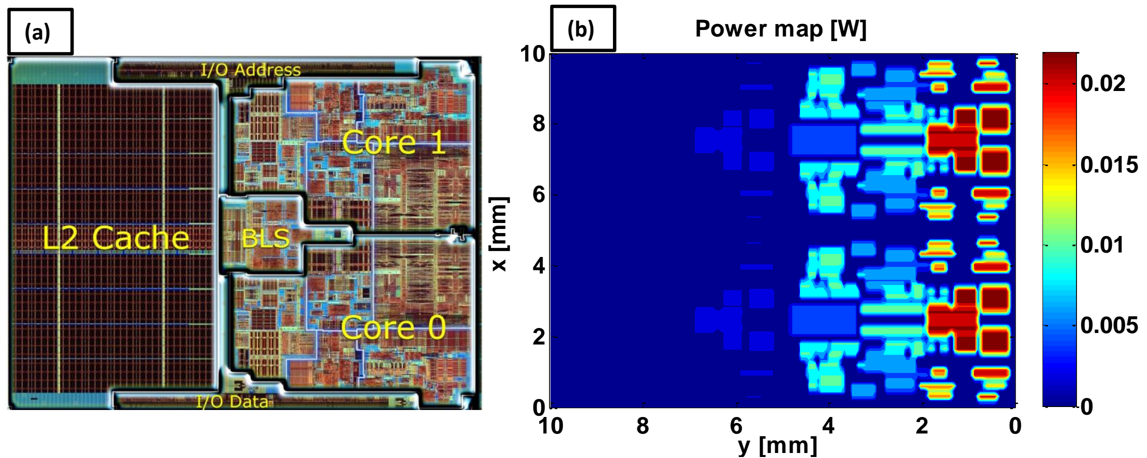

Figure 4-2 (a) Photo of the Intel Core 2 Duo processor code-named Penryn die including major blocks [110], (b) Spatial distribution of power of the die (Power map).

For the package level modeling, forced convection air cooling was assumed on the top surface of the heat sink with heat transfer coefficient of h=30 W/m^{2}K.

Considering that the forced convection cooling from the fins is the dominant path for heat removal, the bottom and surrounding boundaries of the package are assumed to be adiabatic. The initial temperature and surrounding temperature were assumed to be equal to the room temperature T_{amb}= 293.15 K. for simplification of the geometrical model in COMSOL, the fins are removed and an effective heat transfer coefficient, h_{eff}, is calculated and applied on the base area of the heat sink. h_{eff} is formulated based on an equivalent thermal resistant network, consisting of the following:

$$R_1 = \frac{1}{N\eta_f h A_f} \quad , \; R_2 = \frac{1}{h(A_t - NA_f)} \tag{4.1a\&b}$$

where R_1 is the total fin array thermal resistance and R_2 is the surface convection resistance of the base area. N is the total number of the fins; η_f is the fin efficiency; h is the heat transfer coefficient; A_f is the surface area of a single fin, and A_t is the total surface area associated with all the fins and the unfinned portion of the heatsink [104]. The effective heat transfer coefficient can be written based on R_1 and R_2 as:

$$h_{eff} = \left(\frac{1}{N\eta_f h A_f} + \frac{1}{h(A_t - NA_f)}\right)\frac{1}{A_t} \tag{4.2}$$

and is calculated to be 302.71 W/m^2K.

For the chip level modeling (step a5 in Figure 4-1), it is sectioned into three equivalent layers, as demonstrated in Figure 4-3(b): 1. Top Si layer with 0.2493 mm thickness, 2. Middle 5 µm device layer, and 3. Interconnect/Dielectric multilayer at the bottom with 16.72 µm thicknesses. This third layer consists of 21 sub-layers including 10 metal layers, similar to our previous study [65].

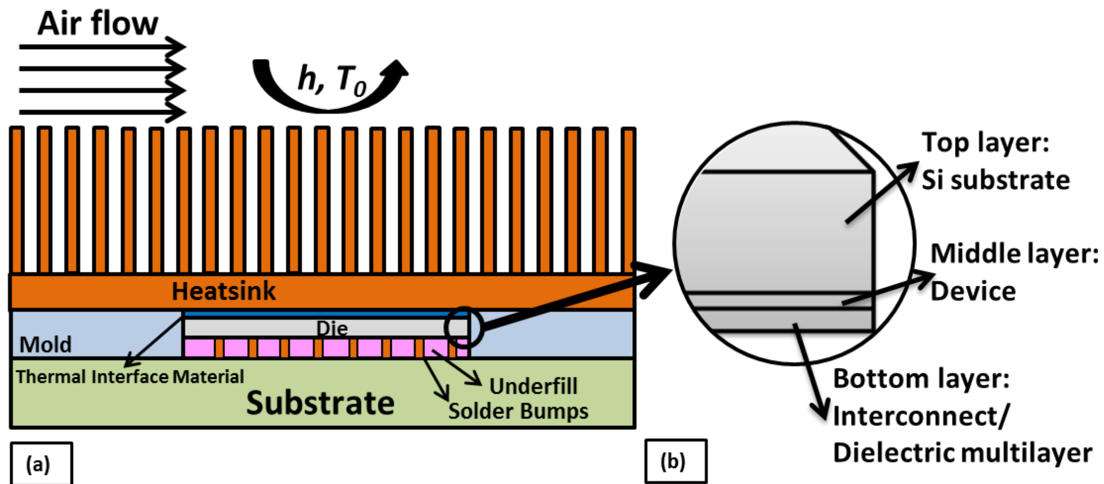

Figure 4-3 (a) Schematic of a simplified FCBGA package used for package level modeling, (b) zoomed-in schematic of the die layer used in chip level modeling.

Table 4-2 lists the calculated material properties of the three layers in the chip level simulations. For Si layer, the values available in literature are used. The corresponding values for the device layer are included from [65]. For the Interconnect/Dielectric multilayer, due to the high level of complexity, a combination of directional volume and surface - averaging methods were implemented to determine the effective properties. Density and specific heat of this layer are calculated using volume averaging. Horizontal thermal conductivity is determined based on the ratio of the volume of the interconnects to the total volume. This is due to the fact that the in-plane thermal transport is governed mainly through the interconnects that are horizontally fabricated. On the other hand, vias are the dominant paths for through-plane heat transfer in the Interconnect/Dielectric multilayer. Therefore, for the vertical thermal conductivity, the value is calculated based on the ratio of the volume of the vias to the entire volume of the multilayer (included in Table 4-2). At this stage, the spatial resolution is limited to each layer of the chip.

Table 4-1 Material properties, thermal properties, and dimensions of the package. * minimum reported value for thermal conductivity of copper released by Intel [103].

	Thermal Conductivity (W/mK)	Density (Kg/m^3)	Specific Heat Capacity (J/KgK)	Dimension (mm)
Die	98.4	2300	721	10x10x266
Underfill	0.6	1820	236	10x10x0.1
Solder Bumps	50	8510	183	40 % of the die surface area
Effective underfill	1.47 (horizontal) 20.24 (vertical)	4496	214.8	60 % of the die surface area
Mold	0.5	1820	236	37x37x1.9
Substrate	0.7	1700	920	37x37x1.1
Heat sink	360*	8960	384	37x37x13.27
Fins	360*	8960	384	Gap=2.1882 Thickness=0.4 Count=17

4.4 Transient Power Map

The highly resolved spatial power distribution of the Penryn processor is obtained from [112]. To produce this power map, a publicly released die photo of the Penryn was examined and the floor plan was generated [113]. The total power of each core and L2 cache is 43.1 W and 4.32 W respectively. The power map is utilized for the modified geometry and simplified FCBGA package used in this paper, such that the total power of the chip is equal to 37.93 W (40% of the original reported power of the Penryn die). 10% of the total power is assigned to the L2 cache. Figure 4-2 (b) shows the spatial distribution of the power in the chip used in this study.

Two types of sudden transient power variations were examined that are categorized as thermal Scenario B:

- **Impulsive**

Impulsive power inputs are rapid changes in voltage or current levels of a chip in either a positive or negative direction for a relatively short period of time. As shown in Figure 4-6 (b), for the package level modeling, an impulsive 20% rise in the existing power of the chip for **3** s was considered (used in step a4 in Figure 4-1).

- **Oscillatory**

An oscillatory power input is an abrupt change in the steady input voltage and/or current of the chip that causes the power to oscillate. An oscillatory disturbance usually decreases to zero in a relatively short period of time. In this study, for the package level modeling, step a3 in Figure 4-1, assumes a suddenly applied oscillatory power input of the sinusoidal form $Q_{Oscillatory} = 37.93 \times (\sin \pi t + 1)$ W for **1.5 s** in the chip power, once steady state is achieved (shown in Figure 4- 6 (c)).

Table 4-2 Material properties, thermal properties, and dimensions of each layer of the die for chip level simulation.

	Vertical Thermal conductivity (W/mK)	Horizontal Thermal conductivity (W/mK)	Density (Kg/m^3)	Specific heat capacity (J/KgK)
Interconnect/ Dielectric Layer	0.48	3.53	1512.17	742.01
Device Layer	34	34	2320	678
Si Layer	130	130	2329	700

4.5 Results and Discussion

A COMSOL model was developed using time step, Δt = 0.1 s. The solution convergence was confirmed with respect to the solver type, time step, and time integration method used for transient analysis. The model of the package consists of

112,227 elements, of which 1,372 are for the die. This grid size was determined after performing mesh independency analysis. As indicated in Section 4, the total spatially uniform chip power used for the package level is 37.93 W, which is applied at the initial time and stays on throughout the length of the simulation (10 s). The spatial distribution of temperature from the FE model after 10 s for the package and the chip is illustrated in Figures 4-4 (a) and (b) respectively.

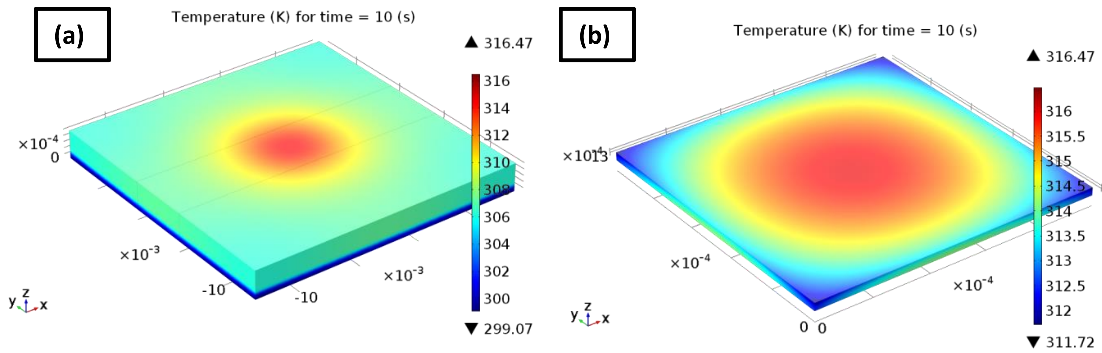

Figure 4-4 Spatial distribution of temperature rise extracted from FE method after 10 s with uniform heating for (a) FCBGA package & (b) chip.

Subsequently, as indicated in Figure 4-1 (step a2) the steady state solution to the package level model with the initial pulsed power input was computed. Figure 4-5 (a) displays the volume-averaged temporal distribution of temperature in the substrate, die, and heat sink until steady state condition is reached. Once the steady state condition was achieved, a sudden oscillatory power input of $Q_{Oscillatory} = 37.93 \times (\sin \pi t + 1)$ W was applied for the first 1.5 s as indicated in step a3 in Figure 4.1 and shown in Figure 4-5 (b). This sinusoidal power is applied for a relatively short period of time, less than 0.01% of the time to the steady state. This allows for the assumption of adiabatic boundary conditions for the die as there is not enough time for the heat to diffuse out through the

package boundaries. The arrow in Figure 4-5 (a) indicates the potential time where power disturbance was applied.

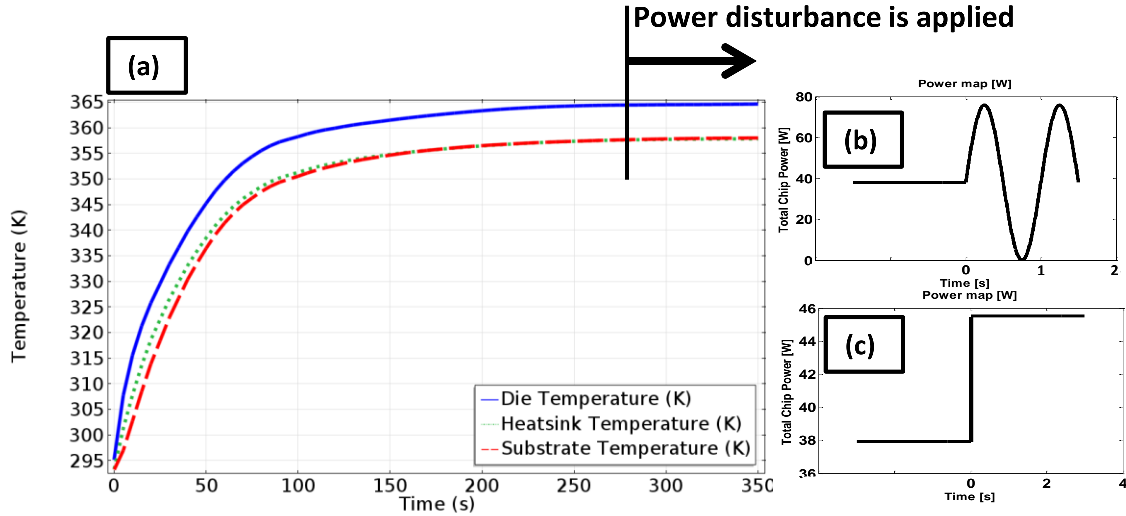

Figure 4-5 (a) Volume-averaged temporal distribution of temperature until steady state condition is reached. The arrow indicates the time where power disturbance is applied. (b) Oscillatory power disturbance applied after steady state is reached. (c) Impulsive power disturbance used to demonstrate the POD prediction capabilities.

The transient temperature field obtained is consequently used to build the package level POD model. Following the algorithm demonstrated in Chapter 2 section 2.3, 30 observations of the transient thermal solution were generated in the first 1.5 s. All of the observations belong to the temperature profiles obtained at different time instants using total chip power of $Q_{Oscillatory} = 37.93 \times (\sin \pi t + 1)$ W. It is important to note that the POD results for the transient thermal behavior for any other scenario with different power inputs are calculated without any new observations. In fact, the POD solutions to different transient thermal scenarios are independent of the initial observations. This ability of the POD method is verified through a comparison of the POD-based and FE model thermal solution for an impulsive power disturbance (step a4 in Figure 4-1).

After generating the observations, the POD basis functions (POD modes) are calculated. To build a reliable reduced order model, the first seven POD modes are kept such that the *cumulative correlation energy*, E_m, was greater than 99.9%. The first three modes, alone, capture over 94% of the energy. POD modes are non-dimensional functions of space. For better visualization, the two-dimensional contour plots of the first three POD modes are expressed in Figure 4-6. Figures 4-6 (a), (b), & (c) correspond to the xy plane crossing the center of the die horizontally at height z = 1.33 mm. Figures 4-6 (d), (e), & (f) correspond to the xz plane crossing the center of the die vertically at y=18.5 mm. The two planes were chosen for their highest temperature gradient that is due to material inhomogeneity and power source being applied only to the die. For comparison purposes, the POD modes are normalized based on the total sum of the modes chosen for this study.

Once the POD modes are defined for a representative case of power input, here an impulsive power, they can be used in developing the transient thermal solution not only for any other types of power input, but also for the projected time. To exhibit this capability of the POD model, an impulsive power input is applied. This power input is a sudden 20% rise in the existing power of the chip and is introduced for 3 s, twice the time of initial observations (shown in Figure 4-5 (c)). Implementing the method of Galerkin Projection, explained in Chapter 2, the POD coefficients were then calculated as functions of time for the new power input.

Calculating the POD modes and the *b*-coefficients the determination of the transient temperature field can be determined. Figure 4-7 (a) exhibits the three dimensional spatial distribution of temperature after 3 s extracted from the POD model.

For a higher precision, the domain is sliced vertically along the xz plane. Four slices are presented that cross through the die area along the A-A cross section shown in Figure 4-7 (b). To validate the results of the POD model, a full field FE model with the time step of 0.03 s was developed in COMSOL using the same grid points, elements, and same power input as that of the POD model. The results are shown in Figure 4-7 (c). It can be seen that the POD model closely predicts the transient thermal behavior of the system not only for the given time domain (1.5 s for the oscillatory power input) but also for the projected time (3> t > 1.5 s of impulsive power input) in future, using just a few POD modes.

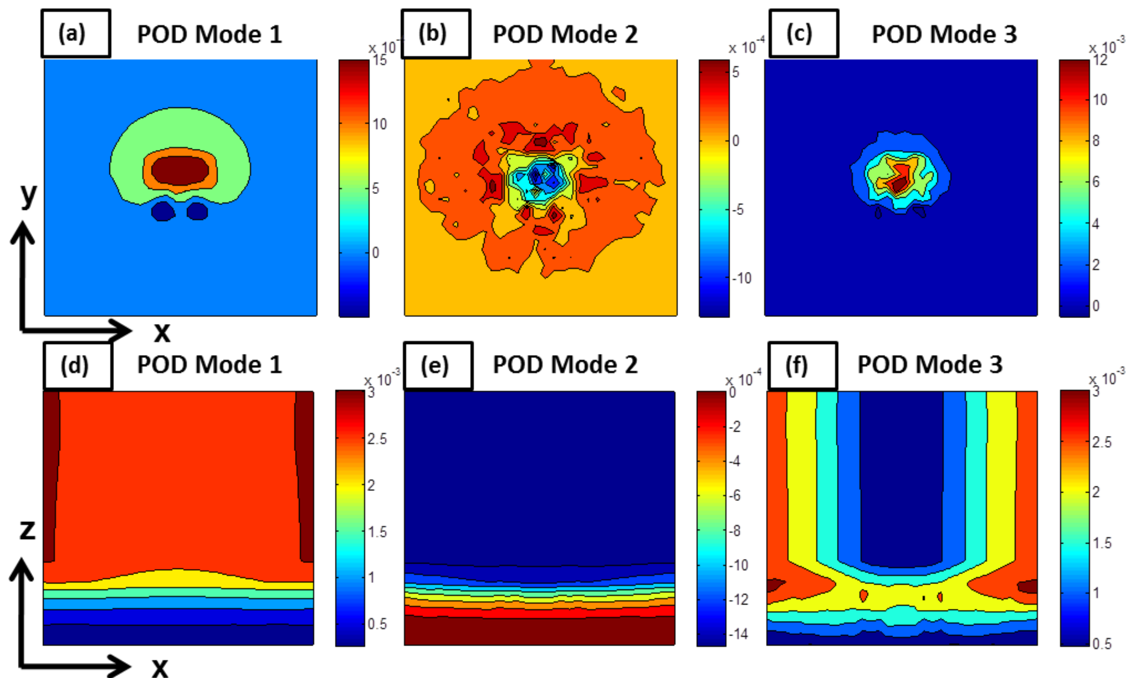

Figure 4-6 Two dimensional contour plots of the first three POD Modes at z = 1.33 mm which is the plane crossing the center of die horizontally (a, b, c) and at y=18.5 mm which is the plane crossing the center of the die vertically (d, e, f).

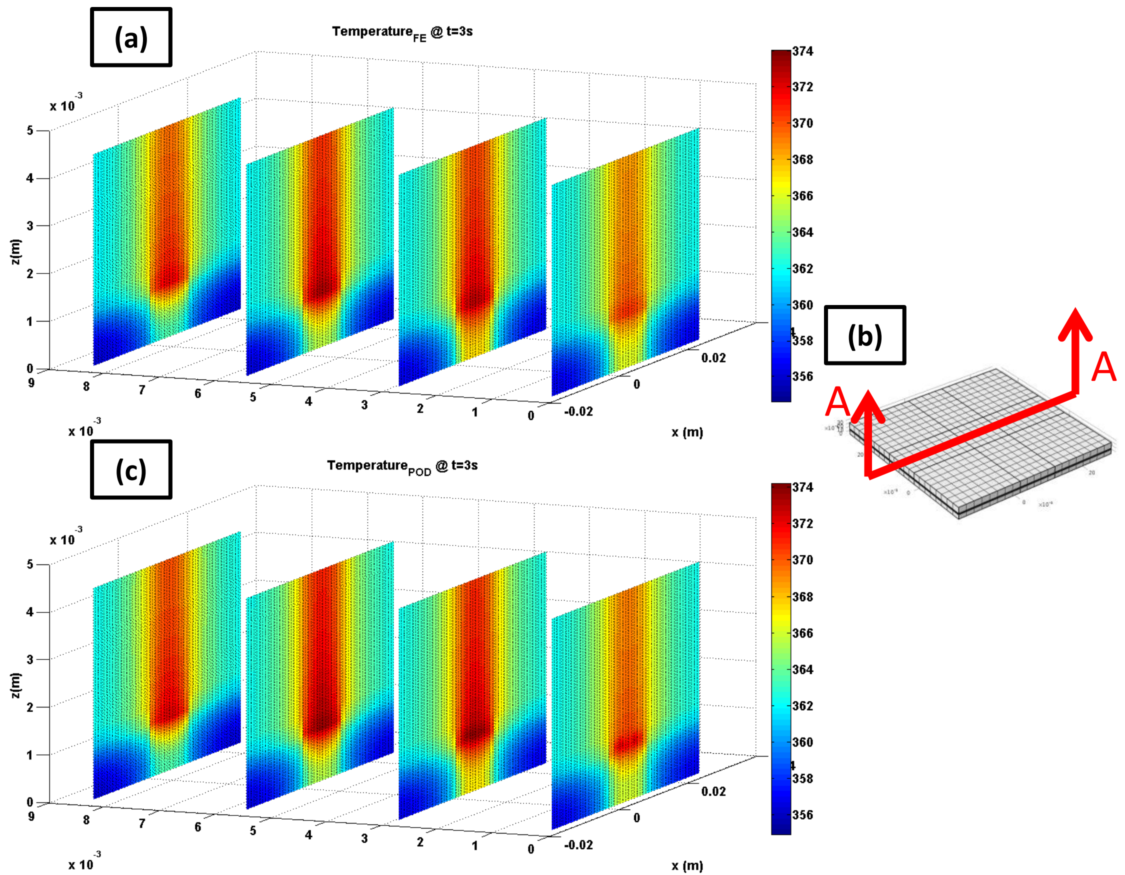

Figure 4-7 Spatial distribution of temperature for the surge-type power after 3 s extracted from the POD model (a) and FE simulation (c). The domain is sliced vertically along the A-A cross section (b) four times in the die area.

Figure 4-8 shows the temporal comparison between the POD model and FE results. The time-temperature history at four different points in the FCBGA package was plotted. The points are chosen at the center of the die, heat sink, and substrate as well as the top of the die. The max error between the POD and FE model is 1.2 K over the entire space and time domain. The computation time for the full field FE simulation is 23.7 minutes. The first POD simulation run-time is 73 s; nevertheless, any further simulation with different power source takes only 20 s to run. The computations are performed on a workstation using an Intel(R) Core (TM) i7 @ 2.20 GHz with 8 GB RAM.

After obtaining the POD-based predicted transient thermal solution at the package level, next step (a2 in Figure 4-1) is to transfer the solution to the chip with the higher spatial resolution in the form of transient boundary conditions. Heat flux on the bottom surface of the die and temperature on the top surface of the die are extracted at 10 time intervals between 0 and 3 s, i.e. every 0.3 s. These profiles are, then, respectively interpolated on a much higher spatial resolution, and applied to the top and bottom surfaces of the die for the chip level simulation. The top and bottom surface of the die consist of 58 elements each for the package level simulation vs. 1,688 elements for the chip level simulation. The side boundaries of the die are assumed to be adiabatic considering the high aspect ratio of the die, as well as the negligible in-plane vs. through-plane heat dissipation.

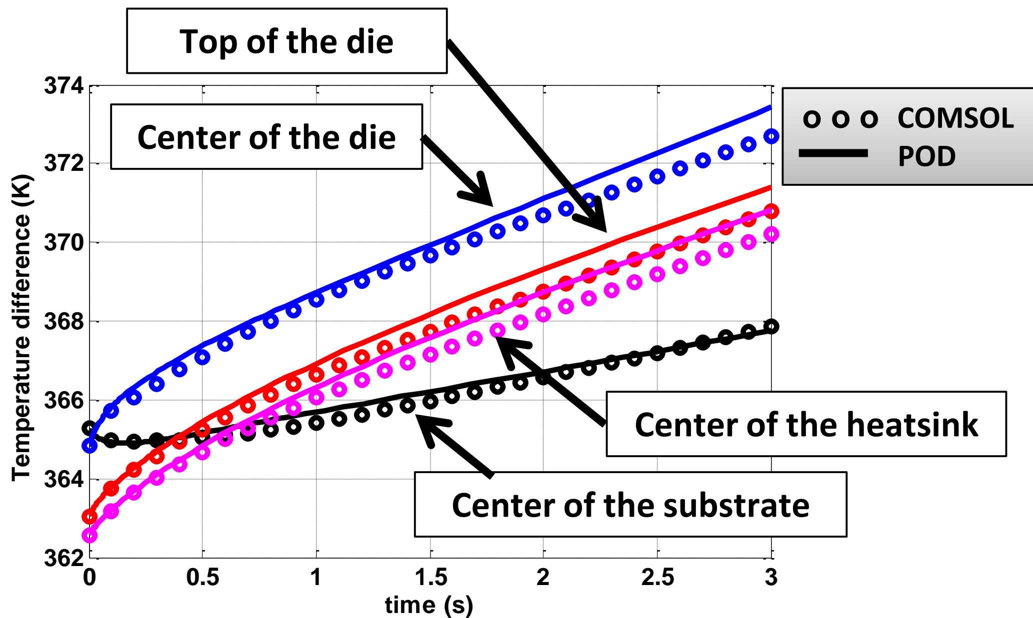

Figure 4-8 Comparison of Temporal dependence of temperature between FE (markers) and POD (solid lines) models four representative points in the domain marked on the plot.

Once the boundary conditions are assigned, the power map shown in Figure 4-2 (b) is allocated to the device layer where the majority of the heat is being generated. For this study the Joule heating produced in the Interconnect/Dielectric multilayer is

neglected and only powered layer is the device layer. The schematic of the layers in the die are shown in Figure 4-3 (b). An FE model is then developed in COMSOL (step a3 in Figure 4-1). The chip level model consists of 199,515 elements. The computational time is 45.5 minutes to run the transient simulation for 3 s. Figure 4-9 exhibits the 2-D spatial distribution of temperature rise extracted from FE solution at three time intervals, at the interface of Device and Si Layer (a and b) and at the bottom surface of the die, i.e. bottom of the Interconnect/Dielectric Layer (c and d). Figures 4-9 (a) and (c) are plotted after 1.5 s and Figures 4-9 (b) and (d) are for the thermal solution after 3 s. Acquiring the spatially resolved temperature profile for chip level model and combining it with the thermal solution at the package level, the full field transient solution is obtained (step a7 in Figure 4-1).

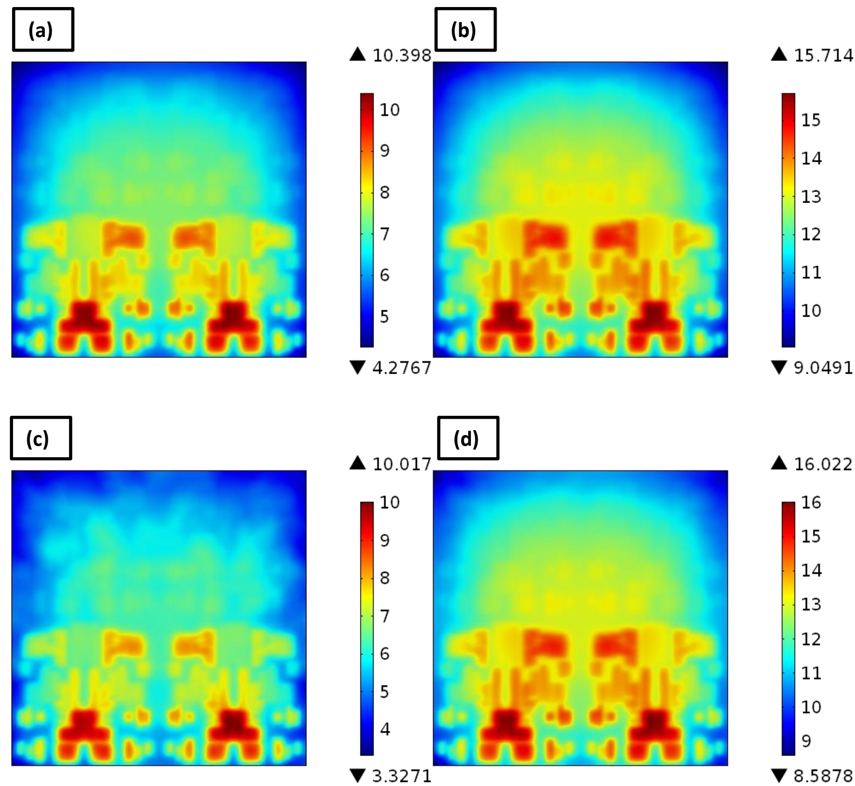

Figure 4-9 Transient temperature rise (K) at the interface of Device and Si Layer (a and b) and at the bottom surface of the die , i.e. bottom of the Interconnect/Dielectric Layer (c and) at t=1.5 s (a & c) and t=3 s (b & d).

This approach can be further extended to multiple levels until the desired spatial resolution on the chip is achieved. As a representation, the results for two steps (package and chip level) are demonstrated in this study.

4.6 Summary and Conclusion

In this chapter the rapid transient thermal analysis of a packaged, forced convection cooled high power microprocessor was studied. A realistic spatially detailed power map for Intel Core 2 Duo processor code-named Penryn was considered. In addition, to resemble realistic power variations of microprocessors in function, two types of transients were chosen. The thermal time constants associated with these two thermal scenarios differ by two orders of magnitude. We utilized our previously developed multi-scale hybrid approach which encompasses two components: *Progressive Zoom-in,* and *Proper Orthogonal Decomposition (POD).* We demonstrated the capability of this approach in modeling several decades of length scale from "package" to "chip" at a considerably lower computational cost, while maintaining satisfactory accuracy. The advantages of using the POD technique was demonstrated in the fast prediction of the transient thermal solution for a sudden impulsive disturbance applied to the chip power, regardless of the power sources used in generating initial observations (oscillatory power disturbance). It was also shown that the POD model closely predicts the transient thermal behavior of the system not only for the time domain considered for the initial observations, but also for the time instants outside this domain. The maximum error between the POD and FE model was 1.2 K over the entire space and time domain. This

characteristic of the POD methodology can be potentially used to define a criterion for the optimal distribution of the current density in the domain.

The hybrid scheme proposed in this study is not restricted to the two levels (package and chip) considered in the present study. One of the strengths of this method is that the algorithm can be scaled to multiples levels of simulation and used to simulate higher spatial resolution for structures comprising the chip while taking advantage of the capabilities of POD method in rapidly predicting different thermal scenarios. In essence, without losing the desired resolution, the hybrid scheme proposes a new methodology to further decrease the computational cost by orders of magnitude.

CHAPTER 5: TRANSIENT THERMAL CHARACTERIZATION OF EMBEDDED NANOSCALE METALLIC FILMS

The continued scaling of transistors and metal interconnects have resulted in high current densities and significant Joule heating in the metal lines, exacerbating thermally driven reliability issues in microprocessors. Therefore, it is imperative to develop an accurate and rapid predictive thermal characterization capability for on-chip interconnect arrays under system operating conditions.

5.1 Objective

The objective of this work is to develop a platform to evaluate rapid transient Joule heating in embedded nanoscale metallic films such as on-chip interconnects. In the current chapter, the effect of rapid transient power input profiles with different amplitudes and frequencies in Cu interconnects were studied using sub-micron resistance thermometry technique. The experimental data are also verified against infrared microscopy and numerical modeling.

5.2 Micro and Nanoscale Thermal Characterization Techniques

Several techniques have been developed and reported in literature for steady state and transient thermal characterization of the modern microelectronic devices. These methods have different application based on their spatial resolutions, thermal time constants, suitability for embedded structures, and ability to be integrated with the device. Table 5-1 provides a summary of the common techniques utilized for high resolution

thermal measurement in microelectronics. Cahill et al. provided a detailed review of thermometry and thermal transport in micro and nanoscale devices [64].

5.2.1 Infrared (IR) Microscopy

One of the most commonly used methods for non-contact thermal imaging is infrared (IR) microscopy. IR microscopy measures the surface temperature based on the fact that any body above absolute temperature emits infrared radiation that is dependent on its temperature. Photons emitted from the surface are focused on the quantum detector of the IR microscope using an optical lens. These photons excite the electrons in the detector creating an electrical signal. The electrical signal is then converted into voltage and is processed to determine temperature. The Schematic diagram of the IR microscopy procedure is provided in Figure 5-1 (b). IR microscopy has the ability to perform steady state and transient thermal measurements. The spatial resolution of IR microscope can go up to 3 μm for steady state and 30 μm for transient measurements. Its relatively low spatial resolution during transient testing is one of the limitations of this technique. The thermal time constant of transient thermal measurements can be as small as 1 μs [44]. Another drawback of this technique is that it requires precise calibration. Also the radiation from the surroundings can cause a major issue in precise thermal readings. Quantum focus Infrascope II was used for thermal imaging and validation of the proposed metrology, shown in Figure 5-1 (a).

5.2.2 Pump-Probe Transient Thermoreflectance (PPTTR) Technique

The (PPTTR) technique was proposed by Paddock and Eesley [62]. PPTTR has the ability to differentiate between the thermal conductivity of thin films and their interface thermal resistance [63, 64]. It is a time resolved methodology that extends the

standard thermo-reflectivity technique [114] to very short time scales by applying optical sampling. Among the advantages of this technique are the fully optical, noncontact, and nondestructive nature of it, together with a high temporal and spatial resolution. This makes PPTTR a prominent methodology in determining the thermal properties of thin dielectric and metal layers.

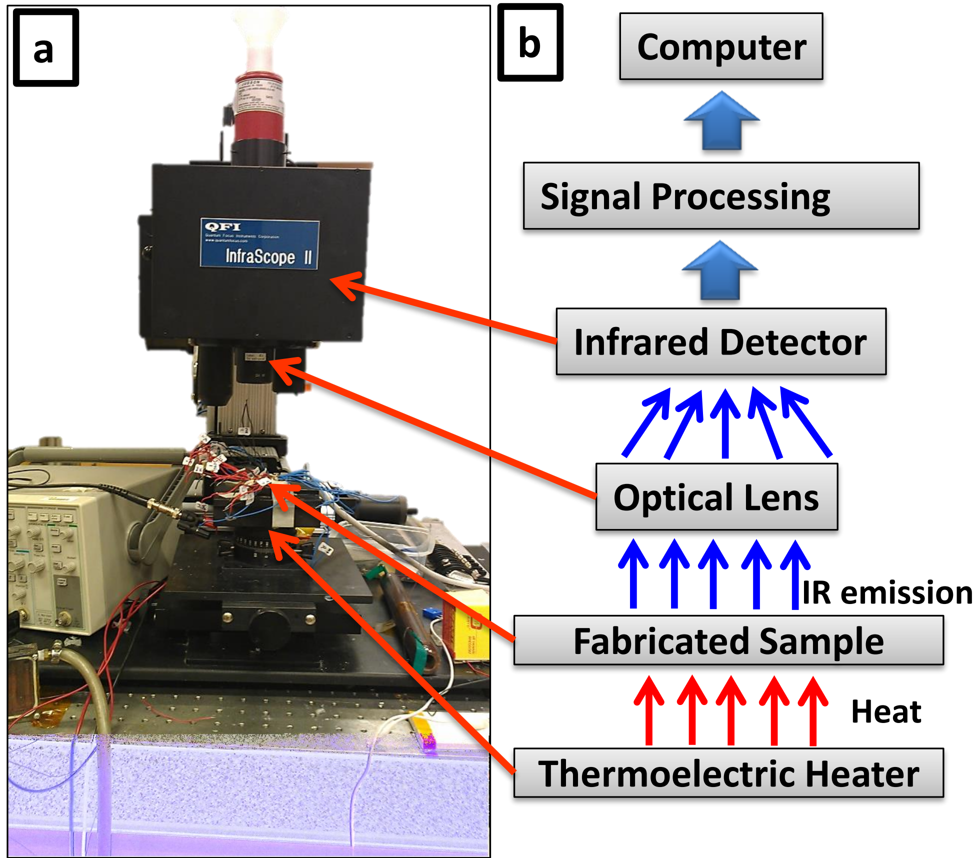

Figure 5-1 (a) Image of the infrared (IR) microscope used for steady state measurements. (b) Schematic diagram of the IR microscopy procedure.

5.2.3 Network Identification by Deconvolution (NID) assisted PPTTR

Another approach to analyze the thermal decay of a PPTTR signal, based on RC network theory of linear passive elements, is Network Identification by Deconvolution (NID). NID was originally introduced by Székely and Bien [115]. This technique, which includes structural information obtained using thermal transient measurements, is used to

analyze the temperature response of semiconductor device packages. Using this technique, one can separate the different contributions to the total thermal resistance and capacitance of the package being studied. Most of the applications of NID have been for the case of step function thermal transient measurement [116, 117]. Recently, Ezzahri and Shakouri considered the case of a delta function excitation applied to the structure [118]. They demonstrated the capability of NID in extracting the thermal conductivity of the thin dielectric layer, as well as the metal layer interface thermal resistance from a single PPTTR signal. One of its strengths is that it does not assume *a priori* (number of layers or interfaces) for the structure of interest.

5.2.4 Scanning Thermal Microscopy (SThM)

Another category for thermal characterization in nanoscale is Atomic Force Microscope (AFM) based approaches. Scanning Thermal Microscopy (SThM), [45] is among these techniques. In this method, a very small thermocouple is fabricated on the tip of an atomic force microscope (AFM) [47-50]. Such a method can provide a maximum resolution of 50 nm. The advantages of SThM are significant spatial resolution, on smooth objects, and high bandwidth. One of the most important drawbacks of this method is that heat transfer between the probe and sample depends on the tip contact with the sample, which can vary with sample hardness, wear, or contact force [51]. Moreover, the interface quality can drastically affect the thermal transport and hence precise calibration on similar surfaces is required.

5.2.5 Scanning Joule Expansion Microscopy (SJEM)

Another AFM based approach is Scanning Joule Expansion Microscopy (SJEM) [46]. SJEM is mainly used to determine the thermal conductivity of thin metallic films.

Therefore, this technique is explained with details in the next chapter, Chapter 6, where the size effect on thermal conductivity of embedded metallic thin films is studied.

5.2.6 Raman Thermometry

One of the common optical thermometry techniques is Raman and micro-Raman thermal imaging method. The energy of the scattered photons from a structure is different than that of the incident photons due to the inelastic scattering and the exchange of energy with lattice vibrations in the material itself. Raman thermometry utilized this phenomenon for the determination of temperature. As the temperature increases, the number of phonons in the excitation mode escalates which will increase the ratio between the anti-Stokes and Stokes peaks. This ratio is can be used to determine temperature [52]. In addition, the shift in the Raman frequency as a function of temperature can be utilized to calculate temperature.

5.3 Proposed Approach: Sub-micron Resistance Thermometry

It is important to note that most of the described methods are used for surface temperature measurement and will not be suitable for buried metallic films within metal/dielectric multi layers such as on-chip interconnects that are not directly accessible. Also, all of aforementioned metrologies are stand-alone measurement tools that require extensive setting up procedures without an ability to be integrated within the structures. Therefore, in order to characterize transient thermal transport in buried metallic film, sub-micron resistance thermometry (RTD) technique was developed. Utilizing this technique, a spatial resolution of 6 μm and thermal time constant of below 1 μs was achieved.

Table 5-1 Summary of micro/nanoscale thermal measurement techniques.

Interconnect Thermometry Techniques	Spatial Resolution	Thermal Time Constant	Suitability for Embedded Structures	Ability to be Integrated with Structure
AFM based approach (SJEM) [119]	10 nm	Steady State Only	N	N
Micro-Raman Thermometry[120]	100 nm	100 µs	Y (Conditional)	N
Scanning Thermal Microscopy (SThm)[64]	30–50 nm	10 µs	N	N
Thermo-reflectance Method[121]	3 µm	100 ps	Limited to Small Depths	N
Infrared Microscopy[44]	30 µm (Transient) 3 µm (Steady State)	1 µs	N	N
Sub-micron RTDs	6 µm	> 1 µs	Y	Y

5.4 Design and Fabrication

5.4.1 Design of the layout

Interconnect architecture have complex geometries and are highly challenging to fabricate. In order to be able to characterize transient thermal transport in burried interconnects, the geometry needs to be simplified while maintaining enough details to chapter the physics of the problem. To achieve this goal, a set of three 70-nm-thick Cu interconnects buried in low-k dielectric materials was designed and fabricated. To measure the transient temperature distribution in the Cu-interconnect arrays, a unique set of 20 embedded sub-micron resistance thermometers (RTDs) were fabricated above the

interconnect layer with a 300 nanometer barrier layer of silicon dioxide. Figures 5-2 (a) and (c) respectively show the top and side view of the proposed layout. A magnified view of the serpentine parts of a typical RTD is shown in Figure 5-2 (b).

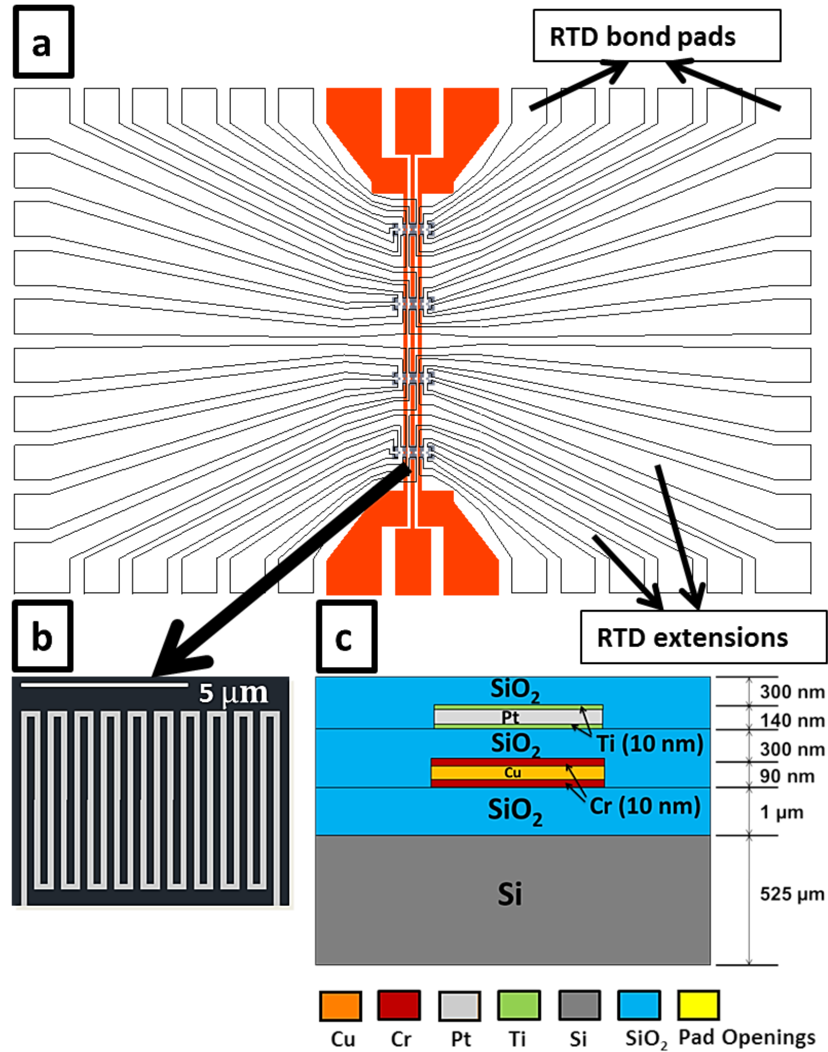

Figure 5-2 (a) Layout of Interconnect and RTD architecture (b) Zoomed-in view of serpentine lines of RTD (c) Cross-sectional view of structure (not to scale).

5.4.2 Fabrication process

5.4.2.1 <u>Interconnect fabrication</u>

The fabrication process was carried in the cleanroom facilities at the Institute for Electronics and Nanotechnology (IEN) at Georgia institute of Technology. The structure

was fabricated on a 4-inch Si wafer of 525 μm thickness. First, a 1.2 μm SiO_2 layer was thermally grown (Lindberg Furnace) on top of the wafer. A layout of three 10 μm-wide, 70-nanometer-thick, and 1.073 μm-long Cu interconnects was then patterned on the sample using standard photolithography process. Negative photoresist (NR9-1500PY) was spun on the sample. The spinner parameters were: Spin speed of 3000 rpm at 500 rpm/s ramp rate for 40 seconds. The sample was soft baked on hotplate at 100 °C for 60 seconds (pre-exposure bake). Afterwards, the photoresist was exposed to UV light using Karl-Suss mask aligner. The sample was then baked at 100 °C on a hotplate 5 minutes (post-exposure bake). The exposed sample was developed for 60 s in diluted RD6 (RD6/DI water 3:1).

Next, three layers of Cr (10 nm), Cu (70 nm), and Cr (10 nm) were consecutively deposited over the patterned negative photoresist layer using E-beam evaporation technique (Denton Explorer). Chromium is used underneath and on top of the Cu film as an adhesion layer between Cu and silicon dioxide. Consequently, the metallic layers were patterned using a lift-off process. Figure 5-3 (a) displays an optical image of the fabricated Cu-Interconnect lines.

A barrier layer is necessary to electrically insulate the interconnects from the RTDs. The breakdown voltage of plasma-enhanced chemical-vapor deposition (PECVD) SiO_2 can be as low as 10^6 V/cm. Therefore, to apply a transient voltage with a peak of 10 V (required in this study), a SiO_2 layer of at least 100 nm is required. To stay well above the breakdown range, a 300 nm SiO_2 layer is deposited over the Cr/Cu/Cr layer using PECVD technique (Unaxis PECVD).

5.4.2.2 RTD fabrication

RTDs fabrication requires more steps compared to that of interconnects due to the different length scales associated with every RTD. The width of the serpentine lines is only 180 nm whereas the RTD extensions which connect the RTDs to the contact pads (shown in Figure 5-2) are 30 to 80 µm wide. Photolithography, however, has the ability to pattern features down to 1-5 µm. Its accuracy will be highly compromised for any smaller features. Therefore, a combination of photolithography and electron beam (Ebeam) lithography technique was utilized. The serpentine section of every RTD was fabricated using Ebeam lithography whereas the RTD extensions and contact pads were fabricated using photolithography. The two layers are overlaid on top of each other with 10 µm tolerance for any misalignment error.

The serpentine sections of all RTDs were patterned on 380 nm of PMMA 6% (an electron beam positive resist) with electron beam lithography (JEOL 9300). The spinner parameters were: Spin speed of 5000 rpm at 2500 rpm/s ramp rate for 60 seconds. The sample was soft baked on hotplate at 180 °C for 90 seconds. Consecutive layers of 10 nm of Ti (an adhesion layer), 120 nm of Pt, and 10 nm of Ti were then deposited using E-beam evaporation (Denton Explorer). The serpentine section of the RTD layer was then formed using a lift-off process. Figure 5-3 (b) displays an SEM image of the serpentine section of RTDs demonstrating the placement of an RTD with respect to the bottom interconnects. The serpentine section of each RTD consists of a 128 µm-long and 180 nm-wide serpentine lines fitted in a 6 µm by 7.7 µm rectangle, shown in Figure 5-3 (c).

The RTD extensions were then patterned on the sample using standard photolithography process. Negative photoresist (NR9-1500PY) was spun on the sample.

The same settings used for the interconnect patterning were applied here. Once the photoresist was patterned, 50 nm of Titanium (Ti) and 350 nm of Gold (Au) were deposited via E-beam evaporation (Denton Explorer). Titanium was used as an adhesion layer between Au and silicon dioxide. Au was chosen for its ability to wire-bond easily. Ultimately, the RTD extensions were developed using a lift-off process.

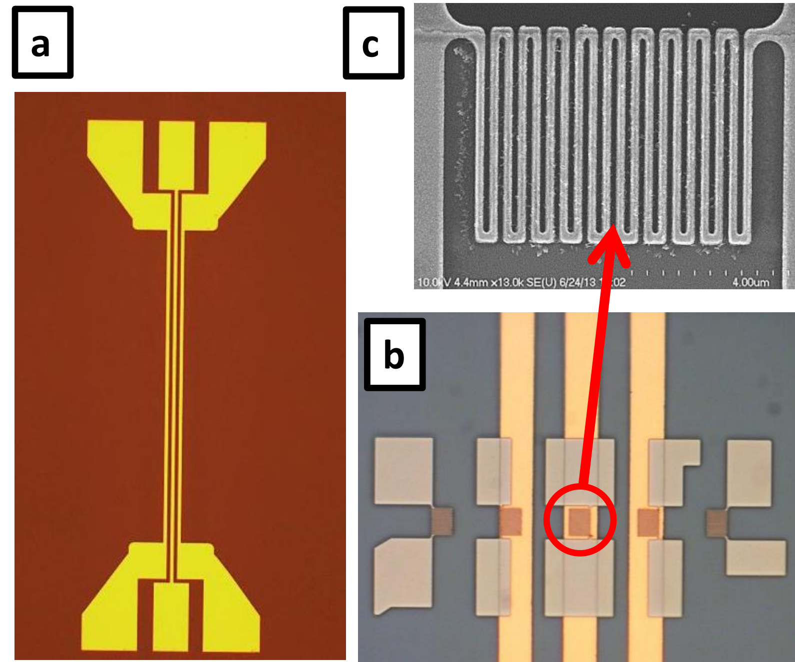

Figure 5-3 (a) Optical Image of Interconnect layer. (b) SEM image of the serpentine section of RTDs fabricated on top of Interconnect layer. (c) Zoomed-in SEM image of the serpentine section of an RTD.

5.4.2.3 Opening the contact pads

In another photolithography step, the areas of the contact pads for the interconnects and RTDs were patterned using negative resist NR9-1500PY. The SiO_2 on top of the openings were then etched away using standard oxide etch by Reactive Ion Etching (Vision RIE) such that the contact pads can be probed with electrical connectivity. The selectivity of PECVD SiO_2 over NR9-1500PY was calculated to be 0.7; i.e., to etch a 300-nm-thick SiO_2 layer, a minimum of 429-nm-thick NR9-1500PY layer is

required. The thickness of the NR9-1500PY is measured to be 1.5 μm that is well above the minimum required thickness. After patterning of silicon dioxide layer, remaining photoresist is removed by acetone and cleaned in ultrasonic bath.

Figure 5-4 (a) shows an SEM image of the final structure. A magnified SEM image of the structure for the interconnects and RTDs is exhibited in Figure 5-4 (b). The serpentine section of an RTD, the RTD extensions, and the overlaying junctions between the two can be seen in Figure 5-4 (b). In order to ensure the repeatability of the experiments, three samples were fabricated and tested.

5.5 Fabrication Challenges

There are some aspects of the fabrication which can highly influence the characteristics of the final device and require further consideration due to their complex nature. Some of these fabrication challenges and their solutions are discussed below.

5.5.1 Pt fabrication process

During the deposition of Pt for the serpentine section of the RTDs, the temperature of the evaporation chamber increases significantly. This is due to the high melting point of Pt, 1768.3 °C with respect to other metal sources for evaporation [122]. As the Pt vapor comes into contact with the surface of the sample, it raises the temperature of the sample considerably. This elevated temperature damages and deforms the patterned Ebeam lithography resist, PMMA 6%, on the sample. If the deformation in the resist is severe, the entire serpentine section will be affected. Figure 5-5 (c) and (d) demonstrate examples of faulty fabricated lines as a result of extreme deformation in the Ebeam resist. Even minor deformations in the resist will result in defects and impurities in the deposited metal layer. Figure 5-5 (a) shows an SEM image of the Pt serpentine lines with defects

and residues between the parallel lines. Figure 5-5 (b) is a zoomed-in SEM image of the serpentine lines exemplifying the defects.

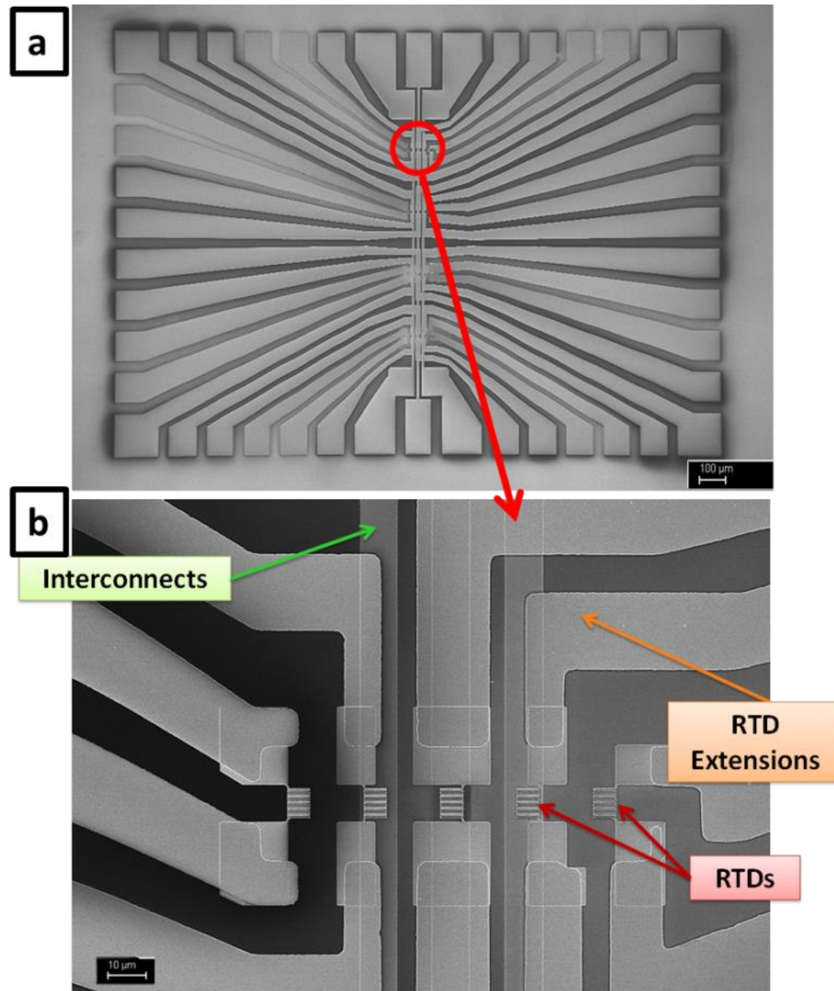

Figure 5-4 (a) SEM image of final device. (b) Zoomed-in SEM image of the structure which shows interconnects fabricated beneath the RTDs, Serpentine section of the RTDs, and the extension of the RTDs fabricated on top.

Any metallic residue between the serpentine lines can cause a short circuit and change the RTD's overall resistance. Therefore, an elemental analysis of the serpentine lines and the surrounding residues was performed via energy dispersive X-ray spectroscopy (EDX). Figure 5-6 (a) displays the SEM image of three adjacent serpentine lines for EDX. Figure 5-6 (b) and (c) show the EDX spectra of the points marked in the

SEM image. It can be seen that the peaks of Pt is detected in both areas. This means that the residues between the parallel lines in the RTD are indeed metallic and form short circuit within the device.

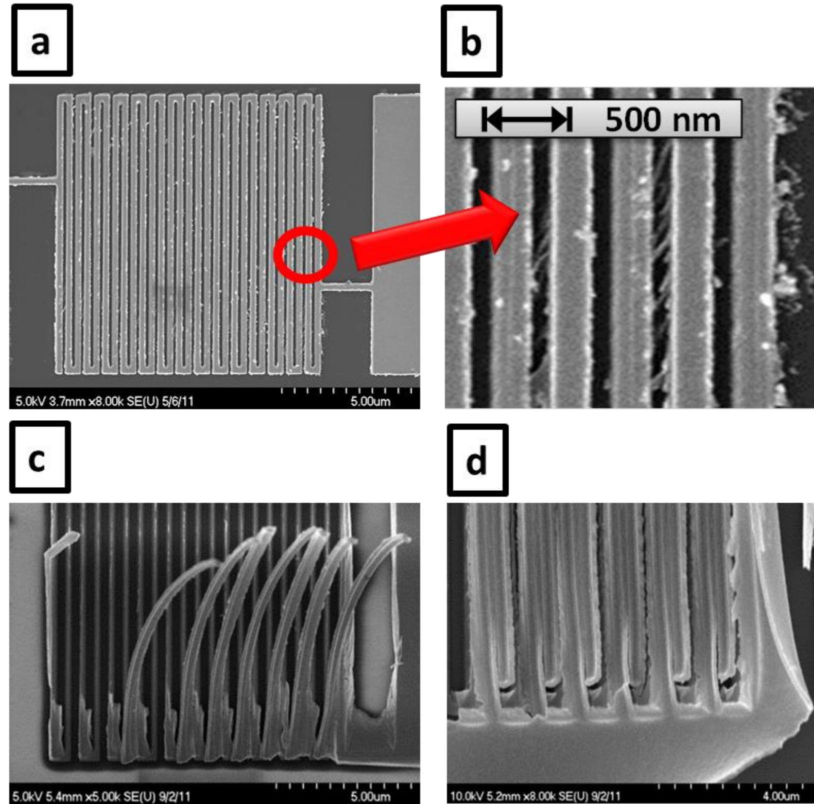

Figure 5-5 (a) SEM image of a representative serpentine section of an RTD with defects. (b) Zoomed-in SEM image of the serpentine lines. (c) and (d) are two examples of faulty fabricated serpentine lines.

To prevent the deformation of the Ebeam resist during Pt evaporation, a conductive Aluminum shadow mask was designed, built, and placed over the sample inside the chamber. Schematic of the setup whithin the chamber is shown in Figure 5-7. The mask has a small opening allowing for Pt vapor to reach the surface of the sample at a much lower quantity. The excess amont of Pt vapor will de deposited on the shadow mask instead. Also, since the mask is highly thermally conductive, the diffused heat from deposited Pt is carried along the mask to the substrate, avoiding the surface of the sample.

Utilizing this approach, the RTDs were successfully fabricated with no defects or impurities.

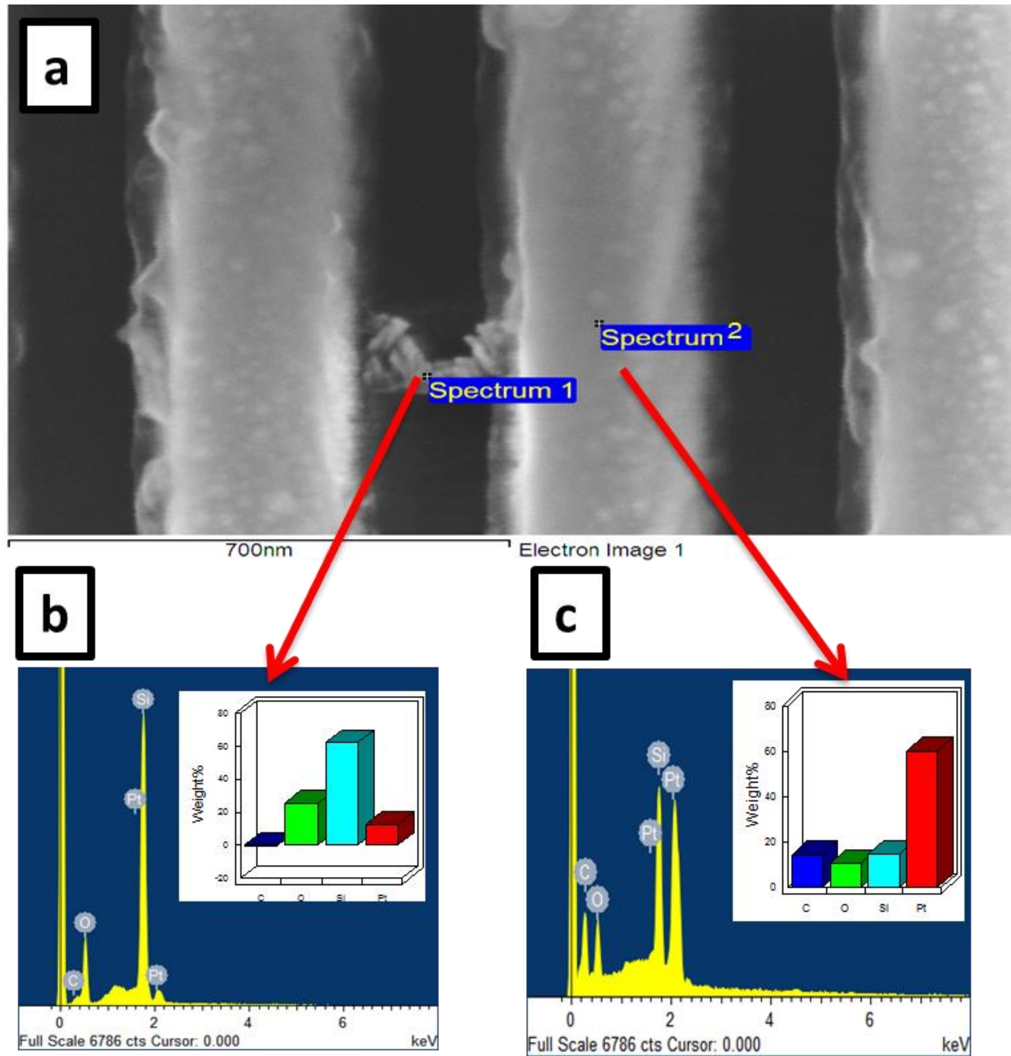

Figure 5-6 (a) SEM image of three adjacent serpentine lines for energy dispersive X-ray spectroscopy (EDX). EDX spectra of the points marked in the SEM image.

5.5.2 Determining the contact resistance

Another important aspect of the RTD fabrication that needs to be further analyzed is the junction area between Pt (serpentine part of RTDs) and Au (connecting extensions of RTDs) of an RTD. Figure 5-8 exhibits an SEM image of a representative junction area between Pt and Au. There is an electrical contact resistance, also known as *ECR*,

associated with the junction of Pt and Au which contributes in the total resistance of each RTD. *ECR* can even be a result of two flat surfaces come into contact [123]. Several studies reported the dependence of contact resistance on variables such as film thickness, surface roughness, and film deposition techniques. [124-126]. However, most of the studies on the theory of contact resistance were performed at macro-scale [127] which does not necessarily predict the behavior of the electrical resistance at micro/nanoscale [128]. Since the junctions of the two overlaid metals in the current design are far from the heated parts of the interconnects, it can be assumed with high approximation that *ECR* does not change with the change in the temperature of the interconnects. Therefore, the *ECR* for each junction needs to be identified and deducted from the total resistance of each RTD used for temperature extraction.

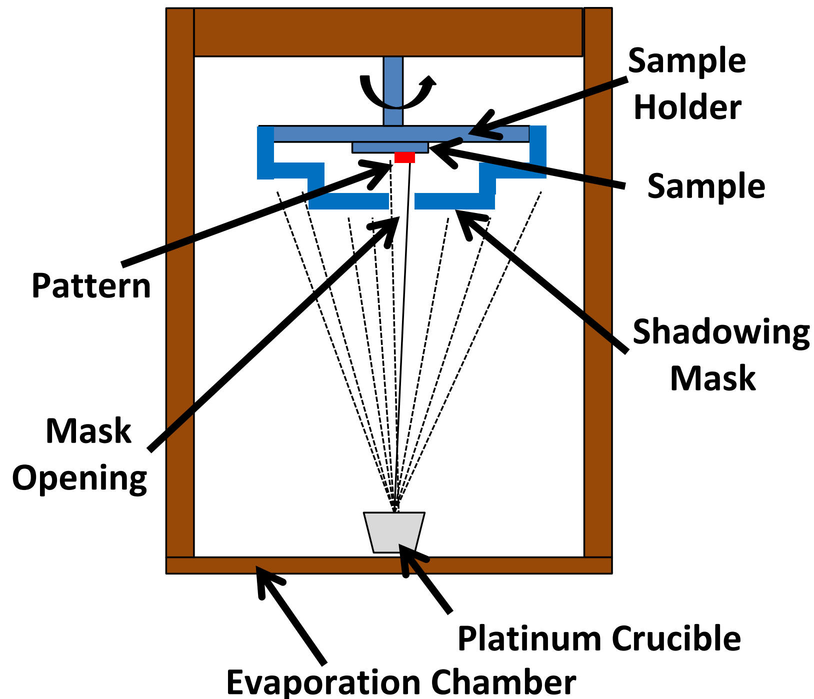

Figure 5-7 Schematic of the sample and the designed setup mounted inside the electron beam evaporation chamber.

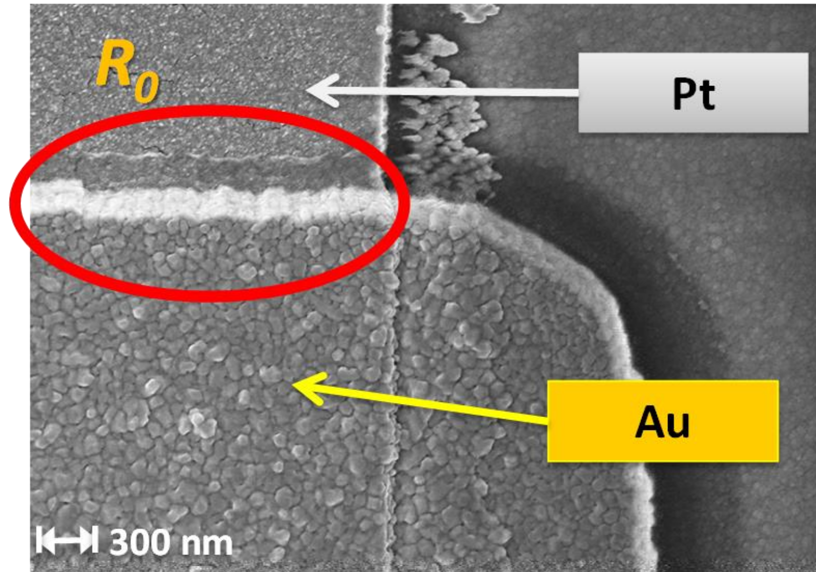

Figure 5-8 SEM image of the junction area between Pt (serpentine part of RTDs) and Au (connecting extensions of RTDs). There is a contact resistance associated with the junction of Pt and Au.

In order to determine the *ECR* value associated with each RTD, two sets of RTDs were fabricated, using the same process as explained in "RTD fabrication" section earlier. The only difference between the two RTDs is that one (R_2) has twice the number of serpentine lines than the other (R_1), shown in Figure 5-9. Therefore, assuming that both RTDs have the same electrical contact resistant, ECR_0, the resistances of the RTDs are correlated as:

$$R_2 = 2R_1 + ECR_0 \tag{6.1}$$

The average value of ECR_0 for the RTDs of the current device was extracted to be 530±10 Ω. The high value of ECR_0 can be attributed to the surface roughness and quality of the deposited Pt film. Figure 5-10 shows SEM images of the cracks in the Pt layer created during the deposition of Pt caused by film stress as a result of its high melting point. These cracks also exist at the junction of Pt-Au and can create local high electrical resistance in the junction which results in a larger value of ECR_0.

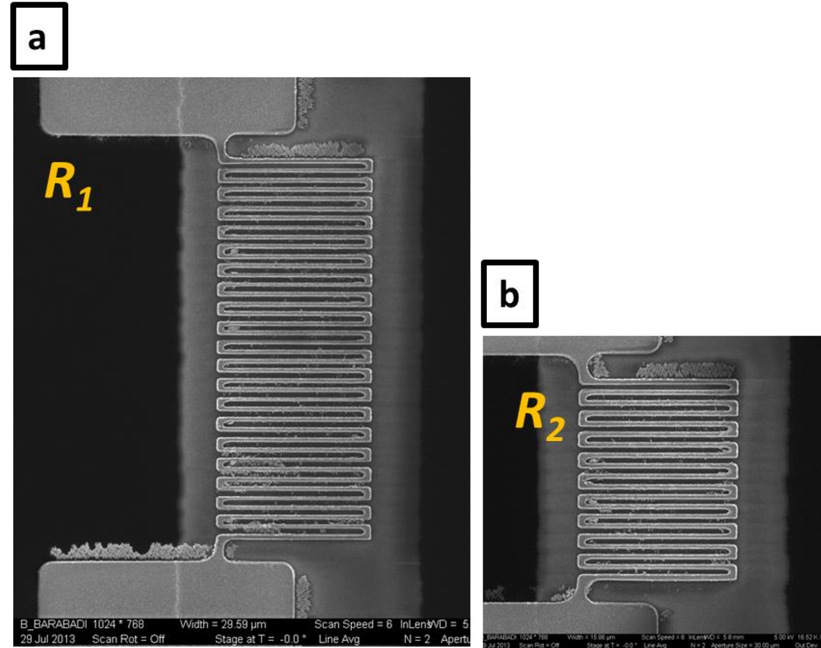

Figure 5-9 SEM image of a typical RTD with (a) resistance R1 and (b) resistance R2. $R_1 = 2R_2 + ECR_0$ where R_0 is the electrical contact resistance for the Pt-Au junction.

5.6 Device Setup and Characterization

Once the devices were fabricated, they were diced and mounted to a specially fabricated PCB board using epoxy. Figure 5-11 (a) shows a mounted device on PCB board. The PCB board has an open slot in the middle where the test device is placed and on the periphery of the board copper traces were machined. Connecting wires were then soldered into the copper traces. The contact pads on the device were wire bonded to the PCB board with 25 µm Aluminum wire. Through testing, the maximum current that can be supplied to the wire bonds before they start melting due to Joule heating was determined to be around 400 mA, which is well above the range used in this work.

In order to determine the temperature dependent electrical resistance of the RTDs, all of the fabricated RTDs on device were calibrated in a temperature controlled forced-convection oven isolated from surroundings. A T-type thermocouple with ±0.1 ºC resolution was placed inside the oven close to the device for temperature measurement.

The heater was set to different temperatures between 20 ºC and 109 ºC and the resistance of the RTDs was measured at each point. Thermal equilibrium was ensured by monitoring the rate of change of temperature at each set point and data was collected only after the rate was less than 0.05 K/min.

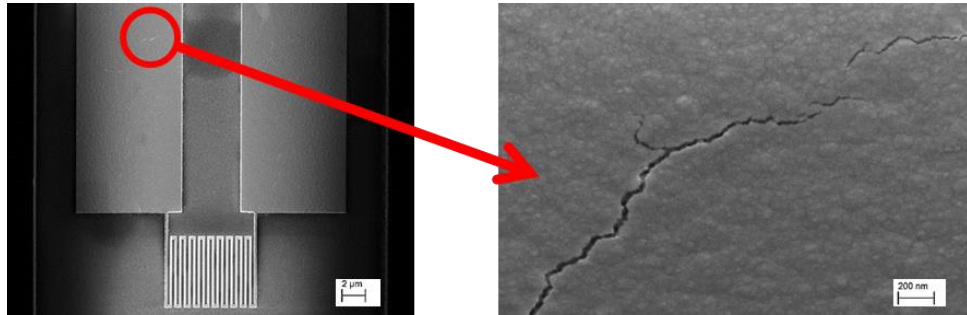

Figure 5-10 SEM images of structure showing cracks created on surface of Pt film during deposition.

A typical calibration curve for one of the RTDs is demonstrated in Figure 5-12. As one would expect, the resistance of RTDs has a linear dependence on temperature with the coefficient of determination, R^2, being greater than 0.999. Additionally, there is consistency in the calibration curve of RTDs across the entire device and wafer. Based on a linear fit to the measurements shown Figure 5-12, the temperature coefficient of the electrical resistance (*TCR*) can then be determined from:

$$\frac{R(T)-R(T_0)}{R(T_0)} = TCR \times \Delta T \quad (5.2)$$

The average *TCR* for the fabricated RTDs were found to be 0.00123 K^{-1}. The bulk value of *TCR* for Pt is 0.00392 K^{-1}.

It is important to recognize that the interconnects can serve dual purpose. In addition to being heaters as a result of Joule heating, they can be used as thermal sensors. By obtaining their calibration curves, similar to those of the RTDs, interconnects can be used to determine temperature. Therefore, in this study, the interconnects were also

characterized and their resistance-temperature calibration curves were acquired. However, they can only provide an average temperature over their entire length as compared to local readings of the RTDs over the interconnect length. Hence, they were utilized as a secondary technique to verify the temperature readings of the RTDs under steady state conditions and not the transient cases.

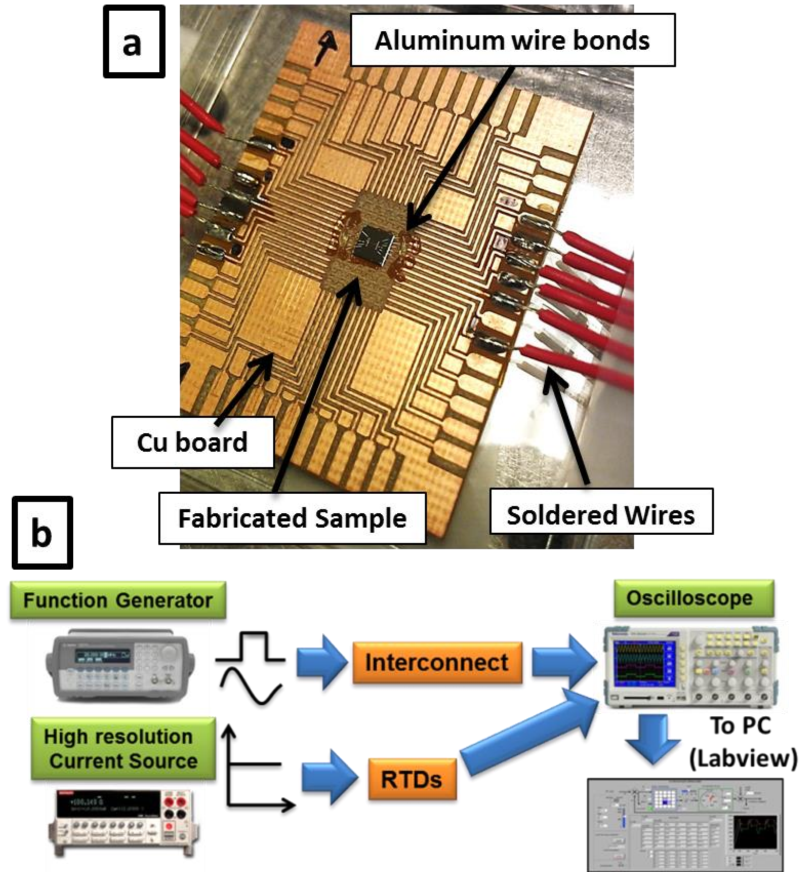

Figure 5-11 (a) Fabricated device mounted and wire-bonded to a printed circuit-board (PCB). The connecting soldered wires are shown. (b) Schematic of experimental setup.

5.7 Experimental Setup and Procedure

As stated earlier, each device consists of a set of three Cu interconnect with an array of 20 RTDs. Due to experimental limitations, only one RTD was utilized in every test. However, the tests were repeated for the same settings using different RTDs each time to reduce experimental errors and ensure consistency in the thermal measurements.

Also, to confirm the repeatability of the experiments, all three devices were tested under the same settings.

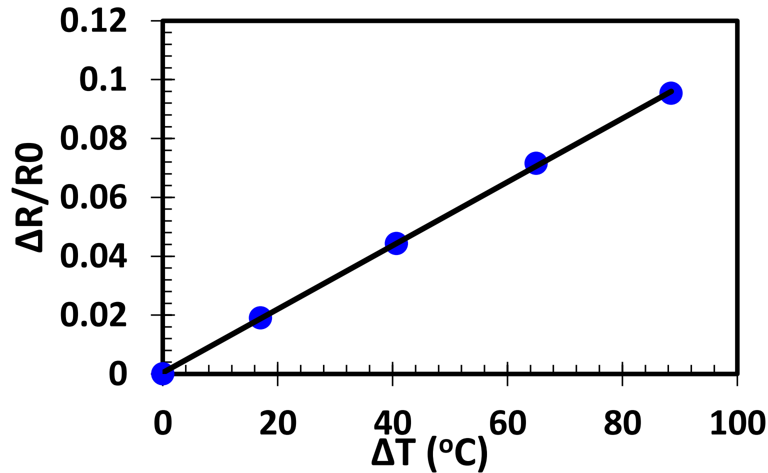

Figure 5-12 Representative calibration measurement for determining the temperature coefficient of the electrical resistance (*TCR*) for RTDs.

A constant current of 200 μA is applied to the RTD that is being used through a Keithly 2400 source meter. This current value was chosen such that the change of resistance across the RTD due to its own Joule heating is negligible. A schematic of the experimental setup is provided in Figure 5-11 (b). The input voltage to the interconnects were supplied using an Agilent 33250A function generator. For the steady state measurements constant voltages with different amplitudes were supplied to the interconnects. Whereas, for the transient tests, square and sinusoidal voltage waves with different amplitude and frequencies were applied to the interconnects. As a result of Joule heating in Cu interconnects, the device heats up causing the change in the resistance of the RTD. The voltage drop across the RTD was measured as a function of time using a Tektronix 2014B oscilloscope. Knowing the current and voltage drop as a function of time across the RTD, its resistance can be determined using Ohm's law and compared to

the calibration curve exhibited in Figure 5-12 to determine its transient temperature. To reduce the effect of the existing noise in the transient measurements, each data point was taken as an average value over 100 identical voltage inputs.

5.8 IR Imaging Setup and Procedure

Through resistance thermometry, the temperature of the interconnects at certain locations can be determined. However, it does not provide the temperature distribution across the device. In order to characterize the interconnect structure and acquire the spatial thermal map of the heaters, IR thermal imaging of the samples via Quantum focus Infrascope II (Figure 5-1 (a)) were performed. The IR thermal measurements were also used to further verify the results obtained by RTDs under steady state conditions.

The maximum spatial resolution of the microscope is 5 μm, which is twice the width of an interconnect. It has a thermal stage to heat the sample during operation and has three optical lenses (1X, 5X, and 15 X). To achieve the maximum spatial resolution, the 15X lens was used during the measurements. Sensitivity of the IR microscope is proportional to the number of photons received by the detector, which is directly proportional to the temperature of the sample. To increase the sensitivity, the stage temperature was set to 50 ^{o}C during operation. A T-type thermocouple was placed on the sample to confirm the set temperature.

One of the downsides of IR microscopy is that the emissivity of the sample is typically not known. The emissivity of the sample, however, plays an important role in the accuracy of the thermal map. A surface with a higher emissivity emits more photons and generates a stronger voltage signal resulting in a more precise temperature reading. Since metals generally have low emissivities, the surface of the devices where coated

with graphite foam which has a high emissivity. The IR camera is then focused on the surface of the device and the emissivity of the surface is determined by comparing the measured irradiance from the surface of the device to its corresponding value predicted by the Planck distribution. The red arrow in Figure 5-13 (a) points to a window on which the IR thermal imaging was performed. For the steady state measurements, at different steps, a constant current was applied to the interconnects using a Keithly 2400 source meter and voltage drop across the Interconnect was measured by a Tektronix 2014B oscilloscope. The IR thermal imaging was conducted at each step after allowing the system to reach steady state. Figure 5-13 (b) displays a representative steady state temperature map of the middle interconnect in a device acquired through IR thermometry. The temperature distribution along the arrow shown on the thermal map is shown in Figure 5-13 (c).

5.9 Results and Comparison

5.9.1 Steady state measurements

The steady state temperature of the middle interconnect as a function of input power is depicted in Figure 5-14 for various techniques. To verify the validity of the thermal measurements obtained from the fabricated RTDs, the results were plotted against IR measurements and the measurements from interconnect as a resistance thermometer. Comparing the results obtained by three different methods, the chance of error in measurements was reduced.

The maximum variation in temperature from RTDs and interconnect measurements vs. IR results are 3% and 13% respectively. Since IR thermometry provides an absolute value for temperature, the experimental errors reported here are

relative to IR measurement. The relatively larger difference in RTD vs. the interconnect measurements can be primarily attributed to the placement of the RTD in the center of the interconnect, which has a higher temperature with respect to the edges of the interconnect. As exhibited in Figure 5-15 (b), the serpentine section of an RTD fits in a 6 μm by 7.7 μm rectangle and does not cover the entire width of the interconnect (10 μm). As a result, the RTD detects a higher value for the temperature than the interconnect itself as the RTD provides the average temperature of only the central portion of the interconnect as opposed to its entire width (Figure 5-15 (c)).

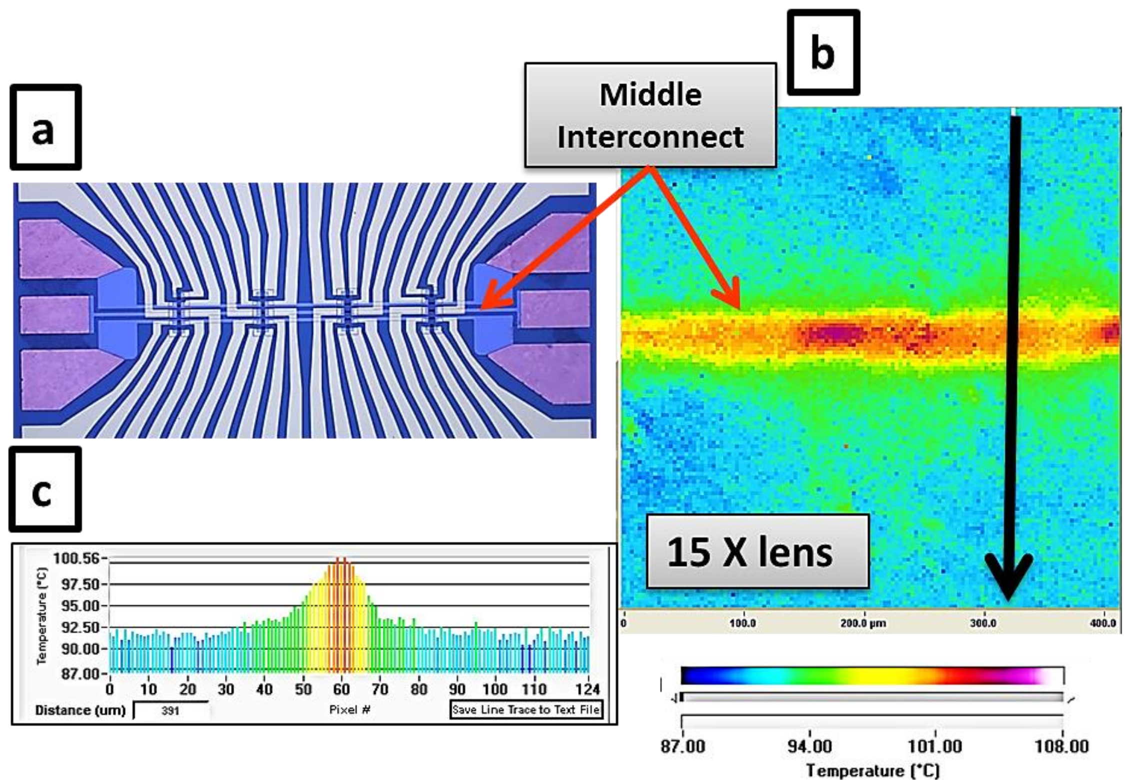

Figure 5-13 (a) Optical image of a representative fabricated device specifying the location were the IR measurements were obtained. (b) A representative temperature map attained by steady state IR thermometry. (c) Plot of temperature distribution along the arrow shown on the thermal map.

Another reason for this temperature different is the experimental error associated with poor signal to noise ratio in the measurements. By increasing the input current of the

RTD, the voltage drop across the RTD will increase, which results in an improved signal to noise ratio. In the current design, however, the maximum allowable current applied to the RTD is restricted because of the low quality of the deposited Pt film shown in Figure 5-10. The cracks in the film create local high resistances in the Pt film which limit the maximum allowable input current. As the input current increases while the device goes through thermal cycling, stresses created in the film cause crack propagation, which in some cases results in device failure [1, 129]. Another explanation is that as the device goes through thermal cycling, cracks can deform and change the total resistance of the RTD. In such cases, the temperature coefficient of the electrical resistance (*TCR*) is no longer a constant and will vary with temperature change resulting in higher error in thermal reading obtained through RTDs.

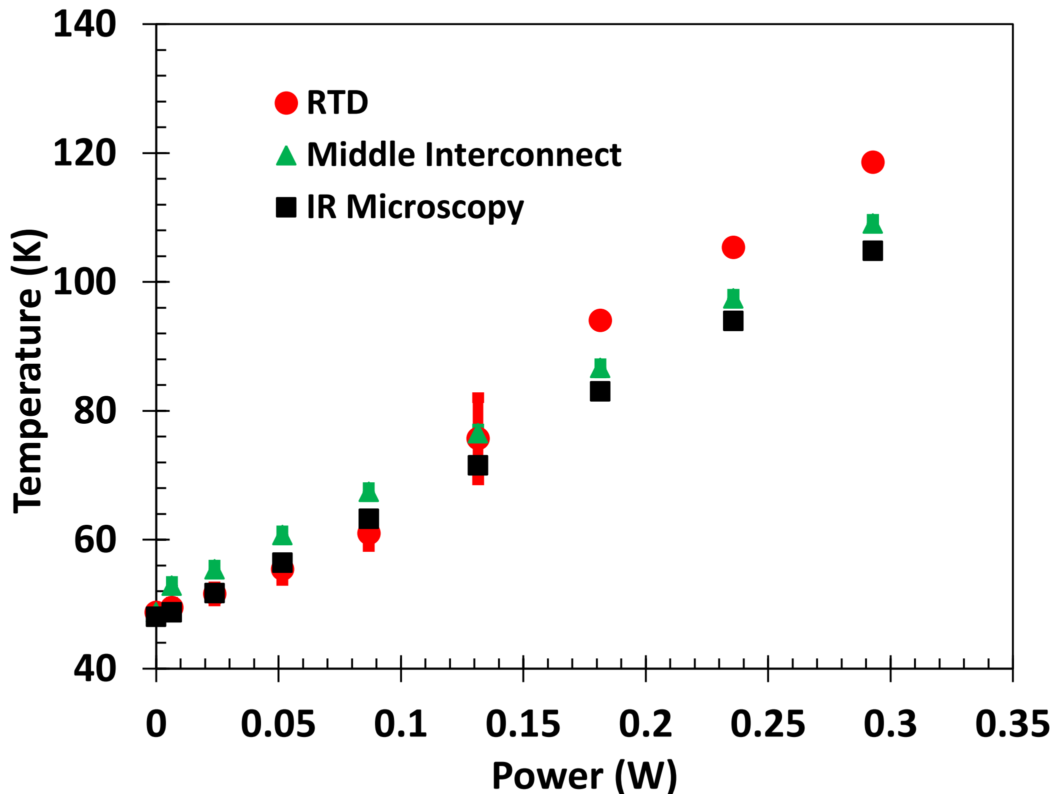

Figure 5-14 Comparison of temperature measurements of the middle interconnect using RTDs, IR Microscopy, and the Interconnect as a resistance thermometer.

5.9.2 Numerical model for verification

Once the RTD measurements were verified against other measurement techniques for steady state condition, a numerical model was developed to evaluate the RTD measurement under both steady state and transient conditions. The numerical modeling is also useful to extract detailed information regarding the characteristics of the transient measurements such as thermal time constant.

For simplification, a 2-dimensional transient finite element (FE) model of the cross-section of the structure, displayed in Figure 5-2 (c), was developed in COMSOL. The convergence of the FE model was verified with respect to the solver type, time step, and time integration method. The FE model consists of 29,942 elements, chosen such that the numerical results are independent of mesh size. Natural convection was assumed on the top boundary and constant ambient temperature conditions were applied to the rest of the boundaries. Left and right boundaries were chosen 100 μm away from the structure, ensuring that isothermal conditions were achieved. Table 5-2 provides the material properties used for numerical modeling. It should be noted that the thermal properties for deposited materials can vary depending on deposition method. Here, these values are taken from Ref. [130] based on their deposition technique. Figure 5-15- (a) shows a representative plot of quasi-steady spatial temperature distribution at the cross section of structure based on a pulsating volumetric heat generation of 657 MW/cm^3 at 10 kHz frequency.

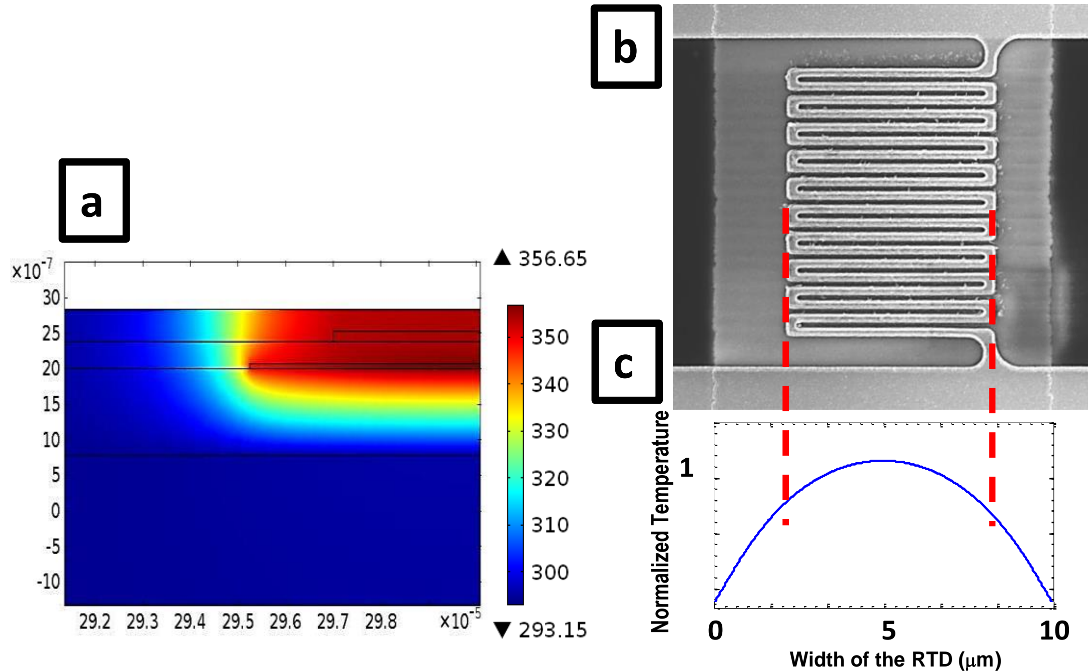

Figure 5-15 (a) A representative quasi-steady spatial temperature distribution at the cross section of structure and (c) along the width of interconnect. (b) SEM image of the placement of a typical RTD over an interconnect line.

Figure 5-16 depicts a representative plot for experimental (diamond shaped black markers) and numerical (circular red markers) steady state temperature of the middle interconnect. The employed RTD used for these data is located at 212.3 μm distance from one edge of the interconnect as indicated in Figure 5-2 (a). The maximum relative error is 3.8% which indicated that there is a good agreement between the experimental and numerical results.

5.9.3 Transient measurements

To characterize the device for transient measurements, the step response and frequency response of the system were determined.

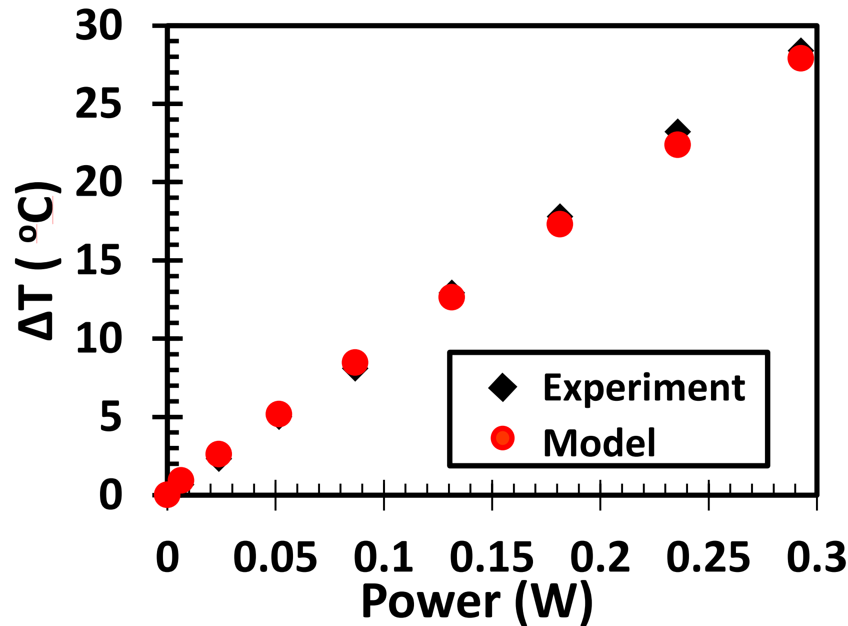

Figure 5-16 Experimental (diamond shaped black markers) and numerical (circular red markers) steady state temperature of the middle interconnect.

5.9.3.1 Step response of the system

In order to determine the step response of the system, a single step voltage was applied to the interconnects. To make a more detailed comparison between the experimental and FE results, the normalized temperature rise time history for a 70 μs window around the rising edge of the input signal is plotted in Figure 5-17. The maximum power input to the middle interconnect was 0.156 W. The thermal response was normalized based on the minimum and maximum temperature in every experiment or numerical model. The solid black line indicates the experimental data and the dashed red line represents the FE results. It can be inferred from Figure 5-17 that the numerical data closely traces the experimental measurements with matching trends. The rise-time (fall-time) is defined as the time for the temperature to rise (fall) from 10% to 90% (90% to 10%) of the steady-state value when a step signal is applied to the interconnect. The rise time is calculated to be 11 μs and 9 μs for the FE modeling and experimental data

respectively. In general, having such a small thermal time constant makes submicron RTDs reliable devices in measuring rapid temperature transients in interconnect architectures. As previously stated the effect of the existing noise in the transient measurements was reduced by averaging the measurements over 100 identical voltage inputs.

Table 5-2 Modeling Parameters

	Thermal Conductivity (W/mK)	Density (Kg/m^3)	Specific Heat Capacity (J/KgK)
Si	163	2330	703
SiO_2	1.4	2200	730
Cu	401	8960	384
Pt	71.6	21450	133

5.9.3.2 Frequency response of the system

Frequency response is the quantitative measure of how the system responds to different frequencies of the input signal. In this study, the input of the system is the applied electrical power to the interconnect and the output is the voltage change of the RTD, which is proportional to the temperature change of the interconnect. To determine the frequency response of the system a series of sinusoidal voltage inputs with frequencies in the range of 2 - 200 kHz were applied to the interconnects. The frequency response of the system was acquired at each frequency.

Figure 5-18 exemplifies a typical frequency response curve of the device. The black markers are the experimentally obtained frequency response data and the solid red

line is a first-order linear time invariant (LTI) system fitted to the data. The bandwidth (also known as the 3dB frequency), f_{3dB}, is defined as the frequency where the signal amplitude drops to $1/\sqrt{2}$ of its steady-state response. The f_{3dB} of the device was determined to be 95 kHz. This means that RTD can follow input power fluctuations with lower than 95 kHz and therefore the transient thermal measurement can be obtained with good accuracy for this range of input frequencies. Therefore, the current design can be utilized as a robust measurement technique for rapid transient thermal events in microelectronics.

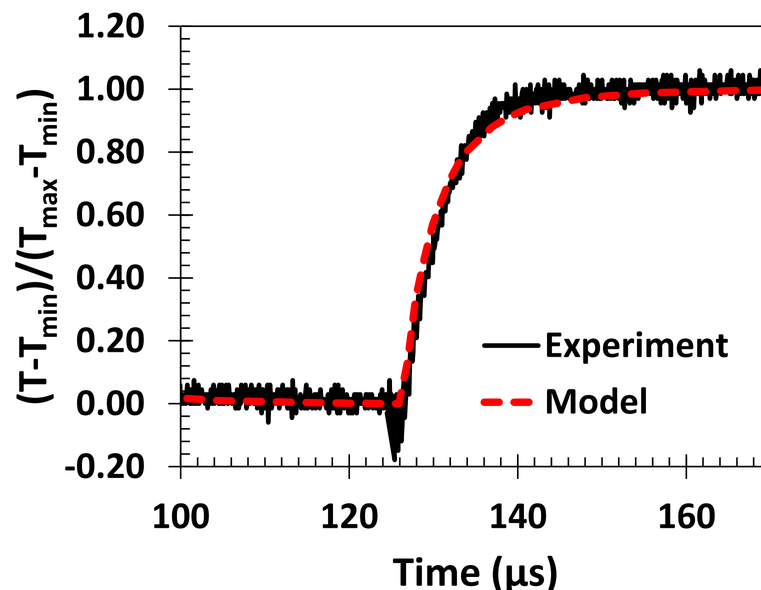

Figure 5-17 Experimental (solid black line) and numerical (dashed red line) normalized step response of the middle interconnect subjected to a 10 kHz square pulse (rising edge). Thermal time constant on the rising edge: Model: 11 µs; Experiment: 9 µs

5.9.3.3 Transient response to different types of input power

Transient thermal response of each device to the input powers of different kinds, frequencies, and amplitudes were also investigated. For a better comparison between the input power, thermal measurements, and numerical results, all data were normalized based on their minimum and maximum values.

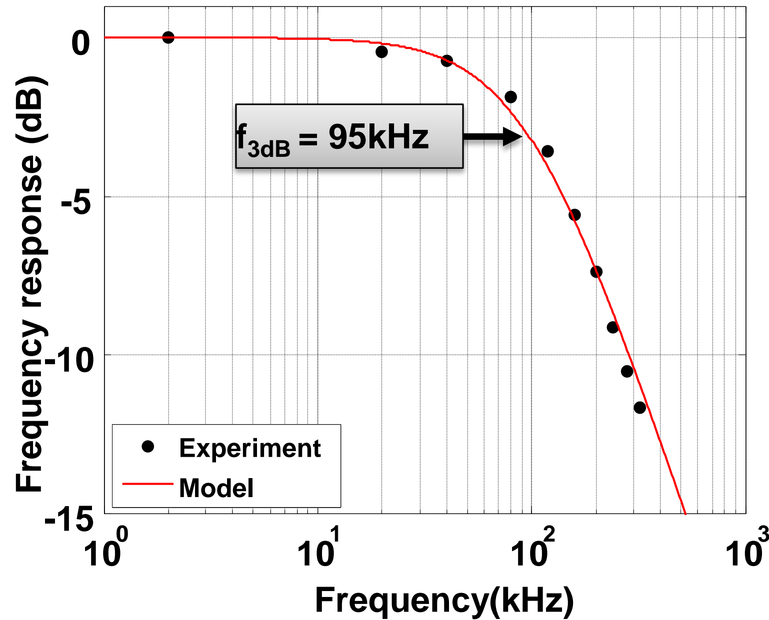

Figure 5-18 Frequency response of the device. Black markers represent the experimentally obtained data and the solid red line is a first order model fitted to the data. The f_{3dB} of the system is 95 kHz.

Figure 5-19 displays the experimental (solid black line) and numerical (dashed red line) normalized transient thermal response of the middle interconnect. The interconnect was subjected to a 10 kHz square pulse (dotted blue line) with the amplitude of 0.156 W at 50% duty cycle. The primary vertical axis is for the input power and the secondary vertical axis is associated with the experimental and numerical temperature data. The results were shown for an arbitrary window of 250 μs. The FE results closely agree with the measurements. However, it can be observed in Figure 5-19 that spikes occur in the experimental measurement exactly when the input voltage pulse is applied or removed. These noises in the measurements are produced because of the way the trigger signal for the oscilloscope is being provided. The trigger signal is supplied from the input voltage to the interconnects that is generated by the function generator. Therefore, at the

beginning of every pulse, spikes are introduced to the measured signals in the oscilloscope. Incorporating a low pass filter can reduce and potentially remove this effect.

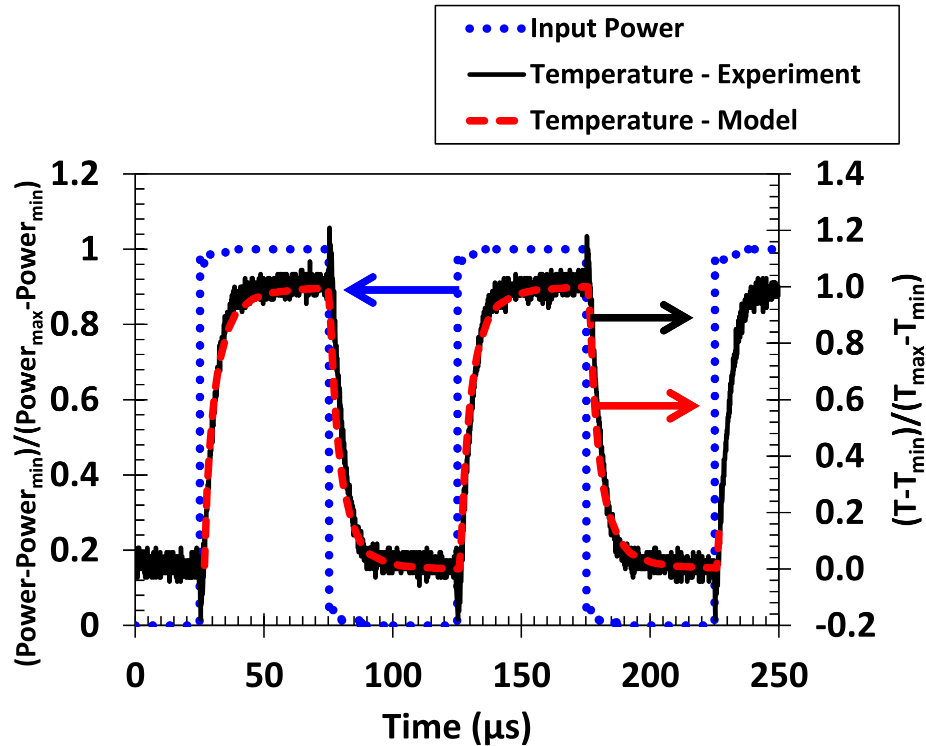

Figure 5-19 Experimental (solid black line) and numerical (dashed red line) normalized transient thermal response of the middle interconnect subjected to a 10 kHz square pulse (dotted blue line). The primary vertical axis is for the input power and the secondary vertical axis is associated with the experimental and numerical temperature data.

Similarly, Figure 5-20 exhibits the normalized transient thermal response of the middle interconnect subjected to a 10 kHz sinusoidal input power with the amplitude of 0.156 W (dotted blue line). The primary vertical axis is linked to the normalized input power and the secondary vertical axis is associated with experimental measurements (solid black line) and numerical results (dashed red line) of the normalized temperature. We observe a thermal delay of 9 µs between the input signal and the temperature (i.e., output). This is very close to the thermal response-time of the system shown in Figure 5-

17. This is not a surprising result as the thermal response-time (i.e., rise/fall time) and the temporal delay to a sinusoidal input are equal in a first-order LTI system.

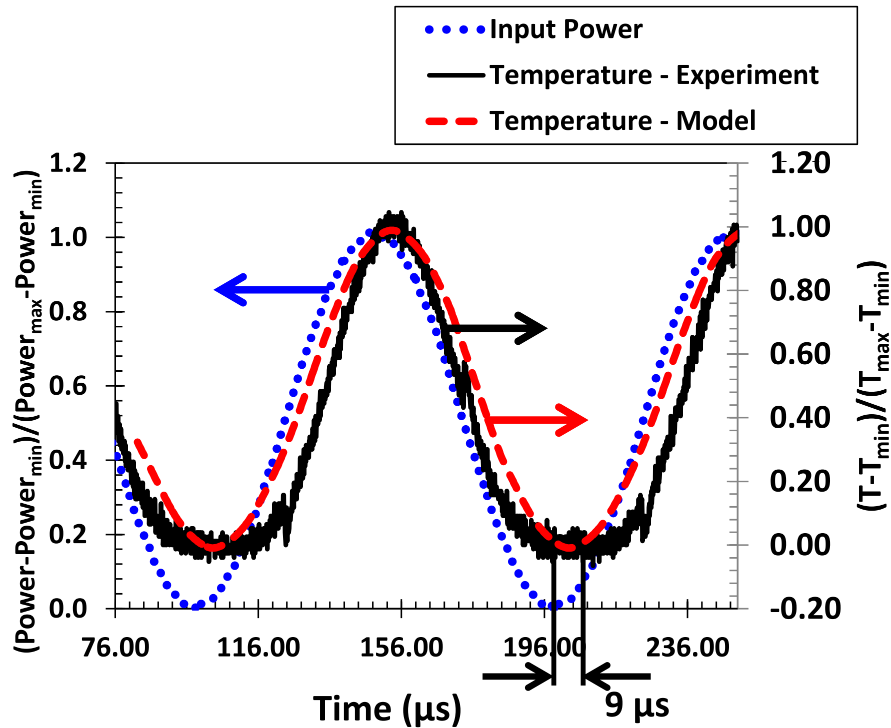

Figure 5-20 Experimental (solid black line) and numerical (dashed red line) normalized transient thermal response of the middle interconnect subjected to a 10 kHz sinusoidal input power (dotted blue line). The primary vertical axis is for the input power and the secondary vertical axis is associated with the experimental and numerical temperature data.

5.10 Conclusion

In this chapter, a new test structure was developed for the investigation of the effect of rapid transient power input profiles with different amplitudes and frequencies in Cu interconnects. A set of three 70-nanometer-thick Cu traces buried in low-k dielectric materials was designed and fabricated to resemble nanoscale on-chip interconnects. In addition, a unique set of 20 embedded sub-micron resistance thermometers (RTDs) were fabricated above the interconnect layer with a barrier layer of silicon dioxide for transient thermal measurements. Each RTD fits in a 6 µm by 7.7 µm rectangular space and their relatively small size provides a high spatial resolution thermal monitoring solution, which

can be employed for the detection of small local hotspots. For steady state condition, RTD measurements were validated by comparing them to other measurement techniques such as IR microscopy. The results of different measurement techniques are within 13% of each other validating the correct calibration of these techniques. In transient analysis, the measurements were validated against FE modeling.

In addition, the step response and thermal time constant of the system was determined (9 µs rise time). The frequency response of the system was also achieved and the f_{3dB} of the device was determined to be 95 kHz. The RTDs fairly small thermal time constants during transient measurements and relatively high f_{3dB} make them powerful measurement devices in monitoring temperature during thermal scenarios where there are abrupt rapid changes in voltage or current levels of the system. The maximum frequency of operation of the structure is only limited by the heat diffusion of the interconnects to the underlying substrate which is enforced by the structure under study and cannot be further increased through design of the RTD. Furthermore, the demonstrated technique has the capability to be integrated within the device and is specifically suitable for embedded structures. Therefore, these devices are very useful for measuring the spatial-temporal response of the interconnects in a buried 3-D IC structure, where measurement techniques such as infrared or thermoreflectance microscopy cannot be used.

CHAPTER 6: THERMAL CONDUCTIVITY MEASUREMENTS OF NANOSCALE EMBEDDED METALLIC THIN FILMS

6.1 Importance of Size Effects

Performance and reliability design of future microelectronic and nanoelectronic systems requires knowledge of the thermal and electrical properties of thin films. As the size of a metal interconnect becomes comparable to, or smaller than the electron mean free path (~40 nm in Cu at room temperature), electron transport becomes dominated by scattering at the metal-dielectric interface, and at grain boundaries. This scattering can reduce the electrical and thermal conductivity to less than half of the bulk value [5-10]. This reduction in conductivity has been explained by the Fuchs-Sondheimer model [5, 6], and subsequently more refined models [7-10]. This confirms the need for experimental methods for measurement of thermal properties of thin film materials for interconnects and dielectrics.

6.2 Objective

This chapter investigates the size effect on the thermal conductivity of embedded thin metallic (here Cu) films. One of the applications of Cu thin films is nano-scale embedded on-chip interconnects typically used in IC industry.

6.3 Measurement of Thin Metallic Film Properties

Various techniques have been investigated for the measurement of thin film thermal properties, more specifically, thermal conductivity for thin metal layers. Some of

the well-established measurement techniques are described briefly below. A thorough review on various micro and nanoscale thermal characterization techniques is provided by Christofferson et al. [121].

Nath and Chopra [53] introduced techniques to measure thermal conductivity of metallic thin films. The outline of their setup is shown in Figure 6-1 taken from [53]. They used both steady state and transient measurements. The only different between the two measurements is the way they measure heat flow. A U-shape copper block and a heating element act as a heat source. At the other end of the structure the lead sheet is considered to be the heat sink with thermocouples measuring the change in temperature. One dimensional heat diffusion equation is used to model the metal film and mica (substrate material) double layer. Thermal conductivity was determined by using a bare mica film and a metal deposited mica film. In a similar steady state approach, Pompe and Schmidt [54] developed a technique for thermal conductivity measurements with different heat source and heat sink.

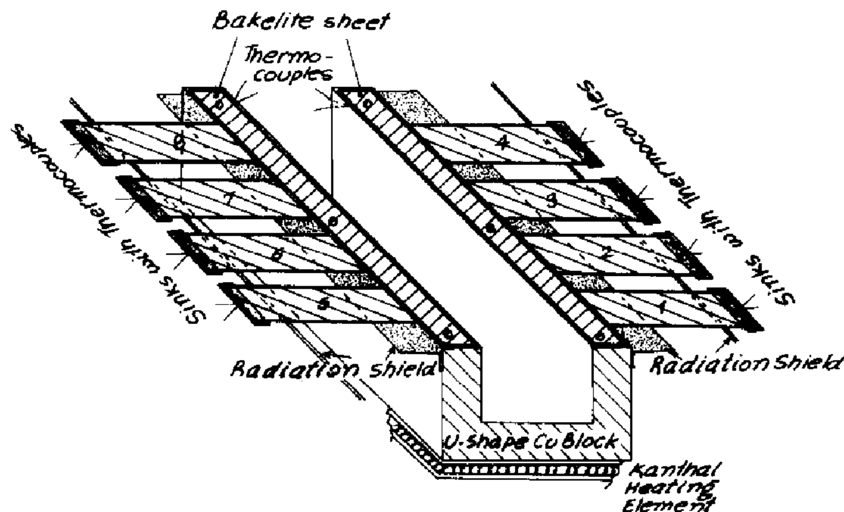

Figure 6-1 Layout of Nath and Chopra's experimental structure used to measure thermal conductivity [53].

Volkelin and Baltes [131] used micromachining to suspend a cantilever structure. The cantilever has a heater and a temperature sensor patterned on one side and is attached to a silicon substrate on the other side. Volkelin and Baltes [131] used this structure to measure the thermal conductivity of poly-silicon. They incorporated a one dimensional heat diffusion model for all the layers on the cantilever. Later on, von Arx et al. [132] applied this technique to measure thermal conductivity of CMOS process materials such as silicon dioxide, silicon nitride, and aluminum.

Shojaei-Zadeh et al. [55], Zhang et al. [56], and Liu et al. [57] used a suspended micro-fabricated metal bridge, shown in Figure 6-2 from [57], for thermal conductivity measurements. More specifically, Liu et al. [57] measured the lateral thermal conductivity of thin copper layers of thicknesses 50 and 144 nm at temperatures between 40 and 400 K, using electrical-resistance thermometry. The temperature gradient in the bridge was caused by Joule heating.

Nonetheless, there are limitations to the discussed steady state approaches. There is an experimental error associated with these methods due to the use of thermocouples and other thermal sensors. Additionally, almost all of these techniques require a suspended metal bridge or a combination of a metal and low-K material bridge that requires significant micro-fabrication. Although, utilizing Joule heating in a suspended metal bridge is a robust technique for thermal conductivity measurements of below 100 nm metal lines, such as interconnects, these structures do not account for the interfacial region. Using fabrication techniques, such as etching, to undercut the substrate can potentially change the quality of the interface. To obtain a realistic value for thermal conductivity that represents a desired structure, the original interface should be preserved.

This demonstrates the need for a technique to measure the thermal conductivity for embedded thin metallic layers without changing the original structure.

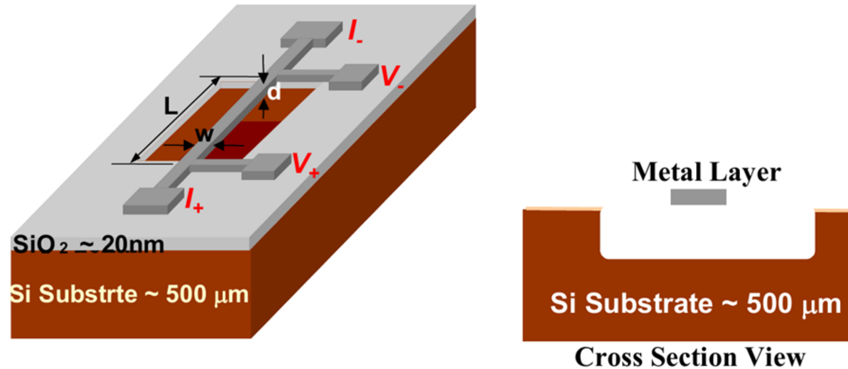

Figure 6-2 Layout of the (a) suspended structure and (b) a cross section of the metal structure used by Liu et al. [57] to measure thermal conductivity.

Transient experiments and modeling have also been reported in literature. Kelemen [58] introduced a transient approach to measure the thermal conductivity of the metallic thin films. A pulsed heat source was applied at one end of the film and the temperature was measured at two points along the film. A one dimensional heat diffusion model was employed to determine the thermal conductivity.

Another method for measuring the thermal conductivity of free-standing thin films was presented by Kemp et al. [133]. In this technique, surface of the film was scanned through a sinusoidally modulated laser beam and the thermovoltage generated at a fixed point on the surface of the film was monitored. In a similar approach, Langer et al. [134] used thermoreflectance from the film to determine the temperature rise. Thermal diffusivity of the thin film was determined through measurements at two different locations.

Another established technique to measure thermal conductivity of bulk substrates and thin films is the 3ω method that is originally developed by Cahill [59]. Lu et al. [60] implemented the 3ω technique to measure specific heat and thermal conductivity of

suspended thin platinum wires. Yang and Asheghi [61] further advanced the technique such that the suspension of the wire was no longer required. Nevertheless, to reduce the heat conduction to the substrate, the underlying silicon dioxide had to be etched away from the sides and yet an extensive three dimensional numerical analysis was performed to account for the remaining substrate.

The main drawback of the described transient methods is that they cannot be directly applied to the sub 100 nm metallic layers such as interconnects except for the method proposed by Yang and Asheghi [61]. Their technique also has some limitations. Yang and Asheghi [61] reported a very low sensitivity due to the effect of substrate in spite of their extensive three dimensional numerical analysis that is required for the model. To obtain the best sensitivity, the length of the interconnect needs to be only a few microns. Similarly, the interfacial defects along the interconnect were not taken into account.

The final two techniques introduced here are considered as high spatial resolution temperature measurement approaches that are mainly used to characterize thermal transport in nanoscale interconnects. First method is known as Pump-Probe Transient Thermoreflectance (PPTTR) Technique was explained in Chapter 5. PPTTR method has a high spatial and temporal resolution and it is fully optical with no contact to the sample.

The other approach is Scanning Joule Expansion Microscopy (SJEM) which is an AFM based technique to extract in-plane thermal conductivity of thin metallic films whose thickness is comparable to the electron mean free path. This technique was briefly introduced in Chapter 5. Figure 6-3, adapted from [46], shows the schematic diagram for an experimental setup that is used for Scanning Joule Expansion Microscopy. SJEM

measures the periodic thermal expansion amplitude at the sample surface, which corresponds to the periodic temperature at the surface. This technique is capable of mapping the temperature gradient near a constriction between wide and narrow metal lines. Thermal conductivity is extracted by using a numerical fit to the measurements. Extracted thermal conductivities of thin gold films showed good consistency with predictions from Wiedemann-Franz for 43 nm and 131 nm gold films. SJEM is a powerful technique as it has no dependence on tip-sample heat flow.

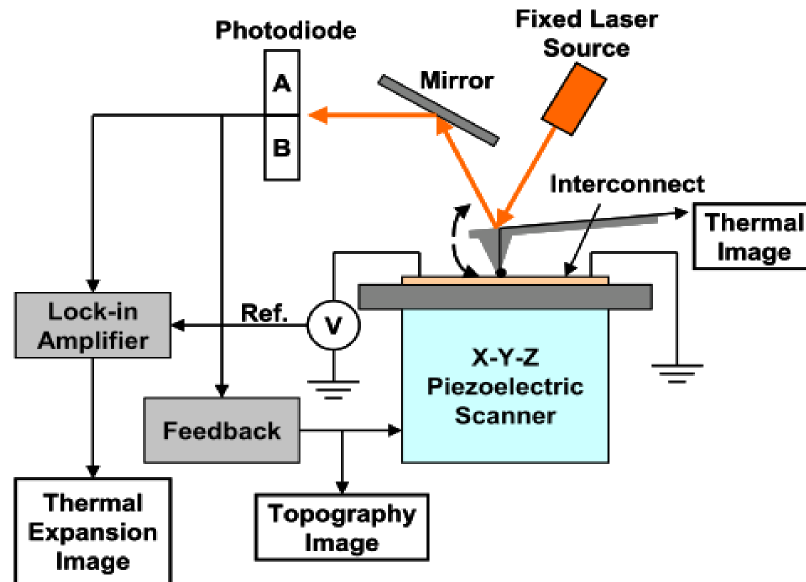

Figure 6-3 Schematic diagram of the experimental setup for the scanning Joule expansion microscopy (SJEM) adopted from [46].

Gurrum et al. proposed an approach to extract in-plane thermal conductivity of the 43 nm and 131 nm gold films using SJEM with 10 nm resolution. They used a three-dimensional FE model of the frequency-domain heat transfer problem, to fit the in-plane thermal conductivity to the measured data. Gurrum, et al. determined that for a heating frequency of 100 kHz, the thermal conductivity of a 43 nm film was 82 ± 7.7 W/mK, and for a heating frequency of 90 kHz, the thermal conductivity of a 131 nm film was 162 ±

16.7 W/mK. These values are significantly smaller than the bulk thermal conductivity of 318 W/mK for gold, showing thermal conductivity dependence on size due to electron scattering [135].

The main disadvantage of this method is that only an AC temperature change can be measured. Also, the amplitude of expansion is highly dependent on the heating frequency, the underlying layers dimensions, and thermal properties. A slight mismatch in the coefficient of thermal expansion (CTE) can result in substantial measurement errors. It is also important to note that both PPTTR and SJEM are stand-alone measurement tools that require extensive setting up procedures. These limitations justify the need for a thermal conductivity measurement technique that can be integrated within the structure and accounts for the interfacial effects.

6.4 Proposed Approach: Steady State "Hourglass" Design

In summary, most of the currently available techniques require a suspended/free-standing bridge that eliminates the effect of heat diffusion to the substrate. Even the methods that partially took into account for the substrate, reported a very low sensitivity and extensive numerical modeling [61]. High spatial resolution temperature measurement approaches (i.e., PPTTR and SJEM) that are mainly used to characterize thermal transport in nanoscale interconnects are also highly dependent on the underlying layers dimensions and thermal properties and a slight mismatch in the coefficient of thermal expansion (CTE) can introduce large inaccuracies. In addition, both PPTTR and SJEM are stand-alone apparatuses and cannot be integrated within the structure. Therefore, the objective of this work is to develop a technique that:

- Accounts for the effect of the substrate and interface.

- Has satisfactory sensitivity to the thermal conductivity of the metallic film.
- Can be integrated within the structure and be used for measurements of embedded or buried structures such as nanoscale on chip interconnects.
- Does not require extensive micro-fabrication.

It is greatly challenging to measure the thermal conductivity of an embedded metallic film due to a much higher resistance of the low-K insulation layer and the substrate than that of the thin metallic film itself. In other words, the thermal conductivity of the metallic film is a fraction (less than 1%) of the total thermal resistance of the structure. This results in very low sensitivity of the overall thermal resistance (in the direction normal to the plane of the structure) to the thermal conductivity of the metallic layer. As a result slight variations in the thermal resistance of the insulation layer (due to the variation in the thermal conductivity or thickness) results in a large error in the estimation of the thermal resistance of the metallic film and consequently its thermal conductivity. Here, we propose a new concept to induce strong sensitivity of the heating structure to the thermal conductivity of the metallic layer. Since the thermal conductivity of the metallic layer is much higher than that of the insulation layer, in-plane heat diffusion is considerably higher in the metallic layer as compared to the insulation layer. In order to exploit this property to devise a structure that is sensitive to the thermal conductivity of the metal we need to induce in-plane heat diffusion, which necessitates lateral heat gradient. For this purpose we propose to use a laterally varying resistor structure to produce lateral heat gradient and to induce lateral heat diffusion in the plane of the metallic layer.

Figure 6-4 (a) shows the layout of the proposed structure used to investigate the lateral thermal conductivity of embedded thin metal films (Not to scale). The structure consists of a Si substrate with a layer of silicon dioxide and a thin metal film on top that is design in a shape of a constriction referred to as "hourglass". The top view of only the metal layer is demonstrated in Figure 6-4 (b). As shown, w_r is the width of the constriction, r_R is the radius of the constricted area and r_i is the radius of the half circles used to create the outer part of the hourglass geometry. Passing a steady electrical current through the metal layer will cause non-uniform Joule heating along the hourglass structure. This will result in a non-uniform temperature rise in the structure and ultimately changes the total resistance of the hourglass metal layer. A numerical model representing the proposed geometry would be developed. By comparing the experimental data of electrical resistance versus input power (R vs. P) with that of the numerical results the thermal conductivity of the copper layer can be determined.

The first step is to study the sensitivity of the structure to the change in the metallic layer's thermal conductivity. Given the complexity of this geometry, analytical solution is not trivial for this problem. Therefore, a three dimensional electro-thermally coupled nonlinear finite element (FE) model was developed in COMSOL representing the proposed structure. Since the electrical resistivity of the metal layer is also dependent on temperature, the model is setup such that it simultaneously solves for electrical and thermal field. Thermal boundary conditions are natural convection on the top surface and room temperature on the bottom. Due to the high aspect ratio of the geometry, the side walls of the structure were assumed to be thermally insulated. Electrical current enters from the left side of the hourglass metal layer.

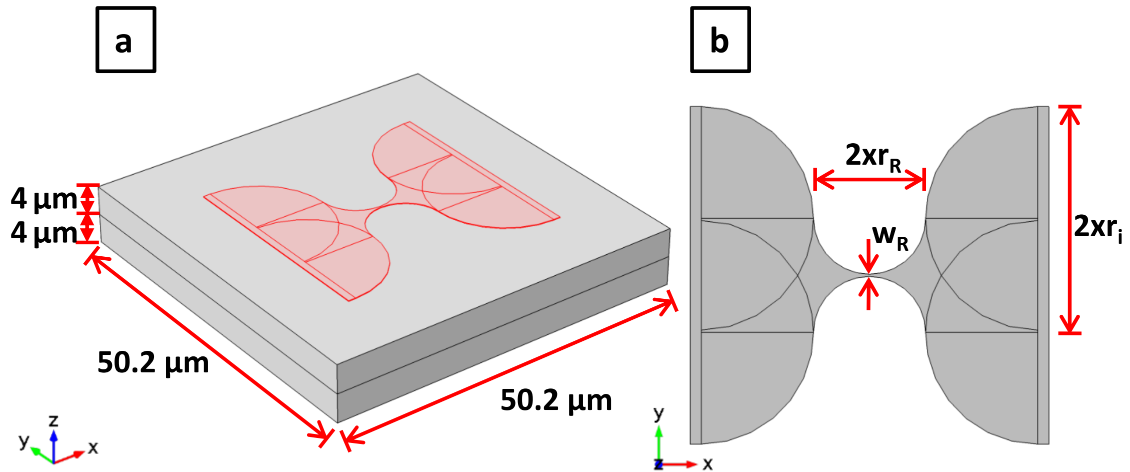

Figure 6-4 (a) Layout of the proposed structure used to investigate the lateral thermal conductivity of embedded thin metal films, (b) Top view of the metal design referred to as "hourglass". As shown, w_R is the width of the constriction, r_R is the radius of the constricted area and r_i is the radius of the half circles used to create the outer part of the hourglass geometry.

Figure 6-5 (a) shows a representative plot of spatial distribution of current density in the proposed geometry. For this analysis w_R = 300 nm, r_R=1 um, and r_i= 5 um. The thickness of the metal and underlying oxide layer is set at 100 nm and 4.5 µm respectively. For demonstration purposes the input current density is set at a high value of 0.1 A. Figure 6-5 (b) exhibits the spatial temperature distribution in the structure with thermal conductivity K=240 W/mK. The value chosen for the input current in this simulation is not realistic and can potentially cause electromigration or defects in the constricted area. However, the effect of thermal conductivity variation is better visible at this current range. Therefore, this value of current was selected for the proof of concept demonstration of the proposed measurement scheme.

Figure 6-6 demonstrated the top view of temperature distribution in the proposed structure for thermal conductivity varying from 140-240 W/mK with equal intervals of 5% change in thermal conductivity. To achieve a better visual comparison when changing thermal conductivity, the input current was set at 0.1 A and the geometrical

values were kept the same as in Figure 6-5, i.e. w_R = 300 nm, r_R = 1 μm, r_i = 5 μm. It can be visually observed that by changing the thermal conductivity of the hourglass metal film from 140 (35% of the bulk value) to 240 (60% of the bulk value), the lateral diffusion depth as well as the maximum temperature varies. More specifically, the proposed structure is noticeably sensitive to even 5% change in thermal conductivity of the metallic film which qualitatively satisfies the initial set requirements for the model. Nonetheless, to quantify and ultimately increase the sensitivity of the structure to thermal conductivity of the metal film, a series of numerical modeling with variations in w_R, r_R, r_i, and the thickness of the underlying silicon dioxide layer were performed.

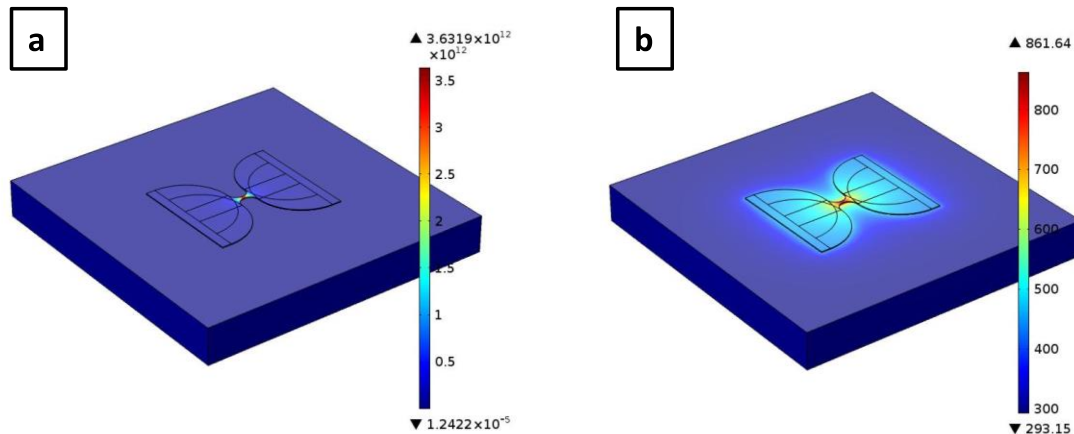

Figure 6-5 Spatial distribution of (a) current density and (b) temperature in the proposed structure. For this model w_R = 300 nm, r_R = 1 μm, r_i = 5 μm, and thermal conductivity *K*=240 W/mK.

Figure 6-7 displays representative plots of resistance vs. input power for different values of thermal conductivity for three combination of w_R , r_R, r_i.

- w_R = 300 nm, r_R = 3 μm, r_i = 10 μm
- w_R = 200 nm, r_R = 3 μm, r_i = 10 μm
- w_R = 200 nm, r_R = 5 μm, r_i = 10 μm

For all three cases, the thickness of the metal and underlying oxide layer is set at 100 nm and 4.5 μm respectively. Electrical resistivity of Cu at room temperature was considered to be 1.72×10^{-8} Ω·m. It can be seen that for every 10% change in thermal conductivity, there is approximately (a) 12.5-15%, (b) 13-15.7%, and (c) 12.4-14.7% change in the slope of resistance vs. input power. The change in the slope of (R vs. P) can also be interpreted as the sensitivity of the structure to the change in the metallic layer's thermal conductivity. By comparing case (a) and (b), it can be inferred that will improve the sensitivity.

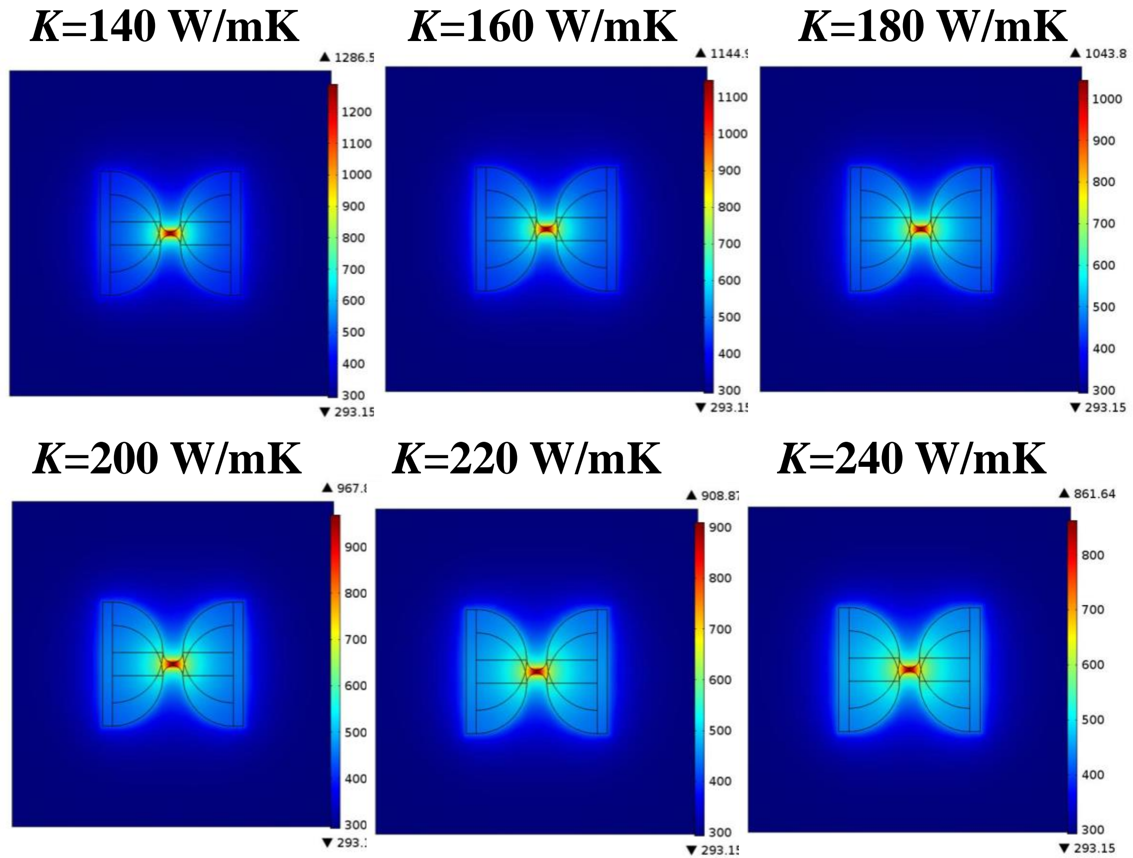

Figure 6-6 Top view of temperature distribution in the proposed structure for thermal conductivity varying from 140-240 W/mK. For this model w_R = 300 nm, r_R = 1 μm, r_i = 5 μm. To achieve a better visual comparison when changing thermal conductivity, the input current was set at a high value of 0.1 A.

As shown in Figure 6-7 (a) vs. (b), the sensitivity can be improved decreasing the width of the constriction, $w_{R,}$ (here from 300 nm to 200 nm). The decrease in w_R results in an increase in the resistance of the constricted area. This will lead to an increase in the temperature gradient in lateral direction and ultimately a rise in the slope of resistance vs. power. Similar analogy holds for an increase in the radius of the constricted area, r_R resulting in an increase in the slope of resistance vs. power (Figure 6-7 (b) vs. (c)). The same conclusion can be drawn for the increase in r_i (not shown in the plot).

6.5 Design and Fabrication

Once the effect of the geometrical parameters on the proposed model was studied, for the actual device, the thickness of the underlying oxide layer was chosen to be 4.5 μm. To have a relatively high sensitivity without further complicating the micro-fabrication steps, the values of r_R and r_i were set at 5 μm and 15 μm respectively. In order to have a higher value for the total resistance of the structure and to reduce the resistance measurement relative error, instead of one hourglass per structure, arrays of 10, 20, and 30 hourglass patterns per structure were fabricated. The measurement of a series of the resistors is equivalent to the statistical average of the measurement results of a large number of devices. Therefore, through a single measurement we are automatically averaging over a large number of devices and reducing measurement variations and errors. In order to further reduce the error and to ensure that the values for thermal conductivity are not dependent on the constriction width $w_{R,}$ three nominal values of 100, 200, and 300 nm were considered for the constriction width. The actual fabricated constriction width was later on measured to be 110, 215, 310 nm. Two different metal thicknesses 60 nm and 112 nm were studied in this experiment.

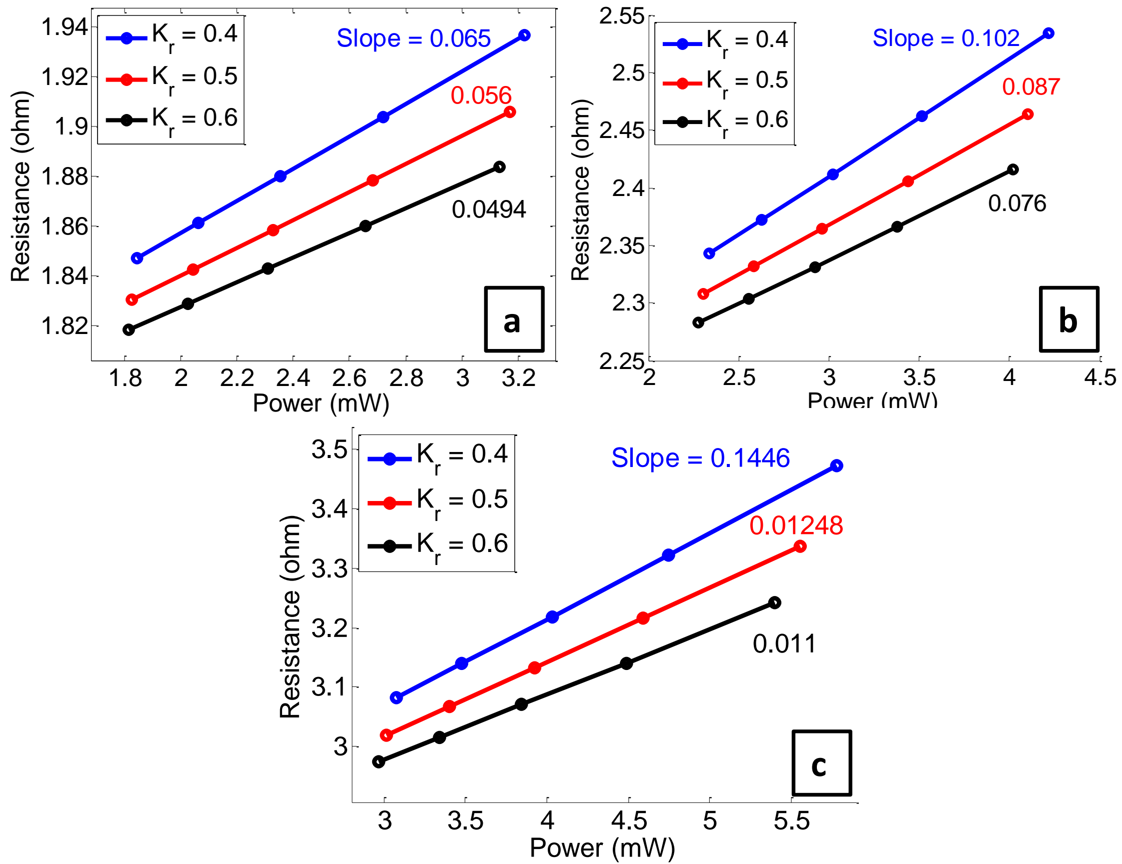

Figure 6-7 FE results of resistance vs. input power for different values of thermal conductivity for three combination of w_r , r_R, r_i. (a) w_R = 300 nm, r_R = 3 µm, r_i = 10 µm (b) w_R = 200 nm, r_R = 3 µm, r_i = 10 µm (c) w_R = 200 nm, r_R = 5 µm, r_i = 10 µm.

By collecting experimental data on various devices with different geometrical values, one can eliminate the error associated with any geometrical features other than the metal film thickness. In other words, by reducing the error in measurement, a more accurate study on the effect of metal thickness on the thermal conductivity can be conducted. As indicated in Table 6-1, a total of 10 devices were fabricated at a metal thickness of 112 nm and 9 devices at a metal thickness of 60 nm. Each device was given an ID based on its metallic film thickness. The specific geometrical parameters for each device including the number of hourglass patterns per structure are given in Table 6-1.

Table 6-1 Fabricated devices' IDs and parameters

Device ID	Thickness (nm)	w_R (nm)	r_R (nm)	r_i (nm)	number of patterns	device number
112-1	112	110	5	15	10	1
112-2						2
112-3					20	1
112-4						2
112-5		215	5	15	20	1
112-6					30	1
112-7						2
112-8		310	5	15	20	1
112-9						2
112-10					30	1
60-9	60	110	5	15	10	1
60-1					30	1
60-2						2
60-3		215	5	15	20	2
60-4					30	1
60-5						2
60-6		300	5	15	20	1
60-7					30	1
60-8						2

As mentioned earlier, one of the reasons in choosing the current structure is its straightforward nano-fabrication process relative to previously reported structures to measure thermal conductivity of metallic thin films. The entire fabrication process was carried in the cleanroom facilities at the Institute for Electronics and Nanotechnology (IEN) at Georgia Institute of Technology. The structures are fabricated on a 4-inch Si wafer of 525 μm thickness. A 4.5 μm thick SiO_2 layer is thermally grown using Lindberg

furnace on top of the wafer. Electron beam resist known as PMMA 6% was then spun on the sample at 2000 rpm and 1000 rpm/s for 60 seconds using EBL CEE Spinner. The sample was baked at 180°C for 90 seconds. The thickness of the resist was measured to be 650 nm. The resist was then patterned using electron beam lithography (JEOL 9300). Next, 5 nm of Titanium (Ti) and 112 nm of Cu were deposited by E-beam evaporation (Denton Explorer) over the patterned PMMA 6% layer. Ti is used as an adhesion layer for Cu. At the end, Ti and Cu layers were patterned using lift-off process. For the 60 nm thick devices, the same process was carried and only 60 nm of Cu were deposited using E-beam evaporation technique. Figure 6-8 shows optical images of fabricated structures with (a) 10, (b) 20, and (c) 30 unit patterns "hourglass" per structure. Figure 6-8 (d) exhibits a scanning electron microscopy (SEM) image of one unit pattern in the final structure. A magnified SEM image of the constriction within the structure is shown in Figure 6-8 (e).

6.6 Device Setup and Characterization

The fabricated devices were then diced and mounted on a 3 by 4 cm^2 Cu block using thermally conductive and electrically insulating epoxy. The devices were wire-bonded to a printed circuit board (PCB) that was placed on top of the Cu block as presented in Figure 6-9 (a). The PCB was then attached to a heatsink using thermal grease and wires are solders to the PCB traces for electrical connection, Figure 6-9 (b). By using a thermally conductive epoxy, Cu block, and a heat sink beneath the fabricated device, it is ensured that there is a high thermally conductive path at the bottom of the Si substrate to a heat sink with a high heat capacity. As a result, the bottom surface of the device (Si substrate) is kept at the room temperature during the experiment. This is very

useful to assure that temperature of the device does not rise during the experiment which facilitates the modeling of the device by enabling us to assume a constant temperature for the bottom surface of the Si with a very good approximation.

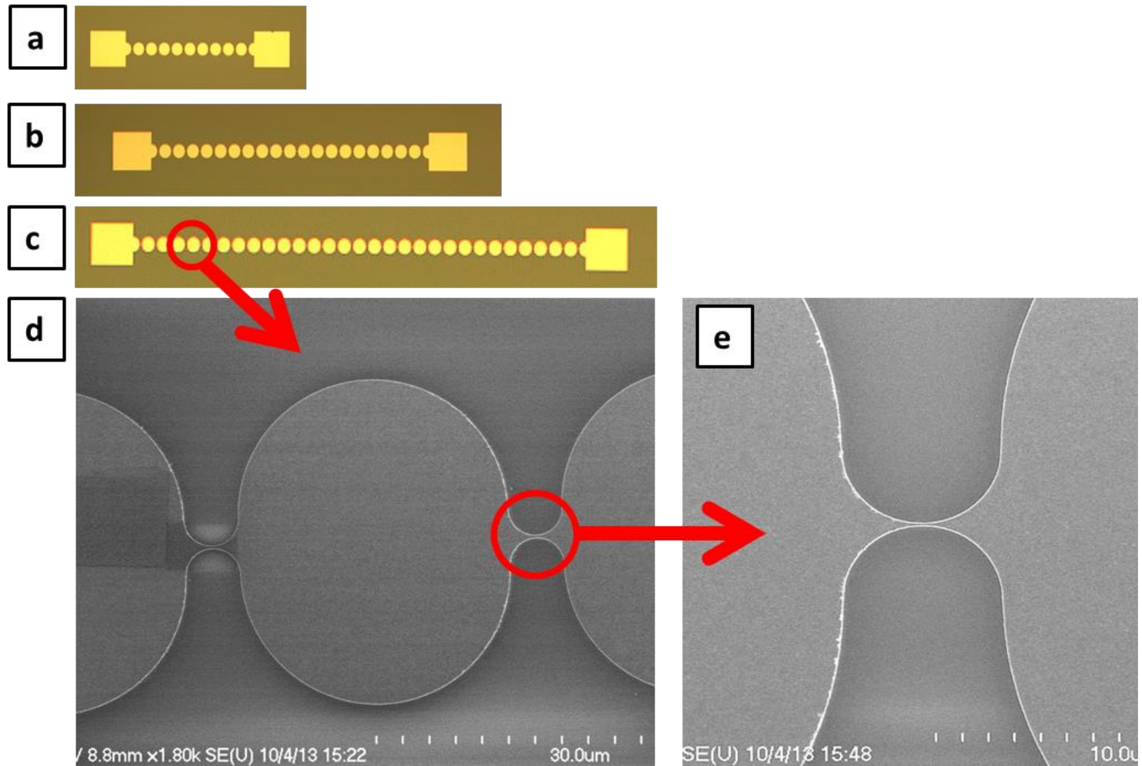

Figure 6-8 Optical images of fabricated structures with (a) 10, (b) 20, and (c) 30 unit pattern "hourglass" per structure. (d) SEM image of one unit pattern in the structure. (e) zoomed-in SEM image of the constriction within the structure.

Prior to testing the fabricated devices, each device was calibrated to determine the temperature dependent electrical resistance of the hourglass metal film. Devices placed in a temperature controlled oven isolated from the environment. A T-type thermocouple with ±0.1 ºC resolution was placed in the oven to further verify the readings of the oven's temperature controller. The oven temperature was set to six different temperatures between 22 ºC and 85 ºC and the resistance of the metal film was measured at each point. Thermal equilibrium was ensured by obtaining the measurements in an insulated, forced

convection heating oven. The change in the resistance with temperature is plotted in Figure 6-10. The diamond shaped markers represent the data from ten 112 nm devices and circular markers indicate the data from nine 60 nm devices. As expected, the calibration curves in Figure 6-10 indicate that the resistance of the hourglass structures has a linear dependence on temperature regardless of the geometrical value of the device other than their metal thickness. More specifically, the coefficient of determination, R^2, is greater than 0.998 for both 60 and 112 nm devices. As explained in Chapter 5, Eq. (5.2), based on the linear fit to the two sets of data in Figure 6-10, the temperature coefficient of the electrical resistance (TCR) can then be determined from:

$$\frac{R(T)-R(T_0)}{R(T_0)} = TCR \times \Delta T \quad (5.2)$$

Average TCR for 112 nm and 60 nm devices are 0.00241 K^{-1} and 0.00187 K^{-1} respectively. The bulk value of TCR for Cu is 0.003862 K^{-1}.

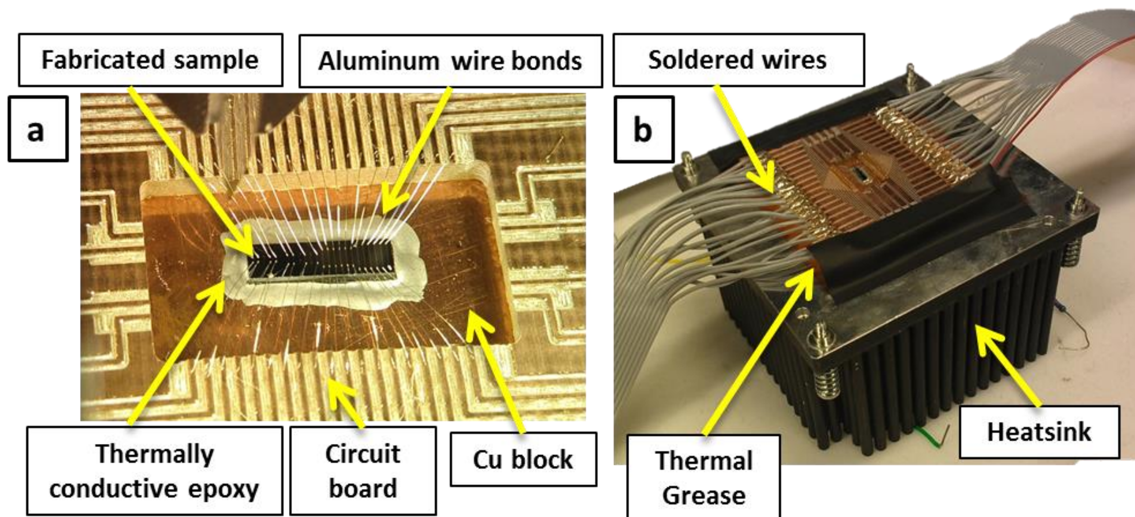

Figure 6-9 (a) Fabricated device mounted on Cu block with thermally conductive epoxy and attached to the circuit-board. (b) The wires are soldered to the board and the board is attached to a heat sink with thermal grease.

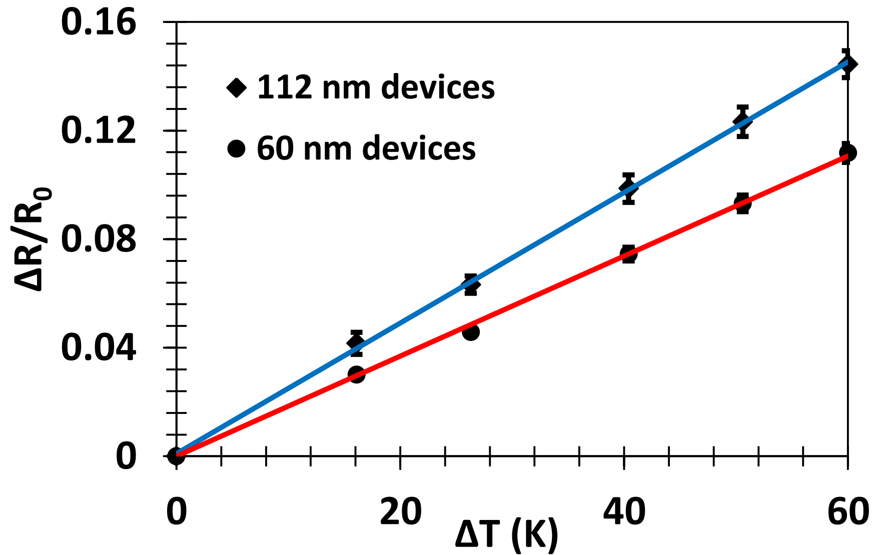

Figure 6-10 Calibration measurements for determining the temperature coefficient of the electrical resistance (TCR) for ten devices with 112 nm Cu film (diamond shaped markers) and eight devices with 60 nm Cu film (circular markers). Average TCR for 112 nm and 60 nm devices are 0.00241 K^{-1} and 0.00187 K^{-1} respectively. The bulk value of TCR for Cu is 0.003862 K^{-1}.

6.7 Experimental Setup and Procedure

Figure 6-11 shows a detailed flowchart of the steady state experimental procedure in determining the thermal conductivity of embedded metal thin films. There are three sets of experimental measurements that are required prior to the development of the numerical model, demonstrated in Figure 6-11.

6.7.1 Determining the initial electrical resistivity, $\rho^E(T_0)$

First step is to determine the geometrical values of each device accurately as the sensitivity of the model is highly dependent on them. w_R, r_R, and r_i were determined through precise SEM imaging with ± 2 nm uncertainty. The thickness of the underlying silicon dioxide layer was measured through refractometry technique with ± 5 nm uncertainty. Finally, the thickness of the metal film was determined by profilometry method with ±3nm uncertainty. Next, the initial electrical resistance (i.e., at room

temperature) of each device was measured by passing a very small current in picoampere (pA) range and measuring the voltage across each device. Based on the initial resistance and the geometrical values, the initial electrical resistivity (ρ_0^E) of each device was then calculated from Pouillet's law:

$$R_0 = \rho_0^E \frac{L}{A} \tag{6.1}$$

where L and A are length and cross sectional area of the resistance, respectively. The geometrical values and the initial electrical resistivity will then be integrated into the numerical model. The values for initial electrical resistance and initial electrical resistivity are provided in Table 6-3.

6.7.2 Determining the Temperature Coefficient of Resistance (TCR)

The next step is to determine the temperature coefficient of the electrical resistance (TCR). The procedure is explained in the previous section "Device Setup and Characterization". After obtaining the value for TCR, the temperature dependent electrical resistivity of the metallic film can be formulated as:

$$\rho^E(T) = \rho_0^E(1 + TCR \times \Delta T) \tag{6.2}$$

Eq. (6.2) will then be incorporated into the numerical model.

6.7.3 Determining the slope of R vs. P for each structure

The third step is to determine the dependence of electrical resistance of each device on the input power. Variation in the heating current will change the electric resistance of the structure which can be acquired via repetitive voltage measurements of the hourglass metallic film. Having experimentally determined the geometrical values, TCR, ρ_0^E, and temperature dependent electrical resistivity, the only remaining variable to

be chosen in the numerical model would be thermal conductivity. Therefore, by using a proper value for thermal conductivity in the numerical model, the experimentally attained slope for (R vs. P) can be matched to its numerical value. This will give the effective value of thermal conductivity for each tested device. As mentioned earlier, this procedure is demonstrated in Figure 6-11.

It is to be noted that both *TCR* and thermal conductivity are dependent on the size of the metallic film. Moreover, they both affect the dependence of resistance of the metallic film to the input power. However, their effects are opposite and partially cancel each other out. In other words, as the film thickness decreases, the sensitivity of (R vs. P) increases (because of the increase in κ) and simultaneously decreases (because of the decrease in *TCR*). Therefore, in order to determine the effect of thermal conductivity on film thickness, the *TCR* value is measured and incorporated in the numerical model. This way, the sensitivity of (R vs. P) with respect to *TCR* is accounted for and the only variable in the numerical results for (R vs. P) is thermal conductivity of the film.

6.8 Experimental Data

Figure 6-12 displays representative plots of resistance (Ω) vs. input power measurements (mW) for (a) 112 nm Cu film (device 112-5) and (b) 60 nm Cu film (device 60-4). The devices specifics are provided in Table 6-1. The solid red line is a linear fit to the experimental data. For short range of power input (shown in Figure 6-12), it can be assumed that there is a linear dependence of resistance over power. The input current was provided to the structure using a Keithly 2400 source meter. The same instrument was utilized for output voltage measurement. Measurement data were collected only after the steady state condition was researched. To reduce the effect of

noise in the measurement, each measurement point was repeated several times. The slope of the linear fit to the experimental data was then calculated for each device. The results are given in Table 6-3. For the two representative plots shown in Figure 6-12 the slopes were found to be 0.2146 (device 112-5) and 0.6875 (device 60-4) over the current range of 1-10 mA.

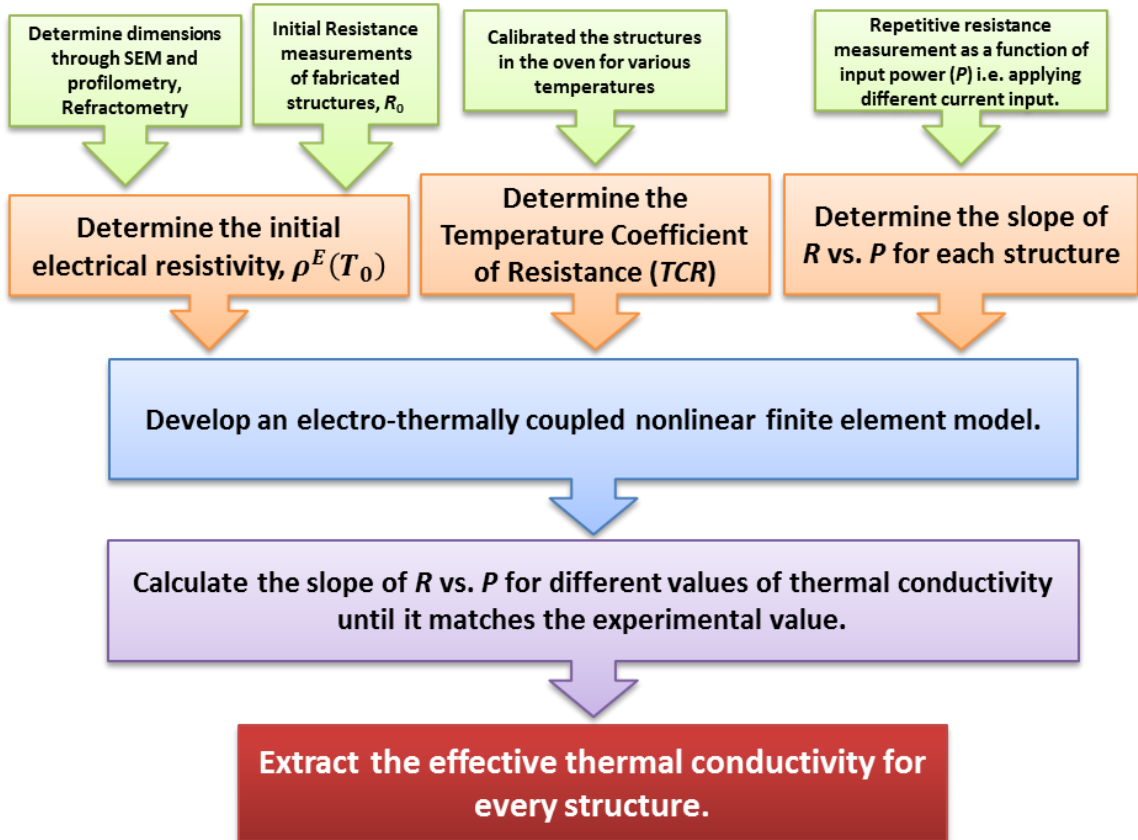

Figure 6-11 Flowchart of the steady state experimental procedure in determining the thermal conductivity of embedded metal thin films.

6.9 Electro-thermally Coupled Numerical Model

In order to accurately extract the thermal conductivity of the metallic films, the numerical model demonstrated in Figure 6-4 (a) should be modified such that it precisely represents the fabricated structure. Figure 6-13 (a) shows the layout of the three dimensional electro-thermally coupled nonlinear finite element (FE) model that was

adjusted to exemplify the unit pattern of the fabricated structure. Figure 6-13 (b) displays the zoomed-in section of the structure where the input current is applied. There are some differences between the originally proposed model in Figure 6-4 and the final model shown in Figure 6-13. In x-direction, the underlying silicon dioxide layer is further stretched to ensure that thermal insulation on the edges was realistic. The other difference is in the thermal boundary conditions in the y direction along the vertical walls. As shown in Figure 6-13 (c) the side walls (including the edge of the hourglass metal film and the substrate) are chosen to be thermally insulated (i.e., zero heat flux). This is the correct boundary condition because the dissipated heat in the high resistive section of each hourglass is equal for all of the resistors in series (equal electric current). Hence, there would be zero heat flux in the y direction at the symmetry plane between each two adjacent hourglass resistors (except for the first and last resistors). Similar to the previous model, the top and bottom boundaries were kept as natural convection and constant temperature respectively.

The FE model consists of 22457 elements with 1266 of them in the hourglass structure. The grid size was determined based on the mesh independency analysis in development of the numerical model. Table 6-2 provides the material properties used for numerical modeling. Since the thermal properties of deposited silicon dioxide layers are highly dependent on the deposition method and parameters, we chose to grow thermal silicon dioxide layers that possess highly repeatable and consistence physical (including thermal) properties. Here, these values are taken from Ref. [130].

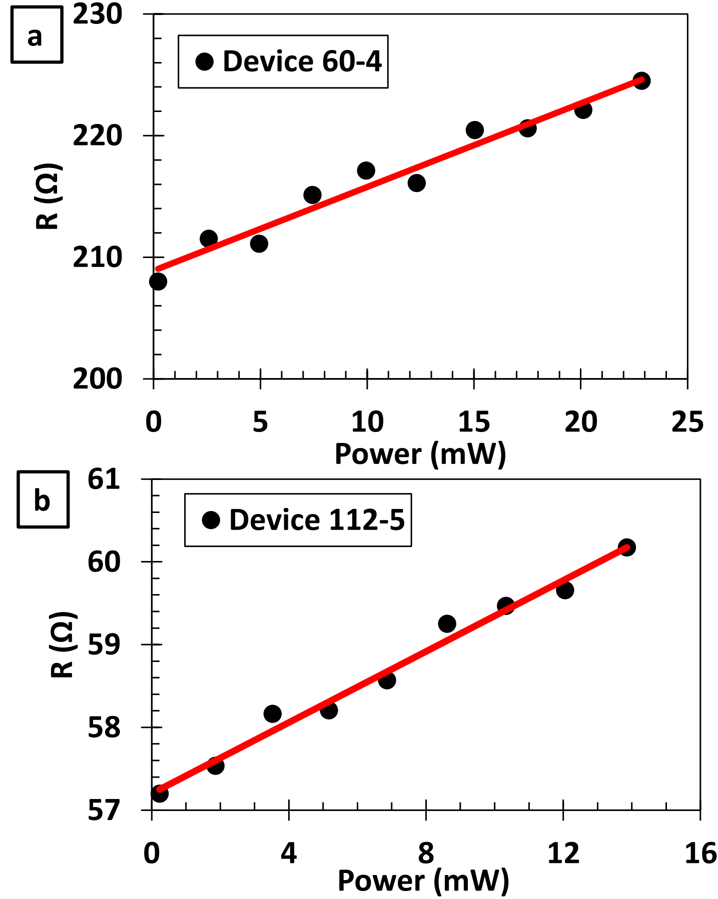

Figure 6-12 Representative plot of resistance vs. input power measurements for (a) 112 nm thick Cu (device 112-5) and (b) 60 nm thick Cu (device 60-4). The solid red line is a linear fit to the experimental data.

6.10 Results and Comparison

6.10.1 Numerical results

Figure 6-14 (a) exhibits a representative FE results for spatial distribution of current density in a unit pattern of the structure for Device 60-4 that was subjected to 10 mA of current once the state of equilibrium was reached. Thermal conductivity is extracted to be κ = 151.62 W/mK described in later sections. The top (x-y) view of current density in the hourglass is shown in Figure 6-14 (b). As a result of current continuity, the current density is much higher within and near the constricted area than in the wider part of the hourglass metallic film, displayed in Figure 6-14 (c).

Table 6-2 Modeling Parameters

	Thermal Conductivity (W/mK)	Density (Kg/m^3)	Specific Heat Capacity (J/KgK)	Initial Electrical Resistivity ($\Omega\cdot m$)
SiO_2	1.4	2200	730	10^5
Cu	401 (Bulk)	8960	384	1.72×10^{-8}
Si	163	2330	703	6.40×10^2

The associated steady state temperature amplitude in a unit pattern of Device 60-4 subjected to the current density, shown in Figure 6-14 (a), is demonstrated in Figure 6-15 (a). As it can be seen in Figure 6-15 (b), higher current density leads to a higher heat generation per unit volume. This non-uniform heat generation causes a large temperature gradient near the constriction as shown in Figure 6-15 (c). On account of relatively high thermal conductivity and low heat capacity of copper, any temperature gradient across the thickness of the constriction is relatively negligible.

6.10.2 Numerical fit to measurements of κ as a function of thickness

As previously stated in "Experimental Setup and Procedure" section, and shown in Figure 6-11, the final step is to find the appropriate value for thermal conductivity in the numerical model for which the slope of the measured (*R* vs. *P*) fitted line matches to its corresponding numerical value.

Figure 6-16 shows a representative plot of numerically generated normalized thermal conductivity versus the slope of (*R* vs. *P*) for 112 nm Cu layers at room temperature. Each blue marker denotes the calculated slope of (R vs. P) and its associated thermal conductivity value. Results presented in Figure 6-16 are for device 112-9. The experimentally measured slope value was determined to be 0.1289. The corresponding

thermal conductivity was found to be 238.19 W/mK compared to the bulk value of 401 W/mK. An important observation of Figure 6-16 is that a polynomial regression model (solid red line) fits very well to data points with the norm of residuals to be 0.01192. This statement holds true for all of the devices that were modeled and tested regardless of their geometrical features or metal film thickness. Therefore, by determining only a few points for thermal conductivity vs. slope, the model for the entire range of slopes can be plotted. This will allow for a more precise extraction of thermal conductivity with much fewer numerical simulations which greatly reduces the computational cost.

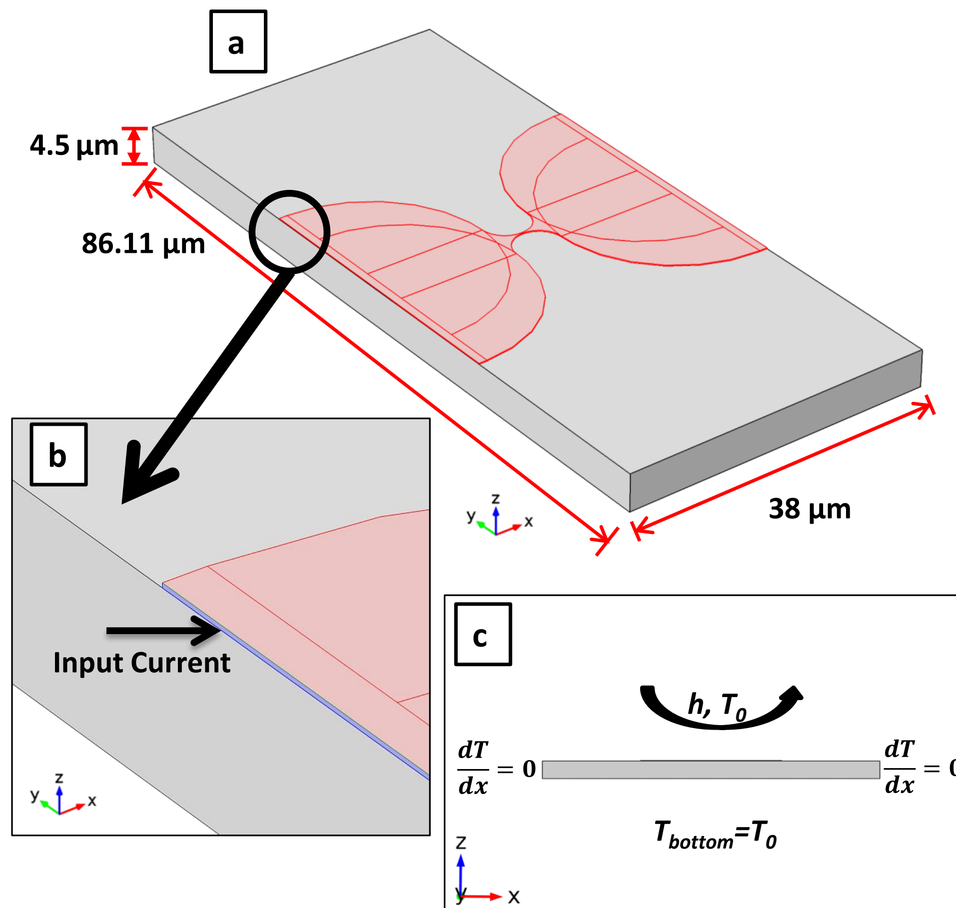

Figure 6-13 (a) Layout of the FE model representing the unit pattern of the fabricated structure. (b) Zoomed-in section of the structure where the input current is applied. (c) Cross sectional view of the structure. The thermal boundary condition on top and bottom are natural convection and constant temperature respectively. In longitudinal direction (x) the walls are considered to be thermally insulated due to symmetry.

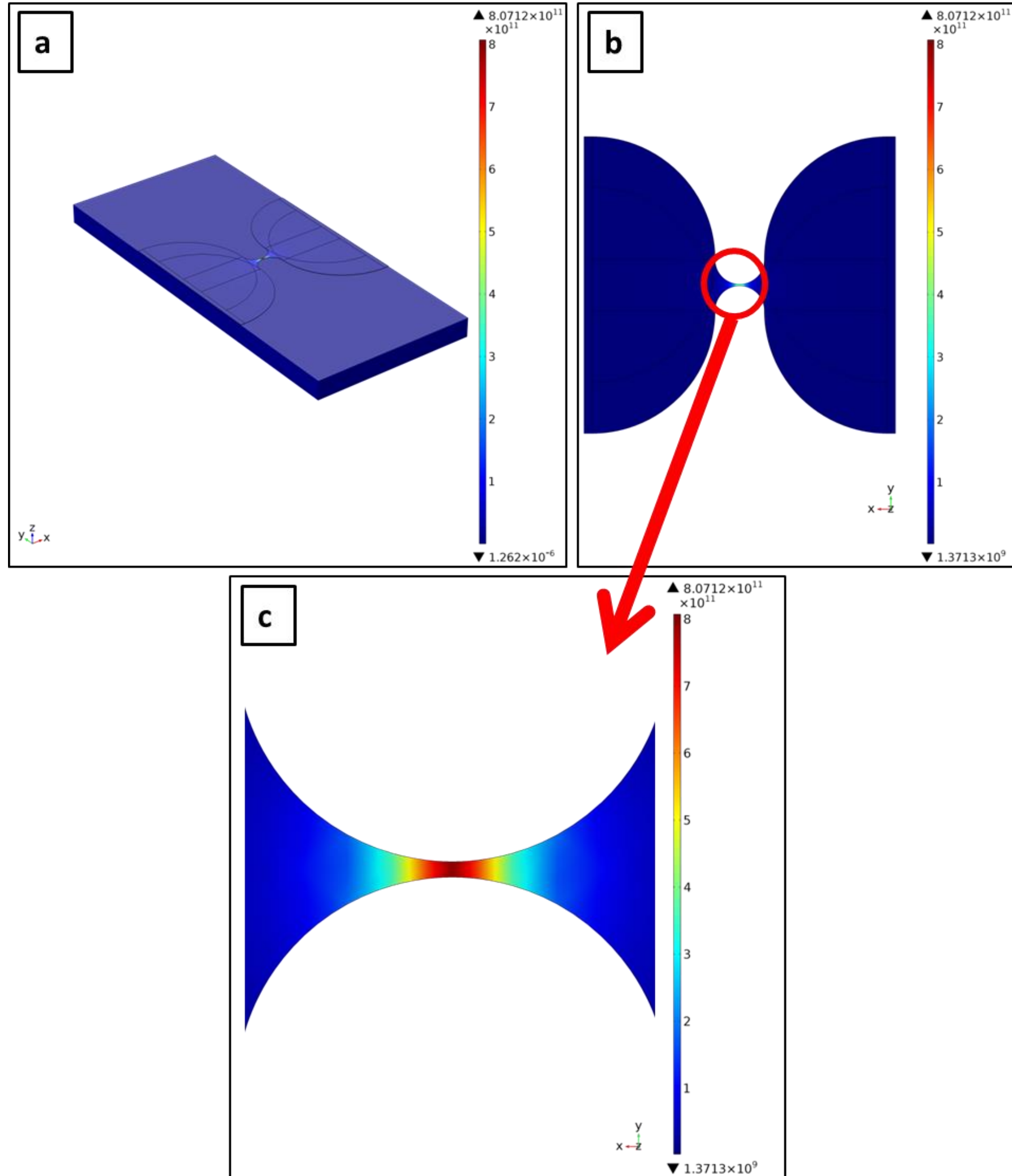

Figure 6-14 (a) spatial distribution of current density in the unit pattern of the structure for Device 60-4 when subjected to 10 mA of current. (b) Top view of current density in the "hourglass" part of the structure. (c) Zoomed-in view of current density distribution in the constricted area within "hourglass".

Similarly, Figure 6-17 displays a representative plot of numerically generated normalized thermal conductivity versus the slope of (R vs. P) for 60 nm Cu layers.

Results are for device 60-4. The experimentally measured slope value is 0.688 and the associated thermal conductivity is determined to be 151.62 W/mK. The solid red line is a second order polynomial regression model fitted to the numerical data with the norm of residuals of 0.00489.

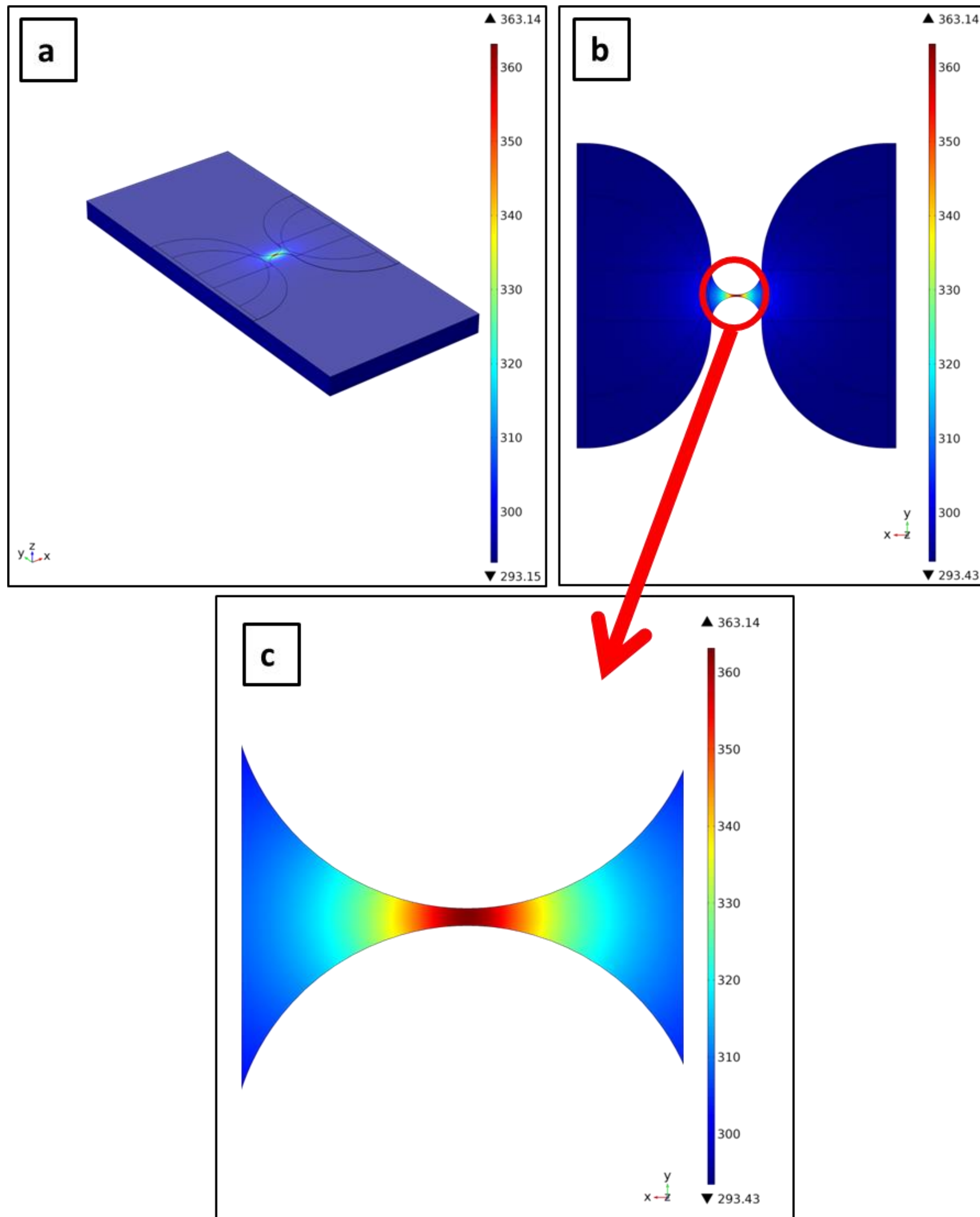

Figure 6-15 (a) Steady state temperature distribution in the unit pattern of the structure for Device 60-4 when subjected to 10 mA of current. (b) Top view of temperature distribution in the "hourglass" part of the structure. (c) Zoomed-in view of temperature distribution in the constricted area within "hourglass". Thermal Conductivity for Device 60-4 is determined to be 151.62 W/mK.

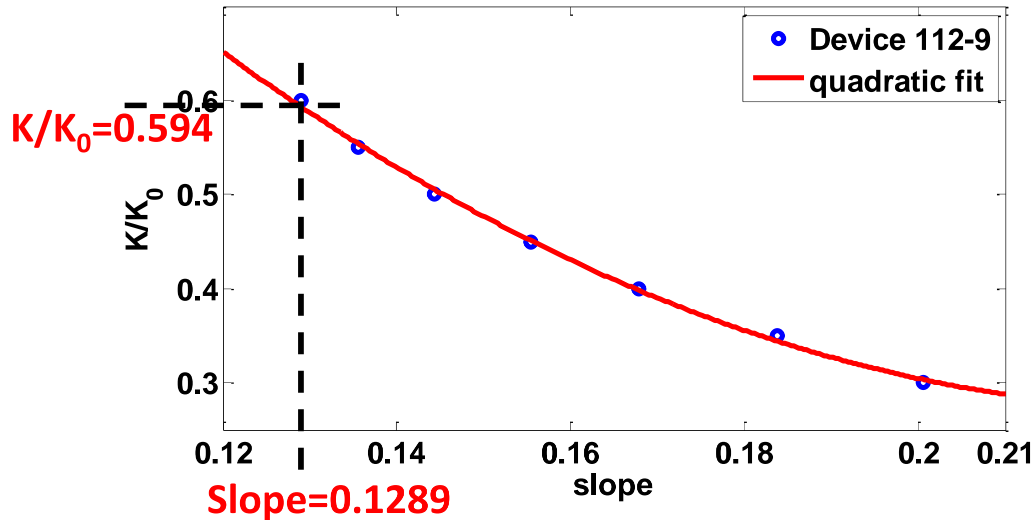

Figure 6-16 Representative plot of numerically generated normalized thermal conductivity versus the slope of (*R* vs. *P*) for 112 nm Cu layers. Results are for device 112-9. The experimentally measured slope value is 0.1289. The corresponding thermal conductivity is determined to be 238.19 W/mK. The solid red line is a quadratic fit to the numerical data.

6.10.3 Effect of constriction width, w_R on the model's sensitivity

The effect of constriction width, w_R, on the slope of (*R* vs. *P*) was also studied. Figure 6-18 shows the normalized thermal conductivity of 112 nm Cu layer versus the slope of (*R* vs. *P*) for different w_R. The representative results are for devices 112-4 (w_R=100 nm), 112-5 (w_R=200 nm), and 112-10 (w_R=300 nm). The solid lines are quadratic fits to the numerical data. It can be seen that the lower the constriction width, w_R, the sharper the decay in normalized thermal conductivity vs. slope. This implies that the devices with thinner constrictions are more sensitive to the change in thermal conductivity. Figure 6-19 presents comparable results for 60 nm instead of 112 nm thick Cu films. Representative results are for devices 60-2 (w_R=100 nm), 60-4 (w_R=200 nm), and 60-8 (w_R=300 nm). As one would expect, by comparing Figure 6-19 with 6-18, it can be concluded that the 60 nm thick Cu data are generally more sensitive to the change of thermal conductivity which results in a better experimental resolution.

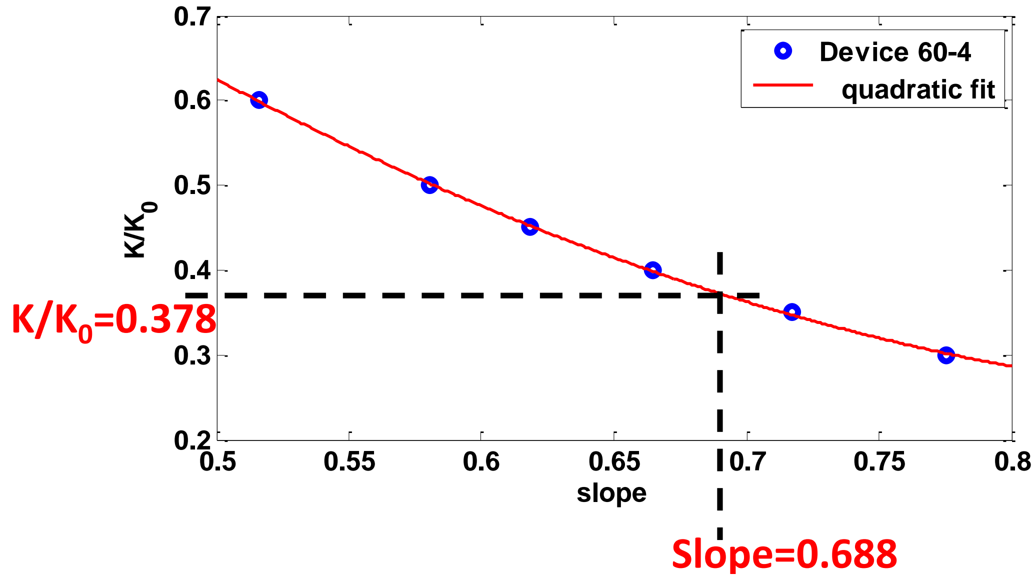

Figure 6-17 Representative plot of numerically generated normalized thermal conductivity versus the slope of (*R* vs. *P*) for 60 nm Cu layers. Results are for device 60-4. The experimentally measured slope value is 0.688. The corresponding thermal conductivity is determined to be 151.62 W/mK. The solid red line is a quadratic fit to the numerical data.

6.10.4 Thermal conductivity as a function of film thickness

Figure 6-20 shows the measured thermal conductivities of copper films as a function of thickness at room temperature. The blue diamond shaped markers indicate the average value for thermal conductivity of devices with 112 nm and 60 nm Cu layers. There are a total of ten devices with 112 nm and eight devices with 60 nm Cu films. The average value for thermal conductivity of for 112nm and 60 nm embedded Cu films were determined to be 216.14 and 160.50 W/mK respectively. The error bars represent maximum of 11% error associated with the measurements. As it can be seen in Figure 6-20, thermal conductivity decreases with a decrease in the metallic film thickness. The presented experimental data agrees very well with the previous studies by Nath and Chopra [136], Klemens [137], and Liu et al. [57]. This both confirms the validity of the proposed approach and the fact that there is not a significant difference between the

thermal conductivity of suspended copper films [57] and those on silicon dioxide layers as shown in this work.

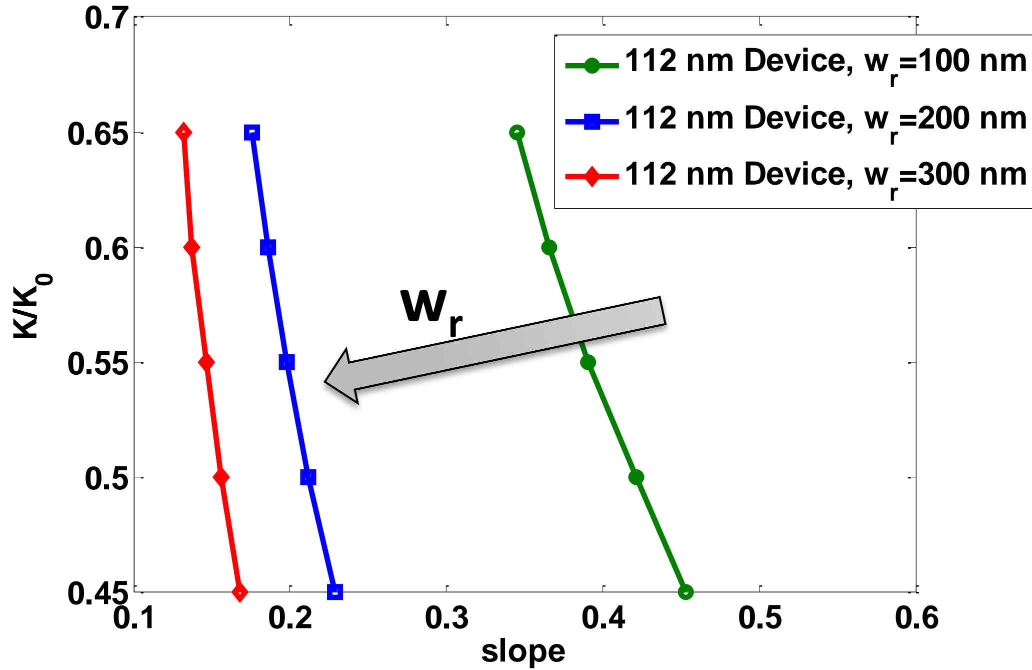

Figure 6-18 Normalized thermal conductivity of 112 nm thick Cu layer versus the slope of (*R* vs. *P*) for different w_R. Representative results are for devices 112-4 (w_R=100 nm), 112-5 (w_R=200 nm), and 112-10 (w_R=300 nm) . The solid lines are quadratic fits to the numerical data.

A summary of the measurements for every device is given in Table 6-3. The results include initial resistance (Ω), slope of (*R* vs. *P*), the initial electrical resistivity, ρ_0^E (Ωm), the range of input current (mW), and finally thermal conductivity (W/mK). It is important to recognize that the measurements of slope of (*R* vs. *P*) are more sensitive to the change of thermal conductivity for lower values of w_R, which result in the lower experimental resolution in determining the thermal conductivity for lower values of w_R. However, as one would expect, the final extracted values of thermal conductivity are not dependent on w_R. This can also be observed from the values of thermal conductivity listed in Table 6-1 for the whole range of w_R.

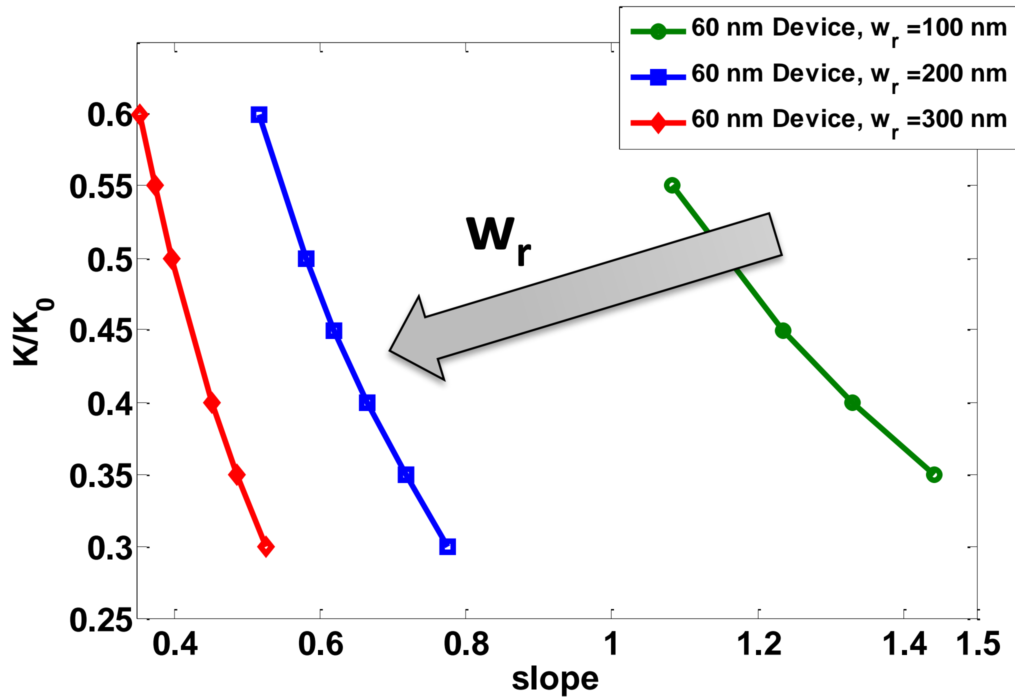

Figure 6-19 Normalized thermal conductivity of 60 nm thick Cu films versus the slope of (*R* vs. *P*) for different w_R. Representative results are for devices 60-2 (w_R=100 nm), 60-4 (w_R=200 nm), and 60-8 (w_R=300 nm). The solid lines are quadratic fits to the numerical data.

It is important to note that the measured values of thermal conductivity are consistent for each metal thickness regardless of their other dimensions, i.e., w_R and number of unit patterns per device which further verifies the applicability of proposed method. In the experimental and theoretical studies so far in this chapter, the dependence of the thermal conductivity on temperature was neglected. However, in order to minimize this effect, these studies were performed at low current densities so that the temperature rise inside the device is not significant (< 30K). In the next section, the effect of temperature dependence of thermal conductivity is included in the experimental and theoretical studies.

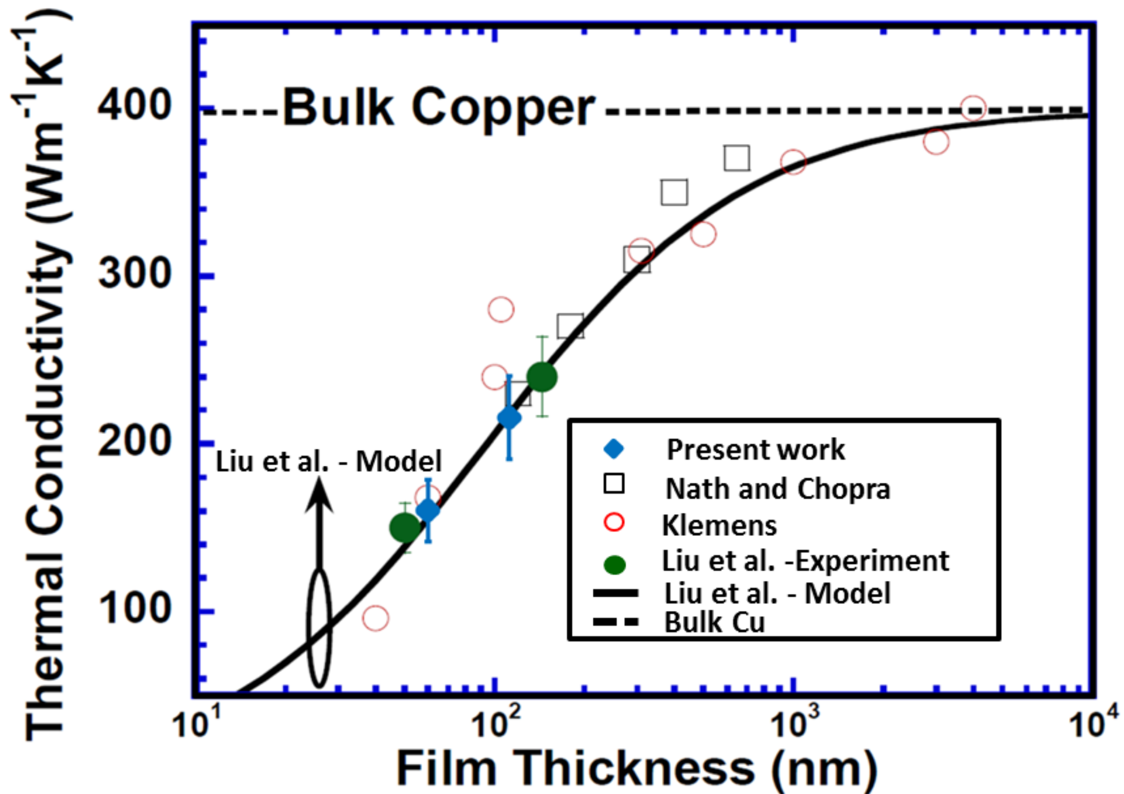

Figure 6-20 Experimental data for the thermal conductivity of 112 and 60 nm Cu layers. There are ten and eight data points for 112 and 60 nm Cu layers respectively. Each data point represents one device.

6.10.5 Study the dependence of thermal conductivity on temperature

In this study, it is assumed that the thermal conductivity of the embedded Cu film has a linear dependency on temperature within the range of 300K to 425 K. This assumption is based upon the analytical correlations and experimental values reported by Nath and Chopra [136] and Liu et al [57] for the thermal conductivity of free-standing copper bridges in the of 300-K450K. Since the error associated with the first order approximation falls well within the 11% uncertainty in the present measurements and also owing to the fact that there are no reported correlations for the temperature dependency of the thermal conductivity of embedded metal films, the linearity

presumption can be justified within the scope of this study. Hence, thermal conductivity can be formulated as

$$\kappa(T) = \kappa(T_0) + \beta_k \times (T - T_0) \quad (300\,K < T < 425\,K) \tag{6.3}$$

where $k(T_0)$ is the measured thermal conductivity of the Cu films at room temperature shown in Figure 6-20 and given in Table 6-3.

In order to determine β_k, Eq. (6.3) together with the measured thermal conductivity of each device at room temperature ($\kappa(T_0)$) are incorporated in to the electro-thermally coupled numerical model. The device is then modeled for a short range of input currents at much higher amplitude to heat the device in the desired temperature range (~ 450 K). Similar to the procedure explained in section "Effect of constriction width, w_R on the model's sensitivity", the numerically attained slope for (R vs. P) can be matched to its experimental value by adjusting the value of β_k in Eq. (6.3). A new set of experiments were performed on all devices at higher current amplitudes to enable the measurement of the thermal conductivity at high temperatures. Table 6-4 provides the current range used for the new sets of experiments and numerical simulations in the process of finding β_k.

6.10.6 Numerical fit to measurements of κ as a function of temperature

Figure 6-21 displays a typical plot of numerically generated for β_k versus the slope of (R vs. P) for 112 nm Cu layers at high power inputs (Current range = 22-27 mA). Each blue marker signifies the calculated slope of (R vs. P) and its correlated β_k value. Results presented in Figure 6-21 are for device 112-7. The experimentally measured slope value was determined to be 0.1945. The associated β_k was found to be 0.103 W/mK^2. The initial thermal conductivity used in the model was previously determined as 233.38 W/mK.

Table 6-3 Experiment specification and results for fabricated devices including thermal conductivities.

Device ID	Thickness (nm)	w_R (nm)	number of patterns	device number	Initial Resistance (Ω)	slope	Initial Electrical Resistivity, ρ_0^E (Ωm)	Current Range (mA)	Thermal Conductivity (W/mK)
112-1	112	110	10	1	24.12	0.2254	$1.72\times10^{-8}\times0.93$	1-10	237.79
112-2				2	24.16	0.222	$1.72\times10^{-8}\times0.93$	1-10	241.00
112-3			20	1	79.43	0.4172	$1.72\times10^{-8}\times1.53$	1-10.7	202.51
112-4				2	119.72	0.422	$1.72\times10^{-8}\times1.53$	1-10.7	199.70
112-5		215	20	1	57.2	0.2146	$1.72\times10^{-8}\times1.53$	1-10	196.69
112-6			30	1	93.07	0.2292	$1.72\times10^{-8}\times1.66$	2-15	198.09
112-7				2	85.65	2.04	$1.72\times10^{-8}\times1.53$	2-15	233.38
112-8		310	20	1	48.48	1.1308	$1.72\times10^{-8}\times1.549$	10-20	233.38
112-9				2	47.85	1.289	$1.72\times10^{-8}\times1.53$	3-13.3	238.19
112-10			30	1	77.99	0.1615	$1.72\times10^{-8}\times1.66$	3-10	191.28
60-9	60	110	10	1	140.44	2.1637	$1.72\times10^{-8}\times3.07$	0.5-7	96.64
60-1			30	1	343	1.7785	$1.72\times10^{-8}\times2.37$	0.5-7	83.40
60-2				2	259.34	1.2411	$1.72\times10^{-8}\times1.8$	0.5-10	179.05
60-3		215	20	2	130.3	0.6694	$1.72\times10^{-8}\times1.85$	1-10	149.97
60-4			30	1	210.6	0.6875	$1.72\times10^{-8}\times2.02$	1-10	151.62
60-5				2	193.66	0.6798	$1.72\times10^{-8}\times1.85$	1.5-10	146.57
60-7		310	30	1	166.87	0.4806	$1.72\times10^{-8}\times1.907$	1.5-10	141.95
60-8				2	168.23	0.481	$1.72\times10^{-8}\times1.92$	1.5-10	143.16

There are a couple of observations regarding Figure 6-21 that should be noted. First, it can be concluded from the present data that for every 1% increase in the slope of (R vs. P), there is approximately 17% reduction in the value of β_k. This implied that the model has a low sensitivity in the determination of β_k which can potentially introduce larger error. Another important observation of Figure 6-21 is that a polynomial regression model (solid red line) fits very well to data points with the norm of residuals to be

0.00029. This is true for all of the devices that were modeled and tested regardless of their geometrical features or metal film thickness. Thus, by determining only a few data points β_k vs. slope, the model for the entire range of slopes can be determined. This will lead to a more accurate extraction of β_k at a much lower computational cost.

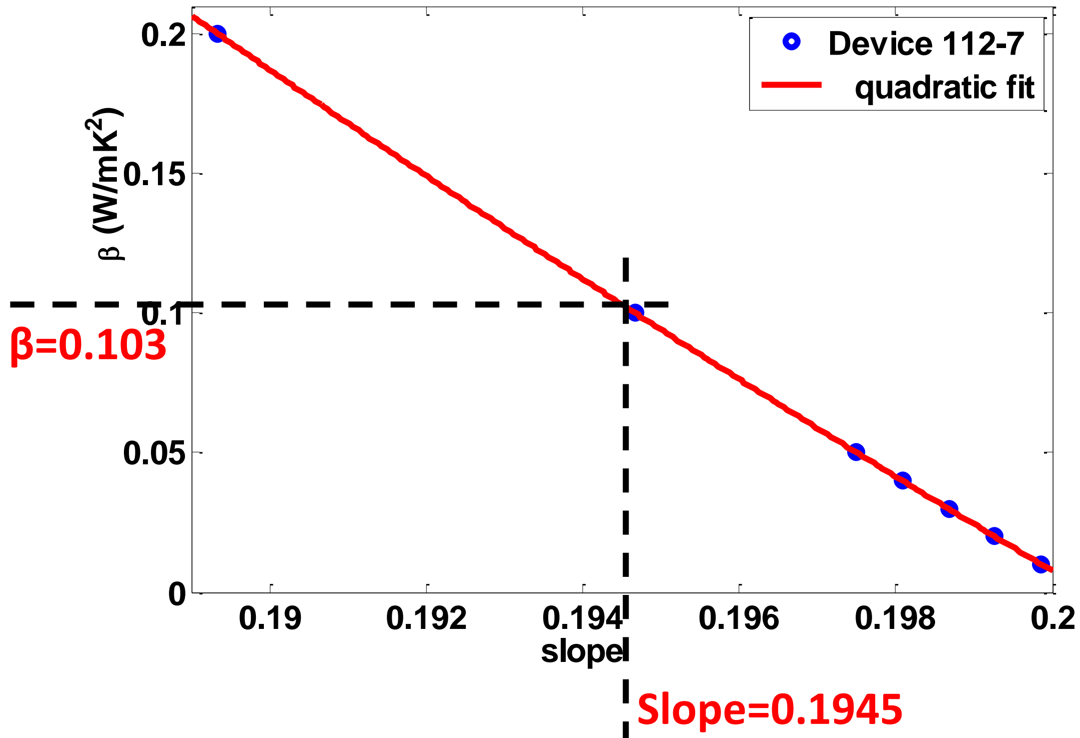

Figure 6-21 Representative plot of numerically generated values for β_k versus the slope of (R vs. P) for 112 nm Cu layers at high power inputs. Results are for device 112-7. The experimentally measured slope value is 0.1945. The corresponding β_k is determined to be 0.103 W/mK^2. The initial thermal conductivity used in the model is 233.38 W/mK. The solid red line is a quadratic fit to the numerical data.

Likewise, Figure 6-22 exhibitions a representative plot of numerically generated normalized β_k versus the slope of (R vs. P) for 60 nm Cu layers (device 60-5) at high power inputs (Current range = 11-15 mA). The experimentally measured slope value is 0.6396 and the associated β_k is determined to be 0.31 W/mK^2. The initial thermal conductivity used in the model was previously determined as 146.57 W/mK. Same remarks as in Figure 6-21 can be stated for Figure 6-22. The model has a low sensitivity

to β_k (6% reduction in the value of β_k for every 1% increase on the slope) which can potentially introduce greater error. As shown with a solid red line, a second order polynomial regression model is fitted to the numerical data with the norm of residuals of 0.0033.

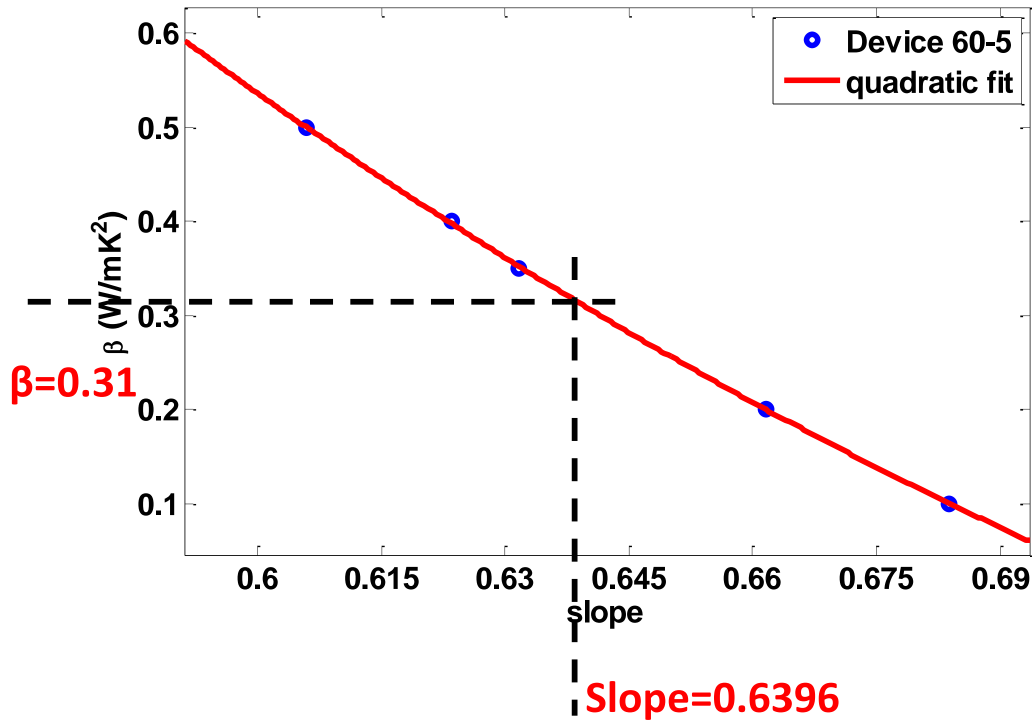

Figure 6-22 Representative plot of numerically generated values for β_k versus the slope of (R vs. P) for 60 nm Cu layers at high power inputs. Results are for device 60-5. The experimentally measured slope value is 0.6396. The corresponding β_k is determined to be 0.31 W/mK^2. The initial thermal conductivity used in the model is 146.57 W/mK. The solid red line is a quadratic fit to the numerical data.

6.10.7 Thermal conductivity as a function of temperature

Thermal conductivity of embedded 112 nm and 60 nm Cu layers as a function of temperature is shown in Figure 6-23. The data are compared to the bulk value of copper thermal conductivity [138] and the theoretical and experimental data for 144nm and 50nm free standing bridges reported by Liu et al. [57]. The blue circular markers indicate the average value for thermal conductivity of devices with 112 nm Cu layers and the

brown square shaped markers represent the 60 nm devices. There are a total of five devices with 112 nm and four devices with 60 nm Cu films. The average value β_k, i.e. the slope of the present data, for 112nm and 60 nm layers were determined to be 0.223 and 0.342 W/mK respectively. Notably, there is a maximum of 10.3% uncertainty for 112 nm films and only 3.5% uncertainty for 60 nm films in the experimental data.

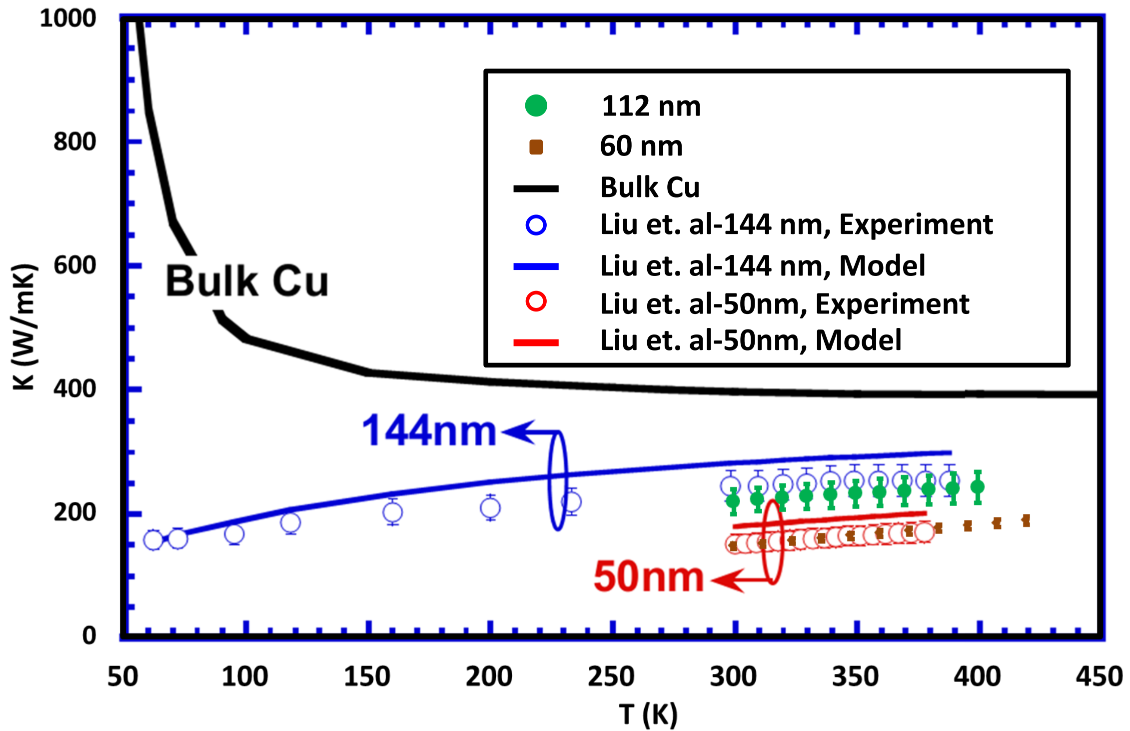

Figure 6-23 Experimental data for the temperature dependent thermal conductivity of 112 and 60 nm Cu films. There are five and four data points for 112 and 60 nm Cu layers respectively. Each data point represents one device.

It is important to note that the average value of β_k is higher for the thinner film. This means that the thermal conductivity decrease with a sharper slope for thinner films. For thicker Cu films (>180 nm), as the temperature reduces, the increase in thermal conductivity is more noticeable and resembles the behavior of bulk Cu. This is verified for 180 nm- 640 nm Cu films by Nath and Chopra [136]. However, the increase in thermal conductivity with reducing temperature becomes less and less evident as the film thickness decreases. Eventually, the temperature dependence of thermal conductivity is

reversed beyond a certain thickness. This behavior is anticipated due to the fact that when the film thickness decreases, the role of size effect is more dominant at lower temperatures (mean free path of Cu at 500 K is 23 nm and at 100 K is 145 nm) [136]. In conclusion, the presented data (shown in Figure 6-23) agree well with the previous studies which further verifies the validity of the proposed method. Table 6-4 provides a complete list of β_k values, their associated slope of (*R vs. P),* current ranges used to determine β_k, and the thermal conductivity values at room temperature for all of the tested devices.

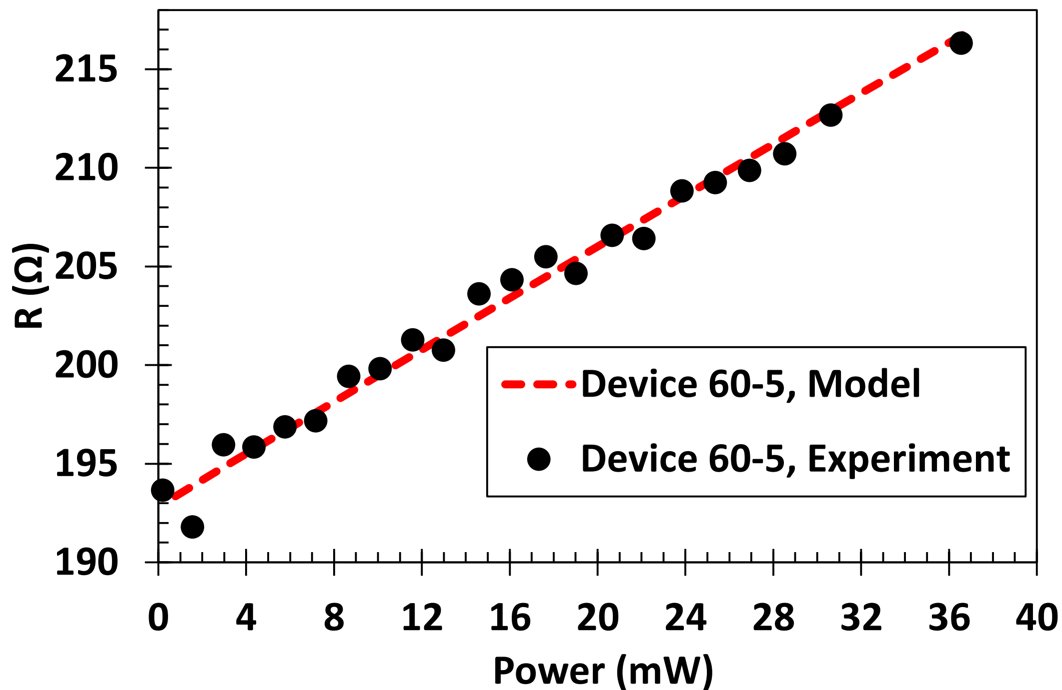

Figure 6-24 Representative plot of experimental data (black markers) and numerical results (dashed red line) for resistance vs. input power. The model is developed using temperature dependent conductivity for the full range of power spectrum. Results are for device 60-5. The initial value of thermal conductivity is κ_0=146.57 W/mK and β_k=0.31 W/mK^2.

In Figure 6-23, the temperature dependent thermal conductivity of thin films was verified against the previous studies. In order to confirm the validity of β_k against the

experimental results for each device, the numerical model was utilized for the full range of input current in each structure. Eq. (6.3) and the determined β_k value were also incorporated into the model for the entire current range and not just the higher current values. The numerical results were then compared against the experimental data for the full range of power to ensure that the chosen value for β_k was indeed accurate and can capture the behavior of the film for the full range of power. Figure 6-24 shows a representative plot of experimental data (black markers) and numerical results (dashed red line with solid markers) for resistance vs. input power. Results are for device 60-5 with the initial value of thermal conductivity being κ_0=146.57 W/mK and β_k =0.31 W/mK^2. It can be seen that the numerical model agrees well with the experimental data. In order to quantify this agreement, the root-mean-square error (*RMSE*) between values predicted by the model and the experimental data was evaluated as:

$$RMSE = \sqrt{\frac{\sum_i^{number}(y_i^{exp} - y_i^{mod})^2}{number}} \tag{6.4}$$

where y_i^{exp} are the experimental and y_i^{mod} the numerical data points. The term $number$ represents the number of the experimental data points that were taken. The *RMSE* value is calculated to be 0.8 over 23 data points which further demonstrates the applicability of the proposed model.

6.11 Conclusion

As discussed earlier, it is highly challenging to measure the thermal conductivity of an embedded metallic film due to a much higher thermal resistance of the low-K insulation layer and the substrate than that of the thin metallic film. The overall through-plane thermal resistance of an embedded structure is very high with respect to the thermal

resistance of the metallic layer. Therefore, a slight deviation in the thermal resistance of the insulation layer can cause a significant error in thermal conductivity measurements. Therefore, in this chapter a new concept was proposed to induce strong sensitivity of the heating structure to the thermal conductivity of the metallic layer. The novel approach proposed here uses a laterally varying resistor structure to produce lateral heat gradient and to induce lateral heat diffusion in the plane of the metallic layer. Ultimately, through steady-state Joule heating and electrical resistance thermometry the thermal conductivity of the films can be identified.

Table 6-4 Experimental specification for determining β_k including the initial thermal conductivities

<table>
<tr><th>Structures ID</th><th>Thickness (nm)</th><th>w_R (nm)</th><th>Number of Patterns</th><th>Device Number</th><th>High Power Slope</th><th>Initial Thermal Conductivity (W/mK)</th><th>Current Range (mA)</th><th>β_k (W/mK^2)</th></tr>
<tr><td>112-3</td><td rowspan="5">112</td><td rowspan="2">110</td><td rowspan="2">20</td><td>1</td><td>0.4282</td><td>202.51</td><td>13-16</td><td>0.141</td></tr>
<tr><td>112-4</td><td>2</td><td>0.4023</td><td>199.70</td><td>13-16</td><td>0.45</td></tr>
<tr><td>112-6</td><td rowspan="2">215</td><td rowspan="2">30</td><td>1</td><td>0.221</td><td>198.09</td><td>22-28</td><td>0.43</td></tr>
<tr><td>112-7</td><td>2</td><td>0.1945</td><td>233.38</td><td>22-30</td><td>0.103</td></tr>
<tr><td>112-9</td><td>310</td><td>20</td><td>2</td><td>0.1131</td><td>238.19</td><td>17-20</td><td>0.28</td></tr>
<tr><td>60-3</td><td rowspan="4">60</td><td rowspan="2">215</td><td>20</td><td>2</td><td>0.6293</td><td>149.97</td><td>11.3-15</td><td>0.364</td></tr>
<tr><td>60-5</td><td>30</td><td>2</td><td>0.6396</td><td>146.57</td><td>11-15</td><td>0.31</td></tr>
<tr><td>60-7</td><td rowspan="2">310</td><td rowspan="2">30</td><td>1</td><td>0.4093</td><td>141.95</td><td>15-17.5</td><td>0.326</td></tr>
<tr><td>60-8</td><td>2</td><td>0.4501</td><td>143.16</td><td>15.5-17</td><td>0.31</td></tr>
</table>

Using the proposed method, the size effect on thermal conductivities of 112 nm and 60 nm embedded Cu films were investigated. The average values of thermal conductivity at room temperature were found to be 216.14 W/mK and 160.50 W/mK for the 112 nm and 60 nm films respectively. This reduction in thermal conductivity is associated with both surface scattering and grain boundary scattering. The experimentally

measured values for thermal conductivity agree well with previous studies on free-standing Cu bridges. The dependence of the thermal conductivity on temperature was also investigated and the results are well within the scope of previously reported value. It is important to note that the measured values for thermal conductivity are not dependent on any of the structure dimensions such as, $w_{R,}$ $r_{R,}$ and number of unit patterns per device which further verifies the applicability of the proposed method.

The proposed technique has several advantages compared to the existing methods. By keeping the underlying substrate, the interface between metal and dielectric will stay unaffected. In addition, any effect associated with the interface is accounted for. Another advantage of the proposed scheme is that it does not require extensive micro-fabrication. It can also be integrated within the structure and be used for embedded or buried structures such as nanoscale on chip interconnects. Although the method is proposed to measure metallic films, it can also be used for any material that can carry current and generate heat such as semiconductors.

CHAPTER 7: SUMMARY AND CONCLUSION

In this work, novel numerical and experimental methods were developed to address transient Joule heating in nanoscale embedded on-chip interconnects. This chapter summarizes the key findings of this work and lists the original contributions of the dissertation. At the end, recommendations are made for the future work to extend the scope of this research.

7.1 Conclusion

Initially, a two dimensional (2-D) reduced order modeling approach based on Proper Orthogonal Decomposition (POD) implementing the Galerkin projection technique was developed to address the transient Joule heating in an inhomogeneous arrangement of interconnects embedded in a dielectric material. To improve the capabilities of the developed POD model and to further reduce the computational cost, a multi-scale reduced order transient thermal methodology called *hybrid scheme* was developed which incorporates 3-D POD technique into another multi-scale modeling approach called *"Progressive Zoom-in"* method. This model was originally developed for low power portable systems, where heat sinks and forced cooling are not employed owing to the compact form factor. Consequently, the hybrid scheme was further advanced to address the transient thermal problem in a packaged high power microprocessor where, in fact, force convection plays a key role in the thermal transport of the structure. Another extension to the model was the implementation of a realistic highly spatially resolved transient power map. The original contributions of this study in the modeling framework are summarized below:

- A computationally efficient and accurate multi-scale transient thermal model, hybrid scheme, was developed with the ability of modeling several decades of length and time scale at orders of magnitude lower computational cost while maintaining satisfactory accuracy.
- Utilizing the proposed model, the computational time is reduced by at least two orders of magnitude at every step of zooming into the geometry.
- The hybrid scheme is not limited to the two levels considered in the present study and can be scaled to multiple levels and be used to simulate more detailed structures on the chip while taking advantage of the capabilities of POD method to avoid any further full field simulation. In essence, without losing the desired resolution, the hybrid scheme proposes a new approach to further decrease the computational cost by orders of magnitude.
- Another distinct benefit of the proposed method is that, for any linear system, the POD solution is independent of the transient power profile (analytically proved in Appendix A). In other words, once the solution to a sample power input is obtained, there is no need to generate new observations or full field FE simulations. This important feature can drastically decrease computational cost, making POD a fast and robust method for reduced order model of transient heat conduction in microelectronic devices.
- One of the most remarkable characteristic of the POD is its optimality. Data sets are expanded for modal decomposition on empirically

determined basis functions in a way that minimizes the least square error between the true solution and the truncated representation of the POD model. This makes the POD method one of the most efficient method of capturing the dominant components of a large-dimensional system with a finite number of modes.

- An additional unique characteristic of this model is that the initial observations can be obtained experimentally which creates the ability of modeling a potentially complex system without generating any numerical model.

In addition, a novel experimental platform was established to investigate rapid transient Joule heating in embedded nanoscale metallic films representing buried on-chip interconnects that are not directly accessible. The effect of rapid transient power input profiles with different amplitudes and frequencies in 70-nm-thick Cu interconnects were studied employing sub-micron resistance thermometry (RTD) technique. The key contributions of this study can be outlines as:

- The novel platform developed in this study has the capability to be integrated within the structure and is especially suitable for embedded structures. Therefore, the proposed approach is suitable for measuring the spatial-temporal response of the interconnects in a buried 3-D IC structure, where techniques such as infrared or thermoreflectance microscopy cannot be used.

- Each fabricated RTD fits in a 6 μm by 7.7 μm rectangular space. Their relatively small size provides a highly spatially resolved thermal solution which can be employed for the detection of small local hotspots.
- The second advantage of these RTDs is their fairly small thermal time constants during transient measurements (9 μs rise time). Thus making them powerful measurement devices in monitoring temperature during thermal scenarios where there are abrupt rapid changes in voltage or current level of the system.
- By obtaining the frequency response of the system, it was shown that the RTDs can follow input power fluctuations up to 95 kHz and therefore the transient thermal measurement can be obtained with good accuracy for this high range of input frequencies.

Ultimately, a new concept was proposed to study the size effect on thermal conductivities of nanoscale embedded metallic films. The novel approach proposed here uses a laterally varying resistor structure to produce lateral heat gradient and to induce lateral heat diffusion in the plane of the metallic layer with strong sensitivity of the heating structure to the thermal conductivity of the metallic layer. Finally, through steady-state Joule heating and electrical resistance thermometry the thermal conductivity of the metallic films can be identified. The dependence of the thermal conductivity on temperature was also investigated and the results are well within the scope of previously reported values. The unique contributions of this study are summarized below:

- The proposed structure is the first device that has enabled the conductivity measurement of embedded metallic films on a substrate. By keeping the

underlying substrate, the interface between metal and dielectric will stay unaffected. As a result, the effect of the substrate and interface can be incorporated while having a satisfactory sensitivity to the thermal conductivity of the metallic film.

- It is important to note that the measured values for thermal conductivity are not dependent on the structure dimensions such as the width of the constriction, radius of the constriction, or the number of unit patterns per device. This observation further verifies the applicability of the proposed method.
- Another advantage of the proposed scheme is that it does not require extensive micro-fabrication. The structure can also be fabricated and integrated within the chip and hence can be used to investigate embedded or buried structures such as nanoscale on-chip interconnects.
- Although the method is proposed to measure metallic films, it can also be used for any material that can carry current and generate heat such as semiconductors.

The novel measurement technique proposed in this study can address some of the challenges that next generation IC industry encounters. Performance and reliability design of future microelectronic and nanoelectronic systems requires knowledge of the thermal and electrical properties of thin films. Among these, interconnects cooling and powering are becoming main bottlenecks. One of the proposed approaches is the use of alternate materials for on-chip interconnects with higher efficiencies has been proposed.

Due to the high thermal conductivity of graphene and carbon nanotubes, they are considered to be strong candidates for future nanoelectronics [139].

Graphene can be conveniently utilized within the integrated circuits due to its planar geometry. It also has high thermal conductivity in the lateral direction [140] and has various valuable electronic properties [141-143]. It is to be noted that the great performance of graphene strongly depends on its architecture and geometrical parameters. Therefore, it is essential to have a metrology that can be integrated within the structure and characterize the behavior of graphene in its real conditions. The proposed measurement technique can be suitably utilized to satisfy the aforementioned requirements.

Similar to graphene, carbon nanotubes exhibit strong thermal performance. Nevertheless, high processing temperatures are required for their synthesis eliminating the possibility of utilizing metallic sensors for thermal characterization and management. The unique approach introduced in this research will allow for the material of interest (here carbon nanotubes) to be used as thermal sensors and there is no need to introduce any other material to the structure. Furthermore, thermal conductivity of carbon nanotubes depends on the tubes diameter and its chirality [144], manifesting the need for their thermal characterization while they are embedded within the structure and are in their actual environmental conditions.

Another challenging situation in next generation ICs where the novel measurement approach can be incorporated is thermal management in military, automotive application, and space industry with extended high temperature ranges [145]. As mentioned before, the proposed metrology eliminates the need for conventional

metallic temperature sensors enabling the thermal management of electronics in harsh environmental conditions.

One of the important aspects of the future generation of Interconnects is their ever decreasing thickness. As the interconnect thickness decreases, the effect of underlying substrate will be more dominant due to the edge scattering effect. Also, the contact resistance between the metallic film and the underlying insulation layer will play a more important role in determining the thermal conductivity of the metallic film. Therefore, the need to have a metrology for thermal property extraction that can be implemented for the embedded or buried interconnects will become more significant.

Overall, the novel technique proposed in this study for the thermal property extraction of nano-scale embedded metallic films can address several important challenges that next generation IC industry will face.

7.2 Future Recommendations

The currently developed hybrid scheme works best for cases with either insulated or periodic boundary conditions. The model can be further expanded to include improvement of the multi-scale model for more realistic boundary conditions. Also, the effect of thermal interface resistance between the die and heatsink were neglected in this study. The proposed method can be advanced to include the effect of thermal interface materials.

In the experimental part of this research, the effect of the dielectric barrier layer on the thermal transient behavior of the interconnects can be investigated by fabricating devices with different dielectric layers and utilizing the developed experimental platform.

A comprehensive understanding of the size effect on the thermal conductivity of metallic films can be achieved by fabricating more devices to cover the film thicknesses that are below the mean free path of electrons in Cu (~ 40 nm). The current hourglass structure is fabricated over a silicon dioxide layer and its top surface is open to atmosphere. In order to resemble the embedded interconnects more closely, it is suggested that new hourglass structure will be fabricated with layers of SiO_2 on top and bottom. Moreover, proposed technique in this research can be further advanced by conducting a series of transient measurements in addition to the current steady state experiments. The time delay in the lateral heat diffusion of short pulses in the interconnects is related to the thermal conductivity and heat capacity of the interconnect. Utilizing this relationship and extracting the thermal conductivity from the proposed metrology, the heat capacity of buried on-chip interconnects can also be determined.

APPENDIX A: ANALYTICAL PROOF FOR POD MODEL

The distinctive characteristic of POD, demonstrated in Chapter 2, which claims that the POD modes will not change by changing the temporal dependence of the heat source for the fixed geometry, is analytically proven here.

The transient heat equation is considered (Eq. (2.6)):

$$\frac{\partial T(x,y,t)}{\partial t} - \frac{1}{\rho c_p} q'''(x,y,t) - \alpha \nabla^2 T(x,y,t) = 0 \tag{2.6}$$

with homogeneous boundary conditions (BCs):

$$\left. \frac{\partial T(x,y,t)}{\partial x} \right|_{y=0} = 0 \tag{A.15a}$$

$$\left. \frac{\partial T(x,y,t)}{\partial y} \right|_{x=0} = 0 \tag{A.15b}$$

and initial condition:

$$T(x,y,t=0) = f(x,y) = const. \tag{A.15c}$$

First, the case where $q'''(x,y,t) = 0$ is considered. For this case, Eq. (A.6) is simplified to

$$\frac{\partial T(x,y,t)}{\partial t} = \alpha \nabla^2 T(x,y,t) \tag{A.16}$$

Since Eq. (A.16) is a homogeneous partial differential equation (PDE) with homogeneous BCs, the method of eigenfunction expansion can be applied to solve for *T(x,y,t)* as

$$T(x,y,t) = \sum_r \sum_s \Gamma_{rs}(t) \Theta_{rs}(x,y) \tag{A.17}$$

where Θ_{rs} (x,y) are the basis functions. $\Gamma_{rs}(t)$ are the time dependent coefficients that can be found by projecting $T(x,y,t)$ onto the basis vectors Θ_{rs} (x,y). Ultimately, using the method of separation of variables, the temperature field can be written as:

$$T(x,y,t)=\sum_{r}\sum_{s}\Gamma_{rs}(t)X_{r}(x)Y_{s}(y) \tag{A.18}$$

Once the basis functions for the homogeneous equation (Eq. (A.16)) are determined, we look at the original nonhomogeneous Equation (Eq. (A.6)) which includes the source term $q'''(x,y,t)$. The basis functions are assumed to be the same for both homogeneous and nonhomogeneous PDEs. In other words, the source term $q'''(x,y,t)$ can also be expanded into the same basis functions for temperature; *i.e.* $X(x)$ and $Y(y)$, because $q'''(x,y,t)$ stays in the same Hilbert space as $T(x,y,t)$. Hence, the source term can be written as:

$$q'''(x,y,t)=\sum_{r}\sum_{s}\Pi_{rs}(t)X_{r}(x)Y_{s}(y) \tag{A.19}$$

Now that the basis functions are the same for both of the Eqs. (A.6) and (A.16), the time dependent coefficients $\Pi_{rs}(t)$ in Eq. (A.19) can be found by projecting $q'''(x,y,t)$ onto the basis vectors Θ_{rs} (x,y) as:

$$\Pi(t)=\left\langle X(x)Y(y),q'''(x,y,t)\right\rangle \tag{A.20}$$

The derivatives of temperature, $\dfrac{\partial^{2}T(x,y,t)}{\partial x^{2}}$, $\dfrac{\partial^{2}T(x,y,t)}{\partial y^{2}}$, and $\dfrac{\partial T(x,y,t)}{\partial t}$ are calculated from Eq. (A.18) and substituted into Eq. (A.6) which yields to:

$$\begin{aligned}&\sum_{r}\sum_{s}\frac{\partial\Gamma_{rs}(t)}{\partial t}X_{r}(x)Y_{s}(y)+\alpha\left(\sum_{r}\sum_{s}\mu_{r}^{2}\Gamma_{rs}(t)X_{r}(x)Y_{s}(y)+\sum_{r}\sum_{s}\gamma_{r}^{2}\Gamma_{rs}(t)X_{r}(x)Y_{s}(y)\right)\\&-\frac{1}{\rho C_{P}}\sum_{r}\sum_{s}\Pi_{rs}(t)X_{r}(x)Y_{s}(y)=0\end{aligned} \tag{A.21}$$

Simplifying this equation provides:

$$\sum_r \sum_s \left[\frac{\partial \Gamma_{rs}(t)}{\partial t} + \alpha \Gamma_{rs}(t) \left(\mu_r^2 + \gamma_s^2 \right) - \frac{1}{\rho C_P} \Pi_{rs}(t) \right] \times X_r(x) Y_s(y) = 0 \tag{A.22}$$

Applying the orthogonality condition to the basis functions *X(x)* and *Y(y)* results in:

$$\frac{\partial \Gamma_{rs}(t)}{\partial t} + \alpha \Gamma_{rs}(t) \left(\mu_r^2 + \gamma_s^2 \right) - \frac{1}{\rho C_P} \Pi_{rs}(t) = 0 \tag{A.23}$$

To find Γ^0_{rs}, the initial condition for Γ_{rs}, we apply Eq. (A.18) into the initial condition for the temperature, (Eq. (A.15c)) as:

$$f(x, y) = T(x, y, t = 0) = \sum_r \sum_s \Gamma_{rs}(t = 0) X_r(x) Y_s(y) = \sum_r \sum_s \Gamma^0{}_{rs} X_r(x) Y_s(y) \tag{A.24}$$

where Γ^0_{rs} can be determined by projecting *f(x,y,)* onto the basis functions *X(x)* and *Y(y)*as:

$$\Gamma^0{}_{rs}(t) = \left\langle X_r(x) Y_s(y), f(x, y) \right\rangle \tag{A.25}$$

Once Γ^0_{rs} is determined, Eq. (A.23) can be solved for Γ_{rs} in the entire domain. The discussed arguments demonstrate the reasoning behind the fact that once the POD modes are determined for the homogeneous problem, the source term (non-homogeneity) can be written in the form of a summation of the POD modes obtained originally. Hence, by varying the time dependency of the heat source in our problem there would be no need for any further observation generation.

REFERENCES

[1] T. Phan, S. Dilhaire, V. Quintard, D. Lewis, and W. Claeys, "Thermomechanical study of AlCu based interconnect under pulsed thermoelectric excitation," *Journal of Applied Physics,* vol. 81, pp. 1157-1157, 1997.

[2] "International Technology Roadmap for Semiconductors (ITRS), 2011 edn," *Executive Summary. Semiconductor Industry Association,* 2011.

[3] J. W. Evans, J. Y. Evans, P. Lall, and S. L. Cornford, "Thermomechanical failures in microelectronic interconnects," *Microelectronics Reliability,* vol. 38, pp. 523-529, 1998.

[4] A. F. Bastawros and K. S. Kim, "Experimental study on electric-current induced damage evolution at the crack tip in thin film conductors," *Transactions of the ASME. Journal of Electronic Packaging,* vol. 120, pp. 354-9, 1998.

[5] K. Fuchs, "The conductivity of thin metallic films according to the electron theory of metals," 1938, pp. 100-108.

[6] E. Sondheimer, "The mean free path of electrons in metals," *Advances in Physics,* vol. 1, pp. 1-42, 1952.

[7] J. M. Ziman and P. W. Levy, "Electrons and phonons," *Physics Today,* vol. 14, p. 64, 1961.

[8] Y. Namba, "Resistivity and temperature coefficient of thin metal films with rough surface," *Japanese Journal op Applied Physics Vol,* vol. 9, 1970.

[9] S. B. Soffer, "Statistical Model for the Size Effect in Electrical Conduction," *Journal of Applied Physics,* vol. 38, pp. 1710-1715, 1967.

[10] S. P. Gurrum, Y. K. Joshi, W. P. King, and K. Ramakrishna, "Numerical simulation of electron transport through constriction in a metallic thin film," *Electron Device Letters, IEEE,* vol. 25, pp. 696-698, 2004.

[11] A. A. Bilotti, "Static temperature distribution in IC chips with isothermal heat sources," *IEEE Transactions on Electron Devices,* vol. ED-21, pp. 217-26, 1974.

[12] Y. L. Shen, "Analysis of Joule heating in multilevel interconnects," *Journal of Vacuum Science & Technology B (Microelectronics and Nanometer Structures),* vol. 17, pp. 2115-21, 1999.

[13] C. C. Teng, Y. K. Cheng, E. Rosenbaum, and S. M. Kang, "iTEM: A temperature-dependent electromigration reliability diagnosis tool," *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on,* vol. 16, pp. 882-893, 2002.

[14] 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, pp. 197-205, 2002.

[15] T. Y. Chiang, K. Banerjee, and K. C. Saraswat, "Analytical thermal model for multilevel VLSI interconnects incorporating via effect," *Electron Device Letters, IEEE,* vol. 23, pp. 31-33, 2002.

[16] K. Banerjee, A. Amerasekera, N. Cheung, and C. Hu, "High-current failure model for VLSI interconnects under short-pulse stress conditions," *IEEE Electron Device Letters,* vol. 18, pp. 405-407, 1997.

[17] K. Banerjee, A. Mehrotra, W. Hunter, K. C. Saraswat, K. E. Goodson, and S. S. Wong, "Quantitative projections of reliability and performance for low-k/Cu interconnect systems," 2002, pp. 354-358.

[18] S. Im and K. Banerjee, "Full chip thermal analysis of planar (2-D) and vertically integrated (3-D) high performance ICs," 2002, pp. 727-730.

[19] T.-Y. Chiang, B. Shieh, and K. C. Saraswat, "Impact of joule heating on scaling of deep sub-micron Cu/low-k interconnects," in *2002 Symposium on VLSI Technology Digest of Technical Papers, June 11, 2002 - June 13, 2002*, Honolulu, HI, United states, 2002, pp. 38-39.

[20] S. P. Gurrum, Y. K. Joshi, W. P. King, K. Ramakrishna, and M. Gall, "A compact approach to on-chip interconnect heat conduction modeling using the finite element method," *Journal of Electronic Packaging,* vol. 130, pp. 031001.1-031001.8, 2008.

[21] Y. Joshi, "Reduced Order Thermal Models of Multiscale Microsystems," *Journal of Heat Transfer,* vol. 134, pp. 031008-1-031008-11, 2012.

[22] F. Christiaens, B. Vandevelde, E. Beyne, R. Mertens, and J. Berghmans, "A generic methodology for deriving compact dynamic thermal models, applied to the PSGA package," *IEEE Transactions on Components, Packaging, and Manufacturing Technology, Part A,* vol. 21, pp. 565-576, 1998.

[23] C. Lasance, H. Vinke, H. Rosten, and K. L. Weiner, "A novel approach for the thermal characterization of electronic parts," *Semiconductor Thermal Measurement and Management Symposium, Eleventh Annual IEEE SEMI-THERM XI.,* pp. 1-9, 7-9 Feb 1995 1995.

[24] Y. Gerstenmaier and G. Wachutka, "Rigorous model and network for transient thermal problems," *Microelectronics journal,* vol. 33, pp. 719-725, 2002.

[25] W. Krueger and A. Bar-Cohen, "THERMAL CHARACTERIZATION OF A PLCC-EXPANDED Rjc METHODOLOGY," 1992, p. 263.

[26] M. R. Stan, K. Skadron, M. Barcella, H. Wei, K. Sankaranarayanan, and S. Velusamy, "HotSpot: a dynamic compact thermal model at the processor-architecture level," *Microelectronics Journal,* vol. 34, pp. 1153-65, 2003.

[27] S. P. Gurrum, Y. K. Joshi, W. P. King, K. Ramakrishna, and M. Gall, "A compact approach to on-chip interconnect heat conduction modeling using the finite element method," *Journal of Electronic Packaging,* vol. 130, pp. 031001-1, 2008.

[28] D. Celo, G. Xiao Ming, P. K. Gunupudi, R. Khazaka, D. J. Walkey, T. Smy*, et al.*, "Hierarchical thermal analysis of large IC modules," *IEEE Transactions on Components and Packaging Technologies,* vol. 28, pp. 207-17, 2005.

[29] C. Christopoulos, "The transmission-line modeling method: TLM," *Antennas and Propagation Magazine, IEEE,* vol. 39, p. 90, 2002.

[30] D. De Cogan, W. O'Connor, and S. H. Pulko, *Transmission line matrix in computational mechanics*: CRC, 2006.

[31] B. Barabadi, Y. K. Joshi, S. Kumar, and G. Refai-Ahmed, "Thermal characterization of planar interconnect architectures under transient currents," in *ASME 2009 International Mechanical Engineering Congress and Exposition, IMECE2009, November 13, 2009 - November 19, 2009*, Lake Buena Vista, FL, United states, 2010, pp. 1381-1389.

[32] B. Barabadi, Y. K. Joshi, S. Kumar, and G. Refai-Ahmed, "Thermal characterization of planar interconnect architectures under different rapid transient currents using the transmission line matrix and finite element methods," in *2010 12th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2-5 June 2010*, Piscataway, NJ, USA, 2010, p. 8 pp.

[33] R. Ait-sadi and P. Naylor, "An investigation of the different TLM configurations used in the modelling of diffusion problems," *International Journal of Numerical Modelling: Electronic Networks, Devices and Fields,* vol. 6, pp. 253-268, 1993.

[34] T. Smy, D. Walkey, and S. Dew, "Transient 3D heat flow analysis for integrated circuit devices using the transmission line matrix method on a quad tree mesh," *Solid-State Electronics,* vol. 45, pp. 1137-1148, 2001.

[35] J. L. Lumley, "The Structure of Inhomogeneous Turbulent Flows," *Atmospheric Turbulence and Radio Wave Propagation,* pp. 166-178, 1967.

[36] P. Holmes, J. L. Lumley, and G. Berkooz, *Turbulence, coherent structures, dynamical systems and symmetry*: Cambridge University Press, 1998.

[37] B. Barabadi, Y. Joshi, and S. Kumar, "Prediction of Transient Thermal Behavior of Planar Interconnect Architecture Using Proper Orthogonal Decomposition Method," *ASME InterPACK,* pp. 213-224, July 6-8, 2011.

[38] Y. Hua and Z. Yu, "Generalized pencil-of-function method for extracting poles of an EM system from its transient response," *IEEE Transactions on Antennas and Propagation,* vol. 37, pp. 229-234, 1989.

[39] Y. Hua and Z. Yu, "On SVD for estimating generalized eigenvalues of singular matrix pencil in noise," *IEEE Transactions on Signal Processing,* vol. 39, pp. 892-900, 1991.

[40] L. Zao, S. X. D. Tan, W. Hai, R. Quintanilla, and A. Gupta, "Compact thermal modeling for package design with practical power maps," *International Green Computing Conference and Workshops (IGCC),* pp. 1-5, 25-28 July 2011 2011.

[41] L. Duo, S. X. D. Tan, E. H. Pacheco, and M. Tirumala, "Architecture-Level Thermal Characterization for Multicore Microprocessors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* vol. 17, pp. 1495-1507, 2009.

[42] L. Tang and Y. K. Joshi, "A multi-grid based multi-scale thermal analysis approach for combined mixed convection, conduction, and radiation due to discrete heating," *Journal of Heat Transfer,* vol. 127, pp. 18-26, 2005.

[43] B. Barabadi, Y. K. Joshi, and S. Kumar, "Prediction of Transient Thermal Behavior of Planar Interconnect Architecture Using Proper Orthogonal Decomposition Method," 2011.

[44] A. McNamara, V. Sahu, Y. Joshi, and Z. Zhang, "Infrared Imaging Microscope as an Effective Tool for Measuring Thermal Resistance of Emerging Interface Materials," in *ASME/JSME 2011 8th Thermal Engineering Joint Conference*, Honolulu, 2011.

[45] A. Majumdar, "Scanning thermal microscopy," *Annual review of materials science,* vol. 29, pp. 505-585, 1999.

[46] J. Varesi and A. Majumdar, "Scanning Joule expansion microscopy at nanometer scales," *Applied Physics Letters,* vol. 72, p. 37, 1998.

[47] G. B. M. Fiege, A. Altes, R. Heiderhoff, and L. J. Balk, "Quantitative thermal conductivity measurements with nanometre resolution," *Journal of Physics D: Applied Physics,* vol. 32, p. L13, 1999.

[48] H. Fischer, "Quantitative determination of heat conductivities by scanning thermal microscopy," *Thermochimica Acta,* vol. 425, pp. 69-74, 2005.

[49] V. Gorbunov, N. Fuchigami, J. Hazel, and V. Tsukruk, "Probing surface microthermal properties by scanning thermal microscopy," *Langmuir,* vol. 15, pp. 8340-8343, 1999.

[50] F. Ruiz, W. Sun, F. H. Pollak, and C. Venkatraman, "Determination of the thermal conductivity of diamond-like nanocomposite films using a scanning thermal microscope," *Applied Physics Letters,* vol. 73, p. 1802, 1998.

[51] H. Pollock and A. Hammiche, "Micro-thermal analysis: techniques and applications," *Journal of Physics D: Applied Physics,* vol. 34, p. R23, 2001.

[52] M. Malyj and J. Griffiths, "Stokes/anti-stokes Raman vibrational temperatures: Reference materials, standard lamps, and spectrophotometric calibrations," *Applied Spectroscopy,* vol. 37, pp. 315-333, 1983.

[53] P. Nath and K. Chopra, "Experimental determination of the thermal conductivity of thin films," *Thin Solid Films,* vol. 18, pp. 29-37, 1973.

[54] G. Pompe and K. Schmidt, "Vapour-deposited lead films and their transport characteristics at low temperatures II. thermal conductivity," *physica status solidi (a),* vol. 31, pp. 37-46, 1975.

[55] S. Shojaei-Zadeh, S. Zhang, W. Liu, Y. Yang, S. M. Sadeghipour, M. Asheghi*, et al.*, "Thermal characterization of thin film cu interconnects for the next generation of microelectronic devices," in *Thermal and Thermomechanical Phenomena in Electronic Systems, 2004. ITHERM'04. The Ninth Intersociety Conference on*, 2004, pp. 575-583.

[56] X. Zhang, H. Xie, M. Fujii, H. Ago, K. Takahashi, T. Ikuta*, et al.*, "Thermal and electrical conductivity of a suspended platinum nanofilm," *Applied Physics Letters,* vol. 86, pp. 171912-171912-3, 2005.

[57] L. Wenjun, Y. Yang, and M. Asheghi, "Thermal and electrical characterization and modeling of thin copper layers," in *Thermal and Thermomechanical Phenomena in Electronics Systems, 2006. ITHERM '06. The Tenth Intersociety Conference on*, 2006, pp. 1171-1176.

[58] F. Kelemen, "Pulse method for the measurement of the thermal conductivity of thin films," *Thin Solid Films,* vol. 36, pp. 199-203, 1976.

[59] D. G. Cahill, "Thermal conductivity measurement from 30 to 750 K: the 3 method," *Review of Scientific Instruments,* vol. 61, pp. 802–808, 1990.

[60] L. Lu, W. Yi, and D. Zhang, "3ω method for specific heat and thermal conductivity measurements," *Review of Scientific Instruments,* vol. 72, pp. 2996-3003, 2001.

[61] Y. Yang and M. Asheghi, "A novel technique for in-plane thermal conductivity measurements of electrically conductive interconnects and nanostructures," in *Thermal and Thermomechanical Phenomena in Electronic Systems, 2004. ITHERM'04. The Ninth Intersociety Conference on*, 2004, pp. 564-569.

[62] C. Paddock, "Transient Thermoreflectance From Thin Metal Films," *J. Appl. Phys.,* vol. 60, pp. 285-290, 1986.

[63] D. G. Cahill, W. K. Ford, K. E. Goodson, G. D. Mahan, A. Majumdar, H. J. Maris, *et al.*, "Nanoscale thermal transport," *Journal of Applied Physics,* vol. 93, p. 793, 2003.

[64] D. G. Cahill, K. Goodson, and A. Majumdar, "Thermometry and Thermal Transport in Micro/Nanoscale Solid-State Devices and Structures," *Journal of Heat Transfer,* vol. 124, pp. 223-241, 2002.

[65] B. Barabadi, S. Kumar, V. Sukharev, and Y. K. Joshi, "Multi-scale Transient Thermal Analysis of Microelectronics," in *ASME 2012 International Mechanical Engineering Congress & Exposition (IMECE2012)*, Houston, TX, 2012.

[66] L. Lennart, "System identification: theory for the user," *PTR Prentice Hall, Upper Saddle River, NJ,* 1999.

[67] T. Söderström and P. Stoica, *System identification*: Prentice-Hall, Inc., 1988.

[68] W. Polifke and A. Gentemann, "Order and realizability of impulse response filters for accurate identification of acoustic multi-ports from transient CFD," *Int. J. of Acoustics and Vibration,* vol. 9, pp. 139-148, 2004.

[69] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec*, et al.*, "Nonlinear black-box modeling in system identification: a unified overview," *Automatica,* vol. 31, pp. 1691-1724, 1995.

[70] P. De Boe and J.-C. Golinval, "Principal component analysis of a piezosensor array for damage localization," *Structural health monitoring,* vol. 2, pp. 137-144, 2003.

[71] G. Kerschen, J.-c. Golinval, A. F. Vakakis, and L. A. Bergman, "The method of proper orthogonal decomposition for dynamical characterization and order reduction of mechanical systems: an overview," *Nonlinear Dynamics,* vol. 41, pp. 147-169, 2005.

[72] U. Feldmann, E. Kreuzer, and F. Pinto, "Dynamic diagnosis of railway tracks by means of the Karhunen–Loeve transformation," *Nonlinear dynamics,* vol. 22, pp. 183-193, 2000.

[73] K. Pearson, "LIII. On lines and planes of closest fit to systems of points in space," *Philosophical Magazine Series 6,* vol. 2, pp. 559-572, 1901.

[74] D. Ahlman, F. Soderlund, J. Jackson, A. Kurdila, and W. Shyy, "Proper orthogonal decomposition for time-dependent lid-driven cavity flows," *Numerical Heat Transfer, Part B: Fundamentals,* vol. 42, pp. 285-306, 2002.

[75] G. Berkooz, P. Holmes, and J. L. Lumley, "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual Review of Fluid Mechanics,* vol. 25, pp. 539-575, 1993.

[76] G. Berkooz, P. Holmes, and J. Lumley, "Turbulence, coherent structures, dynamical systems and symmetry," *Cambridge Monographs on Mechanics, Cambridge University Press,* 1996.

[77] J. Cusumano, M. Sharkady, and B. Kimble, "Experimental measurements of dimensionality and spatial coherence in the dynamics of a flexible-beam impact oscillator," *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences,* vol. 347, p. 421, 1994.

[78] B. Feeny and R. Kappagantu, "On the physical interpretation of proper orthogonal modes in vibrations," *Journal of Sound and Vibration,* vol. 211, pp. 607-616, 1998.

[79] J. A. Atwell and B. B. King, "Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations* 1," *Mathematical and Computer Modelling,* vol. 33, pp. 1-19, 2001.

[80] Y. Liang, W. Lin, H. Lee, S. Lim, K. Lee, and H. Sun, "Proper orthogonal decomposition and its applications-part II: Model reduction for MEMS dynamical analysis," *Journal of Sound and Vibration,* vol. 256, pp. 515-532, 2002.

[81] L. Codecasa, D. D'Amore, and P. Maffezzoni, "An Arnoldi based thermal network reduction method for electro-thermal analysis," *Components and Packaging Technologies, IEEE Transactions on,* vol. 26, pp. 186-192, 2003.

[82] Bia and R. lstrok, "Proper orthogonal decomposition and modal analysis for acceleration of transient FEM thermal analysis," *International Journal for Numerical Methods in Engineering,* vol. 62, pp. 774-797, 2005.

[83] R. Bia ecki, A. Kassab, and A. Fic, "Reduction of the dimensionality of transient FEM solutions Using Proper Orthogonal Decomposition," 2003, p. 2003.

[84] A. Fic, R. A. Bia ecki, and A. J. Kassab, "Solving transient nonlinear heat conduction problems by proper orthogonal decomposition and the finite-element method," *Numerical Heat Transfer, Part B: Fundamentals,* vol. 48, pp. 103-124, 2005.

[85] L. G. Bleris and M. V. Kothare, "Reduced order distributed boundary control of thermal transients in microsystems," *Control Systems Technology, IEEE Transactions on,* vol. 13, pp. 853-867, 2005.

[86] A. P. Raghupathy, U. Ghia, K. Ghia, and W. Maltz, "Boundary-Condition-Independent Reduced-Order Modeling of Heat Transfer in Complex Objects by POD-Galerkin Methodology: 1D Case Study," *Journal of Heat Transfer,* vol. 132, p. 064502, 2010.

[87] I. T. Jolliffe and MyiLibrary, *Principal component analysis* vol. 2: Wiley Online Library, 2002.

[88] R. L. Grossman and C. Kamath, *Data mining for scientific and engineering applications*: Springer Netherlands, 2001.

[89] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 12, pp. 103-108, 2002.

[90] P. Eriksson, C. Jiménez, S. Bühler, and D. Murtagh, "A Hotelling transformation approach for rapid inversion of atmospheric spectra," *Journal of Quantitative Spectroscopy and Radiative Transfer,* vol. 73, pp. 529-543, 2002.

[91] Y. Liang, H. Lee, S. Lim, W. Lin, K. Lee, and C. Wu, "PROPER ORTHOGONAL DECOMPOSITION AND ITS APPLICATIONS--PART I: THEORY," *Journal of Sound and Vibration,* vol. 252, pp. 527-544, 2002.

[92] A. Chatterjee, "An introduction to the proper orthogonal decomposition," *Current science,* vol. 78, pp. 808-817, 2000.

[93] N. W. Rolander, "An Approach for the Robust Design of Data Center Server Cabinets," Georgia Institute of Technology, 2005.

[94] K. Bizon, G. Continillo, L. Russo, and J. Smula, "On POD reduced models of tubular reactor with periodic regimes," *Computers & Chemical Engineering,* vol. 32, pp. 1305-1315, 2008.

[95] M. D. Graham and I. G. Kevrekidis, "Alternative approaches to the Karhunen-Loeve decomposition for model reduction and data analysis," *Computers & Chemical Engineering,* vol. 20, pp. 495-506, 1996.
[96] C. W. Rowley, T. Colonius, and R. M. Murray, "Dynamical models for control of cavity oscillations," *AIAA paper,* vol. 2126, pp. 2126–34, 2001.
[97] P. Ding, X. H. Wu, Y. L. He, and W. Q. Tao, "A fast and efficient method for predicting fluid flow and heat transfer problems," *Journal of Heat Transfer,* vol. 130, p. 032502, 2008.
[98] H. V. Ly and H. T. Tran, "Modeling and control of physical processes using proper orthogonal decomposition," *Mathematical and Computer Modelling,* vol. 33, pp. 223-236, 2001.
[99] J. Rambo and Y. Joshi, "Reduced-order modeling of turbulent forced convection with parametric conditions," *International journal of heat and mass transfer,* vol. 50, pp. 539-551, 2007.
[100] J. D. Rambo, "Reduced-order modeling of multiscale turbulent convection: Application to data center thermal management," Citeseer, 2006.
[101] R. Темам, "Infinite Dimensional Dynamical Systems in Mechanics and Physics," *Applied Mathematical Sciences,* vol. 68, 1988.
[102] Z. M. Zhang, *Nano/microscale heat transfer*: McGraw-Hill Professional, 2007.
[103] B. Barabadi, Y. Joshi, and S. Kumar, "Prediction of Transient Thermal Behavior of Planar Interconnect Architecture Using Proper Orthogonal Decomposition Method," *ASME InterPACK,* July 6-8, 2011.
[104] F. P. Incropera, T. L. Bergman, A. S. Lavine, and D. P. DeWitt, *Fundamentals of heat and mass transfer*: Wiley, 2011.
[105] K. C. Chang, Y. Li, C. Y. Lin, and M. J. Lii, "Design Guidance for the Mechanical Reliability of Low-K Flip Chip BGA Package 1," *International Microelectronics and Packaging Society (IMAPS) Topical Workshop and Exhibition on Flip Chip Technology* pp. 21-24, 06/2004 2004.
[106] B. Barabadi, S. Kumar, V. Sukharev, and Y. K. Joshi, "Multi-scale Transient Thermal Analysis of Microelectronics," *ASME 2012 International Mechanical Engineering Congress & Exposition (IMECE2012),* November 9-15 2012 2012.
[107] B. Barabadi, S. Kumar, and Y. K. Joshi, "Rapid Multi-scale Transient Thermal Modeling of Packaged Microprocessors Using Hybrid Approach," in *The 14th Electronics Packaging Technology Conference (EPTC)*, Sentosa Island, Singapore, 2012.
[108] V. George, S. Jahagirdar, C. Tong, K. Smits, S. Damaraju, S. Siers*, et al.*, "Penryn: 45-nm next generation Intel® core™ 2 processor," in *Solid-State Circuits Conference, 2007. ASSCC'07. IEEE Asian*, 2007, pp. 14-17.
[109] N. Sakran, M. Yuffe, M. Mehalel, J. Doweck, E. Knoll, and A. Kovacs, "The implementation of the 65nm dual-core 64b merom processor," in *Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International*, 2007, pp. 106-590.
[110] (2008). *Intel® Core™ 2 Duo Mobile Processors on 45-nm process for Embedded Applications-Thermal Design Guide*. Available: www.intel.com

[111] K. C. Chang, Y. Li, C. Y. Lin, and M. J. Lii, "Design Guidance for the Mechanical Reliability of Low-K Flip Chip BGA Package 1," in *International Microelectronics and Packaging Society Conference*, 2004.

[112] Y. J. Kim, Y. K. Joshi, A. G. Fedorov, Y. J. Lee, and S. K. Lim, "Thermal characterization of interlayer microfluidic cooling of three-dimensional integrated circuits with nonuniform heat flux," *Journal of Heat Transfer,* vol. 132, pp. 041009-1, 2010.

[113] M. B. Healy and S. K. Lim, "A study of stacking limit and scaling in 3D ICs: An interconnect perspective," in *Electronic Components and Technology Conference, 2009. ECTC 2009. 59th*, 2009, pp. 1213-1220.

[114] S. Dilhaire, S. Grauby, and W. Claeys, "Calibration procedure for temperature measurements by thermoreflectance under high magnification conditions," *Applied Physics Letters,* vol. 84, p. 822, 2004.

[115] V. Szekely and T. Van Bien, "Fine structure of heat flow path in semiconductor devices: a measurement and identification method," *Solid-State Electronics,* vol. 31, pp. 1363-1368, 1988.

[116] P. Szabo, M. Rencz, G. Farkas, and A. Poppe, "Short time die attach characterization of LEDs for in-line testing application," 2007, pp. 360-366.

[117] V. Székely, in *International Workshop on Thermal Investigations of ICs and Systems (THERMINICS)*, Rome, Italy, September 24–26, 2008 (unpublished).

[118] Y. Ezzahri and A. Shakouri, "Application of network identification by deconvolution method to the thermal analysis of the pump-probe transient thermoreflectance signal," *Review of Scientific Instruments,* vol. 80, p. 074903, 2009.

[119] J. Varesi and A. Majumdar, "Scanning Joule expansion microscopy at nanometer scales," *Applied Physics Letters,* vol. 72, pp. 37-39, 1998.

[120] H. Brugger and P. W. Epperlein, "Mapping of local temperatures on mirrors of GaAs/AlGaAs laser diodes," *Applied Physics Letters,* vol. 56, pp. 1049-1051, 1990.

[121] J. Christofferson, K. Maize, Y. Ezzahri, J. Shabani, X. Wang, and A. Shakouri, "Microscale and Nanoscale Thermal Characterization Techniques," *Journal of Electronic Packaging,* vol. 130, pp. 041101-6, 2008.

[122] (November 11, 2013). *platinum (Pt)*. Available: http://www.britannica.com/EBchecked/topic/464081/platinum-Pt

[123] A. Weber, J. Lang, and A. Slocum, "{111} Si Etched Planar Electrical Contacts for Power MEMS-relays," in *Electrical contacts-2007, the 53rd ieee holm conference on*, 2007, pp. 156-159.

[124] R. S. Timsit, "Constriction resistance of thin film contacts," *Components and Packaging Technologies, IEEE Transactions on,* vol. 33, pp. 636-642, 2010.

[125] M. Read, J. Lang, A. Slocum, and R. Martens, "Contact Resistance in Flat-on-Flat and Sphere-on-Flat Thin Films," in *Electrical Contacts (HOLM), 2010 Proceedings of the 56th IEEE Holm Conference on*, 2010, pp. 1-8.

[126] L. Kogut and K. Komvopoulos, "Electrical contact resistance theory for conductive rough surfaces," *Journal of Applied Physics,* vol. 94, pp. 3153-3162, 2003.

[127] R. Holm, "Electrical Contacts-Theory and Applications," *Berlin Germany: Springer-Verlag,* 1967.
[128] R. S. Timsit, "Electrical conduction through small contact spots," *Components and Packaging Technologies, IEEE Transactions on,* vol. 29, pp. 727-734, 2006.
[129] J. R. Lloyd, "Materials Reliability in Microelectronics III," in *Materials Research Society*, Pittsburgh, PA, 1993, p. 177.
[130] A. H. Atabaki, *Reconfigurable silicon photonic devices for optical signal processing*: Georgia Institute of Technology, 2011.
[131] F. Volklein and H. Balles, "A microstructure for measurement of thermal conductivity of polysilicon thin films," *Microelectromechanical Systems, Journal of,* vol. 1, pp. 193-196, 1992.
[132] M. von Arx, O. Paul, and H. Baltes, "Process-dependent thin-film thermal conductivities for thermal CMOS MEMS," *Microelectromechanical Systems, Journal of,* vol. 9, pp. 136-145, 2000.
[133] T. Kemp, T. A. S. Srinivas, R. Fettig, and W. Ruppel, "Measurement of thermal diffusivity of thin films and foils using a laser scanning microscope," *Review of Scientific Instruments,* vol. 66, pp. 176-181, 1995.
[134] G. Langer, J. Hartmann, and M. Reichling, "Thermal conductivity of thin metallic films measured by photothermal profile analysis," *Review of Scientific Instruments,* vol. 68, pp. 1510-1513, 1997.
[135] S. P. Gurrum, W. P. King, Y. K. Joshi, and K. Ramakrishna, "Size Effect on the Thermal Conductivity of Thin Metallic Films Investigated by Scanning Joule Expansion Microscopy," *Journal of Heat Transfer,* vol. 130, p. 082403, 2008.
[136] P. Nath and K. Chopra, "Thermal conductivity of copper films," *Thin Solid Films,* vol. 20, pp. 53-62, 1974.
[137] P. Klemens, "The thermal conductivity of dielectric solids at low temperatures (theoretical)," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* vol. 208, pp. 108-133, 1951.
[138] R. Powell, C. Ho, and P. Liley, "NSRDS-NBS-8: Thermal Conductivity of Selected Materials," *US Department of Commerce, Springfield, VA,* 1966.
[139] A. A. Balandin, "Thermal properties of graphene and nanostructured carbon materials," *Nature materials,* vol. 10, pp. 569-581, 2011.
[140] A. Balandin, "Chill Out: New Materials Can Keep Chips Cool," *IEEE Spectrum,* vol. 29, p. 35, 2009.
[141] C. Berger, Z. Song, X. Li, X. Wu, N. Brown, C. Naud*, et al.*, "Electronic confinement and coherence in patterned epitaxial graphene," *Science,* vol. 312, pp. 1191-1196, 2006.
[142] M. Y. Han, B. Özyilmaz, Y. Zhang, and P. Kim, "Energy band-gap engineering of graphene nanoribbons," *Physical review letters,* vol. 98, p. 206805, 2007.
[143] Y. Zhang, Y.-W. Tan, H. L. Stormer, and P. Kim, "Experimental observation of the quantum Hall effect and Berry's phase in graphene," *Nature,* vol. 438, pp. 201-204, 2005.
[144] X. Yan, Y. Xiao, and Z. Li, "Effects of intertube coupling and tube chirality on thermal transport of carbon nanotubes," *Journal of applied physics,* vol. 99, pp. 124305-124305-4, 2006.

[145] S. V. Garimella, A. S. Fleischer, J. Y. Murthy, A. Keshavarzi, R. Prasher, C. Patel, *et al.*, "Thermal challenges in next-generation electronic systems," *Components and Packaging Technologies, IEEE Transactions on,* vol. 31, pp. 801-815, 2008.