Advanced Thermal Characterization and Temperature Control to Enable the Next Generation of Micro-Electronic Technologies

Assaad El Helou

Southern Methodist University, ahelou@smu.edu

Follow this and additional works at: https://scholar.smu.edu/engineering_mechanical_etds

Part of the Mechanical Engineering Commons

Recommended Citation

https://scholar.smu.edu/engineering_mechanical_etds/31

This Dissertation is brought to you for free and open access by the Mechanical Engineering at SMU Scholar. It has been accepted for inclusion in Mechanical Engineering Research Theses and Dissertations by an authorized administrator of SMU Scholar. For more information, please visit http://digitalrepository.smu.edu
ADVANCED THERMAL CHARACTERIZATION AND TEMPERATURE CONTROL
TO ENABLE THE NEXT GENERATION OF MICRO-ELECTRONIC
TECHNOLOGIES

Approved by:

________________________________________
Dr. Peter Raad
Professor

________________________________________
Dr. Ali Beskok
Professor

________________________________________
Dr. Yildirim Hurmuzlu
Professor

________________________________________
Dr. David Willis
Associate Professor

________________________________________
Dr. Archana Venugopal
Texas Instruments
ADVANCED THERMAL CHARACTERIZATION AND TEMPERATURE CONTROL
TO ENABLE THE NEXT GENERATION OF MICRO-ELECTRONIC
TECHNOLOGIES

A Dissertation Presented to the Graduate Faculty of the
Lyle School of Engineering
Southern Methodist University
in
Partial Fulfillment of the Requirements
for the degree of
Doctor of Philosophy
with a
Major in Mechanical Engineering
by
Assaad El Helou
(M.S., Mechanical Engineering, Southern Methodist University)
(B.Eng., Mechanical Engineering, American University of Beirut)

December 19, 2020
ACKNOWLEDGMENTS

This work could not have been accomplished without the help and constant support of my advisor, Prof. Peter E. Raad, and the long and engaging research discussions during which I had the chance to benefit from his vast knowledge and experience in the field of thermal sciences and engineering. He had a lasting impact on my career and both my personal and professional character, to which I am grateful. I also would like to acknowledge the faculty of the Departments of Mechanical Engineering and Electrical Engineering at SMU who provided their support in both academics and in research throughout my graduate studies.

I would like to thank TMX Scientific for their support in relation to the thermal imaging system and software and the thermal simulation software, the NSRL cleanroom staff at the UT-Dallas for training and supporting me in microfabrication, and my collaborators at Texas Instruments (TI), the Naval Research Lab (NRL), the Australian National University (ANU), and the Bio-microfluidics Research team at SMU (Dr. Ali Beskok and Dr. Anil Koklu) for their contributions and support. I also gratefully acknowledge Texas Instruments for funding my research endeavors at Nanoscale Electro-Thermal Sciences Laboratory (NETS lab) during my PhD, and the Moody Graduate School for granting me the Dissertation Fellowship Scholarship for my last semester at SMU.
In the electronics world, self-heating is an inevitable by-product of electrical activation that has a major impact on device performance and reliability. Thermal technologies have been in constant development to effectively dissipate the generated heat and keep device operation temperatures within reliable limits. Moreover, thermal characterization technologies have been implemented to understand the thermal performance within microelectronic systems, but not without facing experimental and numerical challenges. This work presents advanced thermal investigations, both experimental and numerical, that are adapted and most suited for emerging micro-electronic technologies. Initially, the main experimental and numerical modeling challenges faced in the thermal analysis of micro-electronics are discussed. The advanced thermal characterization techniques are then presented, from high-resolution thermoreflectance based thermal imaging to adaptive multi-grid numerical techniques. Moreover, a coupled experimental and numerical thermal investigation is presented for a more complete thermal assessment of micro-scale semiconductor devices. The coupled approach provides quick and accurate analysis of the thermal aspect of complex micro-electronic device operation and allows the advancement in design and reliability of devices used in numerous applications such as high voltage switching, low power micro-sensing, non-volatile memory, and neuromorphic computing. These thermal characterization strategies are also applied at the integration level of integrated circuits where interconnects reliability is also impacted by their thermal performance. A full review is presented for each of the studied device tech-
nologies in an effort to suitably and effectively apply the presented thermal characterization techniques to address the unique requirements and challenges.
# TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ xii
LIST OF TABLES ........................................................................................................... xxi

CHAPTER

1. Introduction to Microelectronics Thermal Management ......................... 1
   1.1. Trends in Microelectronic Devices Constrict Thermal Constraints ........ 1
   1.2. Temperature Effects on Reliability and Performance ....................... 3
       1.2.1. Failure and Degradation in High Power Devices .................... 4
       1.2.2. Strict Temperature Control for Micro-Sensing Devices ............ 5
       1.2.3. Joule-Heating Induced Electrical Characteristics .................. 5
   1.3. Outline ........................................................................................................... 6

2. Experimental Thermal Characterization ......................................................... 8
   2.1. Thermal Microscopy Techniques: Capabilities and Shortcomings .......... 9
       2.1.1. Infrared Imaging ......................................................................... 9
       2.1.2. Micro-TCs and SThM ................................................................. 10
       2.1.3. Micro-Raman ............................................................................. 10
   2.2. Thermoreflectance based Thermal Imaging ........................................... 12
       2.2.1. Governing Physics of Reflectance-Temperature Dependence ...... 13
   2.3. TR Imaging Acquisition Methodology .................................................. 15
       2.3.1. TR Experimental Setup for Thermal Microscopy ..................... 15
       2.3.2. TR Measurement Procedure ....................................................... 15
           2.3.2.1. Activation Step for \((\Delta R/R)_a\) Map ................................. 18
           2.3.2.2. Calibration Step for \(C_{TR}\) Map ......................................... 18
           2.3.2.3. Extracting Thermal Maps from \(C_{TR}\) and \((\Delta R/R)_a\) ........ 20
6. Compact 3D Thermal Model for VLSI and ULSI Interconnect Network Reliability Verification

6.1. Introduction

6.1.1. Electromigration Failures and the Degradation of IC BEmL Networks

6.2. 2D Model for Thermal Conductance from an Isolated Interconnect

6.2.1. Reduced Mathematical Models for Effective Conductivity

6.2.1.1. 1D Model

6.2.1.2. Bilotti 2D Model

6.2.1.3. Andrews 2D Model

6.3. Reduced 3D Thermal Model for an Interconnect Network

6.3.1. 2+1 D Heat Conduction: The Lossy Wire Model

6.3.1.1. The Thermal Entrance Length

6.3.1.2. The Maximum Attainable Temperature

6.3.1.3. The End Temperatures

6.3.2. Thermal Resistance Network Analysis

6.3.2.1. ICT Entrance Thermal Resistance

6.3.2.2. Via Thermal Resistance

6.3.2.3. Cross Configuration Thermal Resistance

6.3.3. Iterative Approach to Solve Resistance Network

6.3.4. Model Validation with Simulations and Thermal Imaging

6.4. TDTR Conductivity Measurements of BEmL Thin-film Materials

6.4.1. Studied Materials

6.4.2. TDTR Results

6.4.3. TDTR Study Conclusions

6.5. 2D Combined Model Results
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.5.1</td>
<td>Validation Simulations with Thermal Maps</td>
<td>105</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Verifying 2D Reduced model with Simulations</td>
<td>108</td>
</tr>
<tr>
<td>6.6</td>
<td>3D Compact Model Results</td>
<td>110</td>
</tr>
<tr>
<td>6.6.1</td>
<td>Validating Simulations with Thermal Maps</td>
<td>110</td>
</tr>
<tr>
<td>6.6.2</td>
<td>Verifying 3D Reduced Model with Simulations</td>
<td>115</td>
</tr>
<tr>
<td>6.7</td>
<td>Conclusions</td>
<td>118</td>
</tr>
<tr>
<td>7</td>
<td>Standardized Heat Spreader Design for Passive Cooling of Interconnects in the BEOL of ICs</td>
<td>121</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>122</td>
</tr>
<tr>
<td>7.2</td>
<td>Thermal Simulation Model</td>
<td>124</td>
</tr>
<tr>
<td>7.3</td>
<td>Heat Spreader Design</td>
<td>130</td>
</tr>
<tr>
<td>7.3.1</td>
<td>Temperature Measurements of Sample Devices</td>
<td>130</td>
</tr>
<tr>
<td>7.3.2</td>
<td>Experimentally Validated Thermal Model</td>
<td>132</td>
</tr>
<tr>
<td>7.3.3</td>
<td>Simulation and Experimental Results</td>
<td>133</td>
</tr>
<tr>
<td>7.3.4</td>
<td>Parametric Study</td>
<td>135</td>
</tr>
<tr>
<td>7.4</td>
<td>Conclusions</td>
<td>137</td>
</tr>
<tr>
<td>8</td>
<td>Temperature Induced Switching in NbO$_x$ Current-Controlled Negative Differential Resistance (CC-NDR) Memristive Devices</td>
<td>139</td>
</tr>
<tr>
<td>8.1</td>
<td>Memristor: the Fourth Fundamental Passive Circuit Element</td>
<td>139</td>
</tr>
<tr>
<td>8.2</td>
<td>Transport phenomena in Thin Oxides</td>
<td>141</td>
</tr>
<tr>
<td>8.3</td>
<td>Applications of memristor devices</td>
<td>142</td>
</tr>
<tr>
<td>8.3.1</td>
<td>Non-volatile memory</td>
<td>143</td>
</tr>
<tr>
<td>8.3.2</td>
<td>Neuromorphic computing (neuristor)</td>
<td>144</td>
</tr>
<tr>
<td>8.4</td>
<td>Temperature-induced CC-NDR switching characteristics</td>
<td>145</td>
</tr>
<tr>
<td>8.5</td>
<td>Sample Description and Methodology</td>
<td>149</td>
</tr>
<tr>
<td>8.6</td>
<td>Electric Field- and Current-Induced Electroforming Modes in NbO$_x$ [176]</td>
<td>152</td>
</tr>
</tbody>
</table>
8.7. Current Localization and Redistribution as the Basis of Discontinuous Current Controlled Negative Differential Resistance in NbO$_x$ [177] ............ 154

8.8. Schottky barrier induced asymmetry in the negative differential resistance response of Nb/NbO$_x$/Pt cross-point devices [178]. ......................... 161

8.9. Thermal Conductivity of Amorphous Sub-stoichiometric NbO$_x$ for Memristors Applications using Transient-Domain Thermoreflectance ............ 164

8.9.1. Experimental Methods ................................................................. 166

8.9.2. TDTR Results and Discussion ...................................................... 166


9.1. AC Electrothermal Flow for Electrokinetic Actuation of Liquid Media .... 170

9.2. Temperature Control for ACET Applications ................................. 171

9.3. Materials and Methodology ............................................................. 173

9.4. Results ................................................................................................. 177

9.4.1. Frequency-Dependent Self-Heating .............................................. 177

9.4.2. Voltage-Dependent Self-Heating .................................................... 179

9.4.3. Power-Loss Dependent Self-Heating .............................................. 181

9.5. Conclusions ......................................................................................... 184

10. Conclusions ........................................................................................ 185

BIBLIOGRAPHY ......................................................................................... 187
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>a) Logic technology node and gate length b) resistance and c) capacitance of different technologies showing the improvement in performance with device scaling. (Adapted from [248])</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>Graph showing the increased heat flux densities in advance microelectronic devices (represented by bars) resulting from increasing power levels and downscaling in device sizes. Red circles show comparison to other natural and man-made heat generation rates for scale. (Source: DARPA)</td>
<td>3</td>
</tr>
<tr>
<td>2.1</td>
<td>Thermal maps of active GaN HEMT using a) IR and b) micro-Raman imaging from Sarua et al. [220]. c) Device measurement region and Raman measurements adapted from Hiroki et al. [105]. The high-resolution requirements are evident especially as device dimensions near the sub-micron scale.</td>
<td>11</td>
</tr>
<tr>
<td>2.2</td>
<td>Schematic showing optical setup of TR imaging imaging system and the incident and reflected light path from the sample.</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>TMX T°Imager® setup showing activation, temperature, PC, nano-positioning, and light control used for sample thermal imaging.</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Schematic representation of pump-and-probe technique for the time-domain thermoreflectance.</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>a) Schematic of HEMT device structure. The substrate is either Si, SiC, Diamond, or is etched b) Top view image showing Source, Drain and Gate contacts.</td>
<td>43</td>
</tr>
<tr>
<td>4.2</td>
<td>Sample a) calibration map for $C_{TR}$ showing a global reflectance change response from the GaN material, b) reflectance change map $\Delta R/R$ showing the localized response within the GaN Channel Region, c) the resulting temperature rise map $\Delta T$ for the active GaN-membrane and GaN-on-Si device showing self-heating within the channel.</td>
<td>44</td>
</tr>
<tr>
<td>4.3</td>
<td>a) The calibration curve of CTR vs T showing a decrease in TR response at higher temperatures. b) Temperature rise curve as a function of applied power for GaN-on-Si device before and after correction for CTR-T dependence. The correction recovers the linearity of P-T relation for the GaN-on-Si devices.</td>
<td>49</td>
</tr>
</tbody>
</table>
4.4 Calibration results for GaN-on-SiC and GaN-on-diamond showing $C_{TR}$ linear decrease with temperature. .................................................. 50

4.5 A summary of the measured Average (open markers) and Peak (solid markers) experimental temperature rise (in °C) within the active channel of the measured GaN HEMT devices on the different substrates. The linear experimental fit are shown for the measurement data for the average (in dotted lines) and peak temperature (in long dashed lines). The heating rate (total thermal resistance) is reported from the linear regression analysis. .................................................. 51

4.6 a) Measured thermal maps showing temperature rise in GaN HEMTs active region on SiC (left) and diamond (right) substrate (P≈0.5 W, 6.6 W/mm). (b) The local thermal resistance across the Gate-Drain channel shows peak temperature rise at the drain-side edge of the gate. ........ 54

4.7 a) Measured thermal maps for GaN-on-diamond device 2 [239] with a different gate geometry than our measured device b) Comparison of temperature profile for both measured GaN-on-diamond devices with different drain channel length shows similar temperature rise. .................... 55

5.1 a) Schematic of HEMT device structure. The substrate is either Si, SiC, Diamond, or is etched b) intensity image (1024x1024 pixels) of the device under 100× magnification showing the Source, Drain and Gate contacts. Purple regions indicate the GaN drain and source channel where the thermal measurements are conducted c) device model built to scale in T°Solver®. .......................................................... 64

5.2 x-SEM measurements showing thickness and composition of GaN/AlGaN stack. 65

5.3 Converged temperature profiles for the matched numerical and experimental thermal maps for the gate-drain access region of a GaN-on-Si device for a DC power level ($I_{DS} \times V_{DS}$) of 6.4 W/mm. .............................. 69

5.4 Comparison of matched experimental and simulation results for GaN-on-SiC for an activation power level of 15.7 W/mm. ........................ 70

5.5 Sample temperature rise profile for the matched numerical and experimental results for GaN-on-diamond device 1 and device 2. .................. 73

5.6 Variation of temperature profile in GaN-on-Diamond HEMT for (a) the TBR range between 25 and 5 m²K/GW and (b) for the conductivity range $k_{di,xz}$ between 1000 and 1400 W/m.K with $k_{di,y}=1.45 \times k_{xz,di}$. ............ 77
5.7 a) Results of parametric study showing the total thermal resistance for the studies values of diamond conductivities and TBR values, based on maximum and average GaN channel temperatures, and power level of 16.4 W/mm. The through-plane conductivity is taken as $1.45 \times k_{xz,di}$.

b) The corresponding percent change in self-heating temperature rise with reported sensitivity values to changes in $k_{di}$ and TBR.

---

6.1 Image of interconnect showing hillock and void formation due to material transport under Electromigration. (Source: Kyung-Hoae 2011)

---

6.2 a) Schematic of the device structure considered for the 2D thermal model showing the ICT embedded in the dielectric medium b) Corresponding thermal resistive circuit showing the heat source generated $Q'_{gen}$ and the total thermal resistance $R_{th}$ c) The electrical activation circuit for the ICT showing the electrical resistance $R$ and current density $J$.

---

6.3 Schematic showing a) thermally long ($L/\lambda_h > 5$) and b) thermally short ($L/\lambda_h < 5$) ICT with the considered entrance thermal resistance at the nodes and the associated temperature distribution for a given $\lambda_H$. Zero boundary condition is considered $\Delta T_n = \Delta T_0 = 0$ and $\Delta T_k = \Delta T_L = 0$ at both ends $x = 0$ and $x = L$. Note that the temperature distribution in a thermally short ICT does not reach $\Delta T_{max}$.

---

6.4 Schematic showing the resistive network configuration for a) Single Resistive Cross model b) Six Resistors Cross Model based on [40]. The value of R is such that the equivalent network resistance is equal for the two ICTs in both configurations. c) Resistance network schematic showing the six-resistor model [48] connected to the ICT segment ends.

---

6.5 Total thermal resistances obtained for the thin-film samples as function of film thickness. The slope and intercept of each data fit are used to extract the intrinsic thermal conductivity of the film and its interface thermal resistance.

---

6.6 Summary of the measured intrinsic thermal conductivities of the BEoL materials and their interface thermal resistances.

---

6.7 The calculated effective thermal conductivities of the sample films as reconstructed from the extracted intrinsic thermal conductivities and interface resistances.

---

6.8 Thermal maps of the sample BEoL devices micro-resistors of widths 1, 3 and 10 $\mu$m at the different locations: Level 1 of Polysilicon (Poly) shown in the left column and level 2 of AlCu metalization (MET1) shown in the right column.
6.9 Comparison of the heating rate for measured ICT samples as compared with their estimates from the reduced model, with $\kappa_d = 1.1 W/m.K$.

6.10 Comparison between the newly developed combined model and the simulation results, the models of Bilotti and Andrews are also shown for each case to highlight the improvement of using the combined model.

6.11 Combined model chart showing effective conductance $\kappa_{d,eff}/\kappa_d$ from an ICT of width $w$, insulation thickness $t_i$ and passivation thickness $t_p$, all of which are denoted in the added schematic. The Bilotti and Andrews cases show the extreme effects of, respectively, ignoring passivation or making it infinite.

6.12 Schematic showing the resistive network configuration for a) an polyresistor ICT in Level 1 between two via connections to contact leads in the higher MET level, and b) the same configuration with a "sense" lead in the MET layer with the cross coupling. The resistive network that connects the device components is also shown in an overlay that shows the thermal resistance coupling of the different features (entry, via, and cross resistance).

6.13 Graphical user interface for 2D reduced model built in MATLAB GUIDE. The interface shows input electrical and thermal parameters for an isolated interconnect, and provides the allowable current density levels. The model also reports on the effective thermal conductivity of the surrounding dielectric.

6.14 Graphical user interface for the 3D reduced model built in MATLAB GUIDE. The interface shows input electrical and thermal parameters for an isolated interconnect as well as the dimensions of the interconnect and its location within the BEoL stack. The model provides the allowable current density levels as well a temperature distribution along the length of each interconnect segment. The insert graph shows the overlaid temperature of the test samples used in this study.

7.1 a) Schematic showing stages of CMP fill fabrication process with a newly deposited MET layer (top) a deposited dielectric layer for interconnect passivation (middle) and after CMP process provides a level surface for deposition of next layer (bottom). b) Snapshot of BEoL sample device showing microresistor element and contact leads surrounded by the square CMP fill patterns. An exclusion zone around the device is used for optical access to the device.
7.2  a) Device illustrations as seen from the top view. b) Schematic of the three-level BEOL stack showing the resistors devices with the contact and sensing leads (solid color), and the CMP fill regions (hatched color) at the three MET levels: Level 1 Poly (Left), Level 2 AlCu (Middle), Level 3 AlCu (Right).

7.3  a) Top view intensity map of a sample device under test (DUT). The center resistor is at level 1 and surrounded by the CMP fill square pattern. Two contact leads are seen at both resistor ends a central sensing lead is running transversely above the resistor. More CMP fill is also seen at the higher levels of the BEOL stack; b) Top view showing fill pattern unit cell dimensions and in-plane heat transfer path. c) Equivalent thermal resistance circuit model for calculating in-plane effective thermal conductivity. d) Side view of CMP fill pattern unit cell and heat conduction path e) Equivalent thermal resistance circuit model for through-plane thermal conduction. density $w_2/\sqrt{2d}$.

7.4  Simulation results of heat generation limit for an allowable temperature rise $(\Delta T)$ of 15 °C for interconnects at the three MET levels as function of device (resistor) width.

7.5  Schematics showing the design configurations of the microresistor BEOL devices: a) without the heat spreader and b) with the heat spreader. c) The actual device under test intensity image as obtained by the camera. The top row shows the narrow HS configuration while lower part shows the wide HS configuration. d) Side view showing the schematic of the HS features at the different levels of the BEOL e) Top view showing the staggered array structure of the Poly and AlCu lines beneath the resistor.

7.6  Thermal simulations of activated devices (Right) compared to experimental temperature maps (Left) of device with narrow heat sink (Top) and wide heat sink (Bottom) for a heating rate of 1.54 mW/µm.

7.7  Temperature rise profile showing cooling effect and the reduction in overall thermal resistance experienced by the 10 µm interconnect in Level 3 for device with a) narrow HS and b) with wide HS.

7.8  Cooling effectiveness of different widths of HS for different micro-resistor widths (Top) and the associated percent increase in current density $J_{RMS}$ allowance (Bottom).

8.1  The four fundamental passive circuit devices relating the four electrical parameters $(i, v, \phi, \text{and } q)$.

8.2  IV characteristics showing N-type and S-type Negative Differential Resistance (NDR) behavior.
8.3 Metallic filament forms at temperature exceeding the metal-insulator-transition (MIT) temperature (≈1000 K) [196] ................................................................. 142

8.4 Schematic showing the resistive memory behavior (re-RAM) of transition-metal oxides. ................................................................. 144

8.5 Activation synaptic behavior in memristors as compared to an actual neuron [118]. 145

8.6 Biologically inspired neural networks based on the functioning of the synapse [118]. ........................................................................ 146

8.7 Transmission electron microscopy images showing the metallic phase formed after device switching [140]. ............................................. 148

8.8 a) Device structure as seen under 20× magnification showing the four patterned top and bottom electrodes (TE and BE) and the resulting four cross-point devices. b) Intensity map of crosspoint device under 100× as seen with the camera under monochromatic illumination. c) Schematic of the crosspoint devices showing the Pt electrodes and the NbO_x thin-film. 150

8.9 Electroforming and filament distributions: (a) electroforming characteristics of a 5 µm Pt/Nb/NbO_{2.6}/Pt device, (b) electroforming characteristics of a 5 µm Pt/Nb/NbO_{1.92}/Pt device, (c) schematic representation of filament distribution in a 5 µm × 5 µm cross-point device with low-conductivity NbO_x (x = 2.60) films, (d) schematic representation of filament distribution in a 5 µm × 5 µm cross-point device with high-conductivity NbO_x (x = 1.92) films, (e) schematic of an oxygen vacancy filament formed by the field-induced generation, drift, and diffusion of oxygen vacancies, and (f) schematic of a transient current filament due to current bifurcation. Note that each dot in (c, d) represents a single device and the filament distributions represent 75 separate measurements for each film [176]. ........................................................................ 153

8.10 In situ thermoreflectance measurements: a) in situ (points) and ex situ (line) current-voltage characteristics of a 10 µm × 10 µ cross-point device with a high conductivity NbO_x (x=1.92) film, and the temperature of the filamentary and surrounding regions of the film (shown by circles in panel b), b) Two-dimensional temperature maps of the top electrode surface for different device currents. The inset shwos a back-scattered electron image of the filamentary region in the oxide film (i.e, after removing top electrode). ........................................................................ 155

8.11 Finite element simulatio: a) current-voltage characteristics during bidirectional current sweep, b) maximum temperature in oxide film as a function of device current, c) current density distribution before and after bifurcation, d) temperature distribution before and after bifurcation, e) before (point A) and f) after current bifurcation. ............................................. 156
8.12 Core-Shell model and current bifurcation: a) schematic of the CS model of a negative differential resistance device, b) electroforming characteristics of a $5\mu m \times 5\mu m$ device with NbO$_x$ (x=2.22 film), c) electroforming characteristics of a $10\mu m \times 10\mu m$ device with NbO$_x$ (x=2.22 film), d) in situ thermoreflectance maps of $5\mu m \times 5\mu m$ device during electroforming e) in situ thermoreflectance maps of $10\mu m \times 10\mu m$ device during electroforming.

8.13 Structure of the studied NDR crosspoint device showing snapback switching and saturation of filamentary conduction after current bifurcation.

8.14 I-V characteristics showing a) S-type NDR behavior and b) snapback (SB) NDR behavior. c) model of conductive paths in transition metal oxide (NbO$_x$) layer showing core and shell with representative electrical resistance. d) representative electrical circuit with constant current source, and the respective resistance of shell and variable resistance (memristor) of core [177].

8.15 In-situ temperature measurements of S-type and snap-back NDR. a) Current-voltage (I-V) characteristics and average temperature rise of the filament and surrounding area (the area used for averaging is indicated by boxes in b) as a function of applied current, b) 2D maps of the surface temperature rise in a $5 \mu m$ device operating at 1 mA and 6 mA, c) Temperature (current) localization of S-type NDR in the post-formed device at different current levels as shown in a, d) I-V characteristics and average temperature rise of the permanent filament and surrounding area (indicated by boxes in e) as a function of applied current, e) 2D map of the surface temperature rise in a $10 \mu m$ device at pre-threshold (4mA) and post-threshold (10 mA) currents, f) Temperature (current) localization in a post-formed device at different currents as shown in d (the blue arrow indicates snap-back transition), g) I-V characteristics and average temperature rise of the current filament and surrounding area (indicated by boxes in h) as a function of applied current, h) 2D map of the surface temperature in a $10 \mu m$ device at pre-threshold (4 mA) and post-threshold (12 mA) currents, i) Temperature (current) localization due to current bifurcation in the device without permanent filament as shown in g (the blue arrow indicates snap-back transition) [177].

8.16 (a-c) Asymmetric NDRs observed in Nb/NbO$_{x \cdot 99}$/Pt MOM under positive and negative polarity. (d) Matrix representation of dependencies of NDR on stoichiometry and device area.
8.17 a), b) In situ I-V characteristics of a $10 \times 10 \, \mu m^2$ Au (75 nm)/Nb(5 nm)/NbO1.99 (approximately 45 nm)/Pt (25 nm) device showing S-type NDR under positive bias and snapback under negative bias, respectively. Insets show the in situ thermoreflectance ($\Delta R/R$) maps of the $10 \times 10 \, \mu m^2$ device during corresponding bias polarity. c) Temperature profile through the filamentary region for different points in the current-voltage curve during positive bias. d) Temperature profile of the same device for different points in the current-voltage curve during negative bias. Note: the temperature map and the I-V characteristics shown in Figure 3 are also obtained for $5 \times 5 \, \mu m^2$ devices with lower threshold currents, and a snapback response observed under negative polarity and S-type NDR under positive polarity. But in this case, the filament is observed to form around the edges due to the dominance of edge effect in smaller area cross-point devices with a thick (75-nm Au in this case) capping layer [25]. The short-dashed curves in c) and d) represent the experimental data and the smooth lines show Gaussian fits of the corresponding experimental data. A, B, and C denote sub-threshold, threshold, and post-threshold points respectively in the I-V characteristic obtained under positive bias, while D, E, and F denote the same for negative bias polarity [178].

8.18 a) Schematic representing the Pump-and-Probe setup used to conduct transient-domain thermoreflectance conductivity measurements. b) TDTR normalized response showing the heating and temperature decay profile as measured at the gold surface of the blanket film samples.

8.19 a) The TDTR response is presented as a function of temperature for the thickest (100 nm) sample. b) Thermal conductivity as a function of temperature for different stoichiometric NbOx films.

9.1 a) Process flow of electrode fabrication using photolithography technique and b) image of interdigitated electrodes fabricated on glass substrate. The micro-fluidic channel is indicated with red dashed rectangle.

9.2 a) Schematic of the thermal imaging measurement setup used for measuring the temperature rise at the electrodes. b) In calibration, the electrodes are heated using a thermal chuck. c) In activation, the electrodes are self-heated in the presence of applied voltage via Joule heating effect.

9.3 Wavelength scan showing the thermoreflectance response ($\Delta R/R$) at different illumination wavelengths. Only the middle electrode is electrically activated using a constant current source and shows a positive reflectance change response for $\lambda <490$ nm, a negative response for $\lambda >490$ nm, and an optimal response at $\lambda=520$ nm. A 530 nm LED is used for higher intensity.
9.4 a) Impedance magnitude and b) phase angle of 1.4 S/m PBS solution in 400 µm micro-fluidic channel from 100 Hz to 100 MHz c) Experimental temperature rise at the electrode/electrolyte interface with electrode excitation at 20 V<sub>pp</sub> and frequencies varying from 10 kHz to 100 MHz. Maximum temperature rise at these conditions is obtained at 51.5 °C ± 5.5 °C (standard deviations were constant for all data points). 178

9.5 Comparison of experimental (symbols) and numerical (solid lines) temperature fields on the electrode regions a) for different applied voltages at 5MHz. b) Experimentally and numerically obtained average temperature rise on the side and middle electrodes as a function of the applied peak to peak voltage. c) Measured temperature field on chromium-gold surface at 20 V<sub>pp</sub> (Top view). d) Flow field below the electrodes (represented with yellow) bars is shown with streamlines and the velocity magnitude contours given in terms of µm/s. The simulation is performed at 5 MHz and 20 V<sub>pp</sub>. 180

9.6 Temperature rise at the electrode/electrolyte interface as a function of the dissipated power when the side electrodes are excited at 5 MHz and with different voltages. 182

9.7 Temperature rise as a function of applied voltages at 5 MHz for DI water showing negligible self-heating due to low Joule heating in electrically non-conductive fluidic medium. 183
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Source-driven conduction equation governing heating in micro-electronic devices.</td>
</tr>
<tr>
<td>3.2</td>
<td>Temperature-dependent Conductivity Parameter for generic Semiconductor compounds (Data compiled by ATLAS SILVACO).</td>
</tr>
<tr>
<td>3.3</td>
<td>Heat Capacity Ambient Value and Model Parameters (Data compiled by ATLAS SILVACO).</td>
</tr>
<tr>
<td>4.1</td>
<td>Activation conditions for the different GaN HEMTs. Source contact is grounded.</td>
</tr>
<tr>
<td>4.2</td>
<td>Gate (G), Source (S) and Drain (D) dimensions (in $\mu$m) of GaN HEMTs on different substrates.</td>
</tr>
<tr>
<td>4.3</td>
<td>Summary of the extracted thermal resistances (in mm.K/W) of the GaN HEMTs on Different Substrates based on average and peak temperature rise in the active channel region.</td>
</tr>
<tr>
<td>5.1</td>
<td>Gate (G), Source (S) and Drain (D) dimensions (in $\mu$m) of GaN HEMTs on different substrates.</td>
</tr>
<tr>
<td>5.2</td>
<td>Results of Converged Thermal Parameters of the Optimizations conducted for the GaN-on-Si HEMTs.</td>
</tr>
<tr>
<td>5.3</td>
<td>Results of Converged Thermal Parameters of the Optimizations conducted for the GaN-on-SiC HEMTs.</td>
</tr>
<tr>
<td>5.4</td>
<td>Results of Converged Thermal Parameters of the Optimizations conducted for the GaN-on-diamond HEMTs.</td>
</tr>
<tr>
<td>5.5</td>
<td>Comparing Substrate Conductivity and TBR for GaN on Si, SiC, and CVD diamond.</td>
</tr>
</tbody>
</table>
6.1 Temperature axial distribution along the resistor lead showing heating in the activated region and the cross coupling between the resistor and the sense lead. The three graphs show the single resistor configuration (Left), the active resistor with a cross coupled sense lead (Middle), and the active sense lead atop an inactive resistor (Right). Row (1) showing experimental thermal map, row (2) showing simulation thermal map, and row (3) showing a comparison between temperature distributions along the axial direction from the experimental data (Symbols) and the numerical data (Solid lines).

6.2 Comparison between temperature trace of simulation results (Dashed) and approximate model (Solid) for Resistor lead (Blue), Sense Lead (Green), and Contact Leads (Red). (1) shows the results for the initial model without global heating effects. (2) results with global heating (3) results with global heating and corrected cross model The three rows present the original model and two later modifications that are implemented based on presented observations and discussions. The first column (A) shows the case of an active resistor, the second column (B) shows the case of an active sense lead with an inactive resistor, while the third shows the case for active sense and resistor leads.

7.1 Percent reduction in self-heating temperature rise within the BEoL interconnects, and the percent increase in allowable current density ($J$) and heat flux ($Q'\text{'}) showing the cooling effect of the CMP Dummy Fill (*) at different levels of the BEoL.

7.2 Maximum Added Cooling and Increase in Allowable Activation Level for the different width Interconnects, as Compared to the Case with Narrow (N) HS Configuration.

8.1 Deposition conditions for different NbO$_x$ samples
I dedicate this dissertation
...to my mother who, with unwavering love, raised me through sleepless nights and
sacrificed much towards my success and wellbeing.
...to my siblings who were my number-one supporters althrough my academic career.
...to my friends in Lebanon that even while being continents apart, never lost contact and
had nothing but words of constant support and encouragement.
... to my colleagues and friends at SMU and in Dallas for making my five years at SMU
enjoyable and memorable.

In loving memory of my father Elias Assaad Helou (1956-2012), who raised me in good
faith and taught me love, selflessness, and humility.
1.1. Trends in Microelectronic Devices Constrict Thermal Constraints

The luring benefits of faster, more powerful and more efficient electronic systems has been driving the development in the microelectronic industry for the past few decades towards a systematic reduction in physical device scale. By allowing for shorter paths to electrical conduction, smaller features in device scales reduce electrical resistance thus cutting on electrical power losses, and reduce capacitance and switching times enabling higher operation frequencies (Figure 1.1) [248]. With advancement in the fabrication process accuracy and precision facilitating device miniaturization, and with the motivation of the prospective electrical benefits, the size of electronic active components have downscaled exponentially, as observed by Moore’s law.

Motivation by the above electrical benefits has lead to successively more compact microelectronic technologies and increasing power density levels. Moreover, when incorporated within increasingly insulating materials, both electrically and thermally, rates of heat generation densities have exceeded stellar rates (Figure 1.2) with a concerning increase in thermal hotspot generation and in device operating temperatures.

Microelectronic devices, especially those operating in critical and harsh environments, are designed to meet strict standards in performance and reliability. Under the aforementioned arising conditions, thermal design became a limiting factor that is integral in the design process and could no longer be ignored and excluded until the last design stages, due to many damaging consequences of elevated operating temperatures on device performance and reliability. To begin with, at elevated temperature, mechanical and thermophysical properties can change drastically and deviate from design tolerances, critically affecting the perfor-
Figure 1.1. a) Logic technology node and gate length b) resistance and c) capacitance of different technologies showing the improvement in performance with device scaling. (Adapted from [248])
Figure 1.2. Graph showing the increased heat flux densities in advance microelectronic devices (represented by bars) resulting from increasing power levels and downscaling in device sizes. Red circles show comparison to other natural and man-made heat generation rates for scale. (Source: DARPA)

mance and energy efficiency of the device and causing unpredictable electronic behavior. Moreover, devices are built to certain lifetime expectancies where a certain number of active devices under normal or stressed conditions are expected to fail. Elevated temperatures due to excessive self-heating can bring about dramatic failures or accelerate degradation rate in active devices, which inevitably shortens the lifetimes of microelectronic devices [89] or limits activation levels due to reliability concerns especially for mission-critical devices (satellites, antenna and wireless communications, lidar applications and autonomous systems, ...).

1.2. Temperature Effects on Reliability and Performance

Technological development in microelectronics is being driven towards device miniaturization in an effort to harness the potential improvements in electrical efficiency, activation frequencies and production yields. As devices scales shrink, self-heating effects becomes more pronounced, especially for high power and analog devices. Knowing that temperature is an
influencing factor in almost all governing electrical, thermodynamic, and solid-state physics, excessive heating affects device performance and reliability in a variety of ways. Temperature can either affect design considerations, alter device performance, and/or bring upon failure and accelerate device degradation. In other applications, heating is a bi-product to electrical activation and can either be beneficial or detrimental to the operation depending on the technology. A few applications are presented in this work where temperature can pose as a curse or a blessing, and in both cases, the importance of accurate and adequate thermal investigation is stressed due to the many effects temperature have on complex microelectronic system.

1.2.1. Failure and Degradation in High Power Devices

The activation temperature rise has been observed to affect reliability and performance across a wide spectrum of electronic devices and structures, each with its specific challenges and thermal characterization and management strategies. At the front-end (component level), elevated temperatures can cause device operation to deviate from design expectations [40, 64], and can compromise device reliability and lifetime [45]. Therefore, thermal concerns are most critical when devices are to operate under high-power high-temperature applications, and when the devices are required to operate reliably with minimal failure outcome. These requirements, when considering the afore-mentioned thermal challenges, often require some form of thermal management to minimize self-heating effects. As an example, high power Gallium Nitride (GaN) high electron-mobility transistors, used in high-frequency and high-voltage switching applications, are limited in operation due to self heating effects that cause sudden failures and accelerated degradation [168, 200]. GaN technology have therefore required the development of innovative thermal management strategies such as high conductivity substrates and heat spreading materials [11, 239, 256] and micro-channel cooling [44].

Another region of thermal concern is the back-end (interconnect level) of integrated circuits (ICs), where elevated temperatures in high-current carrying conductors can accelerate
electromigration [222, 246]. This form of degradation in the interconnect network, which delivers power to front-end devices, can eventually cause the performance to drop below required standards and can also lead to short circuit and open circuit failures [40, 111, 270].

1.2.2. Strict Temperature Control for Micro-Sensing Devices

Other applications in microelectronics may not require high-power activation, but rather robust and accurate operation which may be sensitive to device temperature, and thus require precise thermal control. This is the case for micro-sensing applications where any temperature-induced fluctuations in device performance such as piezo-sensors [62, 84], chemical or bio-sensors [108, 219], if not characterized and accounted for accordingly would lead to erroneous readings and faulty operation. Two applications are presented in this documentation. GaN HEMTs with an etched Silicon substrate leaves a suspended GaN membrane which have been observed to be sensitive to external stimuli. Therefore, GaN HEMT membranes have found several micro-sensing applications [62, 121, 243], but are thermally challenged by an excessive self-heating rate due to the absence of a substrate as a medium for heat dissipation.

The second presented micro-sensing temperature study is for a lab-on-chip application, where micro-fluidic manipulation results in significant self-heating that must remain within a certain limit to insure that bio-fluidic samples (blood cells, bacteria...) are not damaged or denatured [129, 219].

1.2.3. Joule-Heating Induced Electrical Characteristics

While thermal effects are detrimental to some device technologies, in other cases, the thermal response is a part the governing mechanisms behind some distinct electrical characteristics, such as negative-differential resistance (NDR) and snap-back switching characteristics observed in memristive devices (i.e., memristors or memory resistors) [63, 234]. These electrical characteristics have thus been the topic of development in several applications such as neuromorphic computing [86, 118, 259] and non-boolean non-volatile storage (Resistive
Random Access Memory ReRAM) systems \[85, 224]. For that reason, accurate characterization of the thermal response of such devices are essential to understand and model their electrical performance, bring about further developments in the technology, and eventually aid its integration in system applications.

1.3. Outline

Present developments in microelectronic technologies accompanied with emerging thermal management challenges have demanded advanced thermal characterization techniques, both experimentally and numerically, in order to assess the thermal effectiveness of device structures and thermal management strategies. Experimental techniques, mainly implemented for thermal mapping, is required to identify hot spots, investigate the thermal performance of structure designs and characterize their heat dissipation effectiveness. These measurement techniques are developed with different specifications, and in this work, we will present the advanced experimental capabilities relevant to microelectronic devices that are able to provide reliable high-resolution surface temperature maps with good accuracy and a suitable device accessibility.

On the other hand, numerical simulations are a necessity to apply the governing physics in microdevices to either predict self-heating rates in a micro-device, or to understand the effect of different parameters on the thermal response, and thus guide thermal designs and developments. Similarly, presented in this work are the different modeling challenges faced in thermal problems of microelectronics and advanced modeling capabilities for representing the thermal response in microelectronics devices and structures.

Thermal heating effects are an inevitable bi-product of electrically active micro-devices. However, their effect on device reliability and performance may vary in criticality or benefit in each particular application. While being a minor inconvenience in many applications, the thermal aspect could ultimately be a determining factor in some microelectronic applications. The common consent is that, to ensure reliable and accurate operation, the thermal perspective is becoming a more determinant design factor and cannot be exempt from, or
kept until the last stage of the microelectronic design process.

The next two chapters (Chp. 2 and 3) will present the advanced thermal experimental and numerical analysis techniques used in this work to meet the distinct challenges and requirements of each of the studied device technologies, which are presented in the remaining part of this work. Chapters 4 and 5 present a novel approach to fully characterize high-power GaN HEMT transistors through a combined experimental and numerical approach. The characterization method is applied to investigate the self-heating and cooling effectiveness of different substrate configurations, and extract important thermal properties to guide future developments in device fabrication. Chapter 6 presents the investigation of the thermal characteristics of structures within the back-end of line (BEOl). Available models are assessed and a faster, more-accurate reduced model is developed to complement reliability model and provide a fast thermal check on IC interconnect networks. Chapter 7 builds upon the developed understanding of the thermal response within the BEOl to develop and optimize a heat spreader design to provide passive cooling to critical interconnects in the BEOl. Two additional device technologies are presented in Chapter 8, in which the thermal response of memristive devices are investigated to understand the thermally-governed distinct electronic behavior, and in the last Chapter 9, which investigates the self-heating of lab-on-chip electrodes in AC electro-thermal flow used in micro-fluidic sensing applications. In each of the technologies presented, we discuss the motivation behind thermal management and temperature control, as well as the unique characterization challenges that are addressed. We finally present the main outcomes and expanded understanding of the technologies that was attainable by our advance thermal characterization techniques.
Chapter 2
Experimental Thermal Characterization

Thermal criteria have become more stringent with continued downscaling of microelectronic devices, further hindering performance due limit overheating that causes premature failures and accelerated degradation. As a result, the thermal aspect has become an integral part of the design of microelectronic devices, especially for those operating under high power densities or temperature-sensitive applications. For any thermal investigation of a device structure or a suggested thermal management strategy, experimental techniques are a primary requirement to observe and evaluate the actual response of the test device. These observations of the real world are what would also validate developed simulation models, and help build an understanding around the governing physics to better predict device performance.

Within thermal investigations, temperature measurements serve as the instrumentation basis from which other thermal parameters and properties can be deducted, such as heat flux, thermal conductivity and thermal diffusivity. For that reason, the temperature measurement techniques are the corner-stone of experimental thermal investigation. Temperature instrumentation methods are naturally indirect and inferred by measuring secondary phenomena, such as but no limited to: volume expansion (thermometer), radiation emission (infrared), electrical resistance (thermistor or RTDs), or a developed potential difference (thermocouple). Given the many instrumentation methods, the selected technique should be tailored and suited to the application. For the case of microelectronic devices, exceedingly more advanced techniques are required to meet stringent requirements of resolution, material compatibility, and adequate device accessibility.

In this Chapter, we present some of the measurements techniques most suitable for microelectronic device thermal imaging applications, while discussing some of their benefits
and limitations. In the second Section, we present the experimental technique that is used for the investigations presented in this report, starting with the underlying instrumentation physics, the measurement setup, and finally the acquisition methodology. The challenges facing the experimental approach for each of the specific applications will be presented in the respective upcoming Chapters.

2.1. Thermal Microscopy Techniques: Capabilities and Shortcomings

2.1.1. Infrared Imaging

The most commonly used thermal imaging technique is the infrared thermometry (IR). Pyrometers also known as radiation thermometers operate on the basis of measuring the amount of infrared radiation emitted from a body. The temperature is then inferred by Planck’s law. [60] This method the most commercially popular since it is most suitable for acquiring non-contact temperature maps at the macro-scale and provides quick thermal maps at a relatively moderate cost.

The major setbacks of IR imaging techniques is that IR-based cameras cannot detect spatial resolutions below a few microns due to the diffraction limit of IR radiation. The diffraction limit (D) of optical imaging is determined by the wavelength of illumination (\( \lambda \)) and numerical aperture (\( NA \)) as\( D = \frac{2\lambda}{NA} \). For IR, the resolution is diffraction limited to between \( 3-10 \mu m \) depending on NA, which is insufficient to detect temperature variations in microscale devices. Moreover, IR emission is proportional to the materials surface emissivity \( \epsilon \), which is lacking for reflective materials present in microelectronic devices rendering them inefficient at emitting IR radiation. As a result, IR thermometers often require that the sample devices be coated with a thin “target” of black paint to enhance the emissivity of the surface. This is not desirable since it may alter the thermal properties and response of the sample under test and induce additional systematic measurement errors.
2.1.2. Micro-TCs and SThM

To achieve temperature measurements at the microscale, thermal engineers resorted to micro-thermocouples, either embedded in the structure of the device or mounted on a tip that engages contact at the surface of the device. These methods utilize contact methods to obtain single point measurements. Mapping a surface for a temperature distribution further requires additional translation stages for navigating the tip across the measurement area. Although the temperature detected is accurate to 0.01 K, the spacial resolution is limited to the size of the thermocouple tip (25-50 microns) [60].

Scanning thermal microscopy (SThM) is another technique that can improve on the spacial resolution by using nanometer size thermocouple fabricated at the tips of an atomic-force microscopy (AFM) probe. The method is highly accurate but proves to be expensive and not commercially available. The measurement is also time consuming and, being a contact method, could introduce additional heat conduction paths and result in systematic errors. Moreover, systematic error can be introduced for light emitting devices which may radiate and result in heating of the measurement probes.

2.1.3. Micro-Raman

To obtain complete thermal maps of the region of interest (ROI), non-contact methods are preferred over contact methods since the temperature map is obtained remotely without interfering with the device’s thermal system. An optical method used readily in sub-micron thermal imaging is micro-Raman spectroscopy. The method uses visible light illumination and can thus reach spatial resolutions of less than 1 μm. The basis of operation is that the surface material of the device under test scatters illumination in-elastically thus shifting the wavelength and intensity. This shift is temperature dependent and can be used to infer a device’s surface temperature.

The physics of Raman instrumentation is highly complex since it requires the decoupling of the effects of temperature, mechanical stress, and electric field on the measured Raman shift [14]. The Micro-Raman method is also inapplicable for materials with no phonon lines in
their spectrum, like metals and packaging plastics [220], and is most applicable for III-Nitride materials that have good Raman scattering effects. Since the method uses high intensity laser, the measurement remains a single point method and often required XY translation stage, adding to the complexity and duration of acquisition, and usually resulting in low resolution imaging (Figure 2.1) [105].

Figure 2.1. Thermal maps of active GaN HEMT using a) IR and b) micro-Raman imaging from Sarua et al. [220]. c) Device measurement region and Raman measurements adapted from Hiroki et al. [105]. The high-resolution requirements are evident especially as device dimensions near the sub-micron scale.
2.2. Thermoreflectance based Thermal Imaging

The thermal metrology technique chosen for the studies presented in this work is based on the thermoreflectance (TR) effect, which provides an optical measurement method that is non-invasive, non-contact and non-destructive [207–209, 211]. The physical basis of the measurement relies on the temperature-varying optical properties of semiconductor materials, and is mainly acquired through the reflectance change of a surface material subject to a rise in temperature, under visible light illumination. By detecting the changes in reflected light intensity ($\Delta R/R$) from a surface, the changes in temperature ($\Delta T$) can be inferred after the R-T dependence of the test sample is known before-hand through calibration.

In addition, the method is more suited than IR for thermal mapping of microelectronic devices, since most of the semiconductor and metallic materials constituents are of low emissivity $\epsilon$ and high reflectivity $R$. Moreover, by measuring with a shorter wavelength illumination, the TR imaging can theoretically reach down to a diffraction-limited spatial resolution of around 0.3 to 0.5 $\mu$m ($\approx 10 – 50 \times$ that of IR) [61].

Initially, the method was implemented as a single-point measurement setup, and later for surface scanning with the use of a translation stage. Made possible by the advancements in camera technology, the TR method is currently implemented as a surface temperature imaging technique, cutting acquisition times to a fraction of scanning single-point measurements methods. Current camera-based TR imaging systems can detect temperature-induced surface reflectance changes with relatively low noise levels allowing for a temperature resolution reaching as low as 0.1 K [61].

The TR response is also implemented in pump-and-probe techniques for measuring thermal properties of thin-film material stacks, such as the time dependent thermoreflectance (TDTR) measurement technique. In TDTR, the surface temperature response is monitored while a laser excites a transducer gold surface atop a blanket film stack. The energy absorption and the temperature decay through the stack is later analyzed to infer the thermal properties of the constituting materials [51, 57, 268]. The experimental details will be presented in a following section while the reverse modeling equations to extract thin-film
properties will be presented with the numerical techniques in Chapter 3.

2.2.1. Governing Physics of Reflectance-Temperature Dependence

The surface reflection of light is a complex interaction between incident electromagnetic wave and the energetic structure of the material. The electromagnetic wave can be either absorbed, transmitted or reflected, with the proportion of each depending on the optical properties of the surface, determined by the energy of the lattice, and the energy of the incident photons. The portion of light reflected from a surface is related to the complex refractive index \( N = n + ik \) as presented by the Fresnel equations (2.1).

\[
R = \left| \frac{1 - N}{1 + N} \right| = \frac{(1 - n)^2 + k^2}{(1 + n)^2 + k^2} \tag{2.1}
\]

From the above expression, the reflectance change \( \Delta R \) due to any change in the refractive indexes can be expressed by differentiating (2.1).

\[
\frac{\Delta R}{R} = \left( \frac{4N}{N^2 - 1} \right) \frac{\Delta N}{N} \tag{2.2}
\]

The refractive indexes can be determined from the complex dielectric function \( \epsilon = \epsilon_1 + i\epsilon_2 \)

\[
n = \left( \frac{(\epsilon_1^2 + \epsilon_2^2)^2 + \epsilon_1^2}{2} \right)^2 \tag{2.3a}
\]

\[
k = \left( \frac{(\epsilon_1^2 + \epsilon_2^2)^2 - \epsilon_1^2}{2} \right)^2 \tag{2.3b}
\]

Consequently, the reflectance of a surface can be expressed in terms of the change in the real and complex part of the dielectric function by combining (2.1) and (2.3) and differentiating with respect to \( \epsilon_1 \) and \( \epsilon_2 \). \[225\]

\[
\frac{\Delta R}{R} = \frac{1}{R} \frac{\partial R}{\partial \epsilon_1} \Delta \epsilon_1 + \frac{1}{R} \frac{\partial R}{\partial \epsilon_2} \Delta \epsilon_2 \tag{2.4}
\]

The partial derivatives expressed in terms of \( \epsilon_1, \epsilon_2 \), and \( \epsilon = \sqrt{\epsilon_1^2 + \epsilon_2^2} \) are shown in (2.5)
\[
\frac{1}{R} \frac{\partial R}{\partial \epsilon_1} = \frac{\sqrt{2(\epsilon_1 + \epsilon)}}{\epsilon((\epsilon_1 - 1)^2 + \epsilon_2^2)}(2\epsilon_1 - \epsilon - 1)
\]

\[
\frac{1}{R} \frac{\partial R}{\partial \epsilon_2} = \frac{\sqrt{2}}{\epsilon\sqrt{\epsilon_1 + \epsilon}}(2\epsilon_1 + \epsilon - 1)
\]

The above relation has allowed the complex dielectric function \( \epsilon = \epsilon_1 + i\epsilon_2 \) to be calculated from modulated reflectance measurements in spectroscopic ellipsometry. The dielectric function determines the energy structure of metals and semiconductors. The optical reflectance is dependent on the structure of the energy band and on the process involved in its interaction with incident electromagnetic waves. For intraband processes, the free carriers are the major contributors in defining the interaction, which occur mostly in metals and conductors. For interband processes, the optical properties are defined by electrons at low energy levels absorbing the incident radiation and transitioning to higher energy levels, which is the case for semiconductors [78].

Temperature change has many complex effects on the optical properties of a certain material and thus changes the complex dielectric function and the energy structure of solids through different physical mechanisms [215].

1. Volumetric expansion causes the electron density to change and the Fermi level to shift
2. Shear stress resulting from sample constraint in the substrate can also cause shifts and warps in energy bands
3. Broadening electronic distributions around the Fermi level
4. Phonon population increase causes less electron relaxation times and shifts and warps due to electron-phonon interaction
5. Temperature changes causes small increase in the Fermi level
6. If the temperature is self induced by Joule’s effect, current flowing displaces the distribution function
In a simplified approach the change of the dielectric function can be caused by a change in temperature and is related in solid state physics by a shift in the band gap energy $E_g$ and the broadening parameter $\Gamma$ in the case of semiconductors, and to the DC electrical conductivity in the case of metals. However at large temperature changes the above mentioned effects become more significant and result in nonlinear variation in the $C_{TR}$ coefficient and affect the accuracy of thermoreflectance microscopy techniques [78]. For thermoreflectance experimental studies, and unlike theoretical approaches, the overall combined effect of temperature on the reflectance is acquired by calibration at a given light wavelength (phonon energy), without the need for full photon energy spectrum analysis.

2.3. TR Imaging Acquisition Methodology

2.3.1. TR Experimental Setup for Thermal Microscopy

Thermal imaging experimental studies presented in this work are conducted using the T°Imager® system which is commercially available [207]. The setup schematic consists of an optical setup that directs an illumination source (LED or monochromatic light) to a sample which is vacuum-held in place on a thermally controlled wafer chuck (Figure 2.2). The stage is also mounted on an an XYZ nanopositioner to be able to reposition the DUT and counteract thermal motion. The incident light is first sampled through a beam splitter where a photodetector monitors the changes in the illumination source intensity for data corrections. The light then passes through an objective lens (5× to 100× magnification) and reflects from the sample surface back to a camera that monitors the reflected light intensity.

Electrical connections to the device are established using the Cascade Microtech nanopositioners which can provide 2-probe and 4-probe activation. The instrumentation setup is controlled through the T°Imager user interface which is used for both the thermal imaging both in activation and calibration of the device (Figure 2.3).
Figure 2.2. Schematic showing optical setup of TR imaging imaging system and the incident and reflected light path from the sample.
Figure 2.3. TMX T°Imager® setup showing activation, temperature, PC, nano-positioning, and light control used for sample thermal imaging.
2.3.2. TR Measurement Procedure

The measurement acquisition of the device’s thermal map is performed in a two-stage procedure. Initially, the device is electrically activated, where self-heating results in a temperature distribution to be determined. The reflectance change resulting from that temperature rise is mapped using the camera in the activation step. Next, the reflection-temperature relation of the device surface features are determined using a calibration step in which the reflectance change is acquired after a known set of uniform temperature rises are applied.

2.3.2.1. Activation Step for $(\Delta R/R)_a$ Map

The reflectance of a surface has been demonstrated to be weakly sensitive to temperature changes where the coefficient of thermoreflectance lies in the range of $10^{-5}$ to $10^{-4}$ per Kelvin, which proves the detection of such changes to be challenging [27]. To be able to acquire a sufficient signal to noise ratio, the surface is required to be reflective, and to be sufficiently thermally reflective. For most metallic surfaces used in microelectronic fabrication like Gold and Aluminum, the reflective surface requirement is satisfied where a sufficient amount of light is reflected back to the light sensing equipment. However, for a surface to be thermally reflective, it must exhibit a measurable reflectance changes with temperature under some specific wavelength of illumination. Incidentally, different photon energies interact differently with the surface varying the thermal reflectance response for different wavelengths of incident light. Thus by measuring the TR response subject to a wavelength scan, the optimal illumination can be determined to ensure the best TR signal prior to activation.

In activation, and upon determining the optimal wavelength for TR measurements, the sample is electrically activated to the required power levels using a Source/Monitor Unit while measuring the surface reflectance change $(\Delta R/R)$ maps between the ON and OFF states of the device (i.e. hot and cold frames). The reflectance change maps are recorded and will be used in conjunction with the calibration maps to calculate the temperature rise distribution.
2.3.2.2. Calibration Step for $C_{TR}$ Map

The relation between the reflected light intensity and the surface temperature has a dominant first order dependence as shown in equation 2.6. For small temperature variations (O(10 K)), the relation is assumed linear. Weak non-linear effects may occur but are more evident for larger temperature variations (O(100 K)).

$$R(T) = R(T_0) + \frac{dR}{dT}(T - T_0) + \frac{1}{2} \frac{d^2R}{dT^2}(T - T_0)^2 + ...$$  \hspace{1cm} (2.6)

$$R(T) = R(T_0)(1 + C_{TR}(T - T_0))$$  \hspace{1cm} (2.7)

The linear parameter $C_{TR} = \frac{1}{dT} \frac{dR}{R}$ (units $/\text{Kelvin}$) is often referred to as the coefficient of thermo-reflectance (similar in form to the coefficient of thermal expansion relating strain to temperature $C_{TE} = \frac{1}{dT} \frac{dL}{L}$). The thermo-reflectance coefficient naturally depends on the surface material composition and its roughness, and on the light’s wavelength and its optical path. Thus any transparent passivation also adds interference effects and varies the TR response. For that reason, the coefficient is indeterminate except for few cases of bare metals, like Au, Al, Pt and Ni [131]. For an arbitrary device, with various surface material and features, the coefficient of thermo-reflectance is difficult to determine analytically, and thus must be acquired by in-situ calibration where the surface reflectance change is measured as a known temperature change is applied. As complex microelectronic devices contain different material and features on its surface, the $C_{TR}$ map is also obtained for each pixel independently.

In calibration, a known temperature rise $\Delta T$ is applied to the sample device through the Peltier heating stage, and the reflectance change $(\Delta R/R)$ from the sample surface is acquired using the camera and used to calculate the thermoreflectance coefficient map $C_{TR}(i,j)$ as given in (2.8). Realignments of the DUT are performed between the hot and cold set temperatures after thermal motion (X and Y directions) and to refocus (Z direction). This ensures that the measured location refers to the same position on the actual device.
2.3.2.3. Extracting Thermal Maps from $C_{TR}$ and $(\Delta R/R)_a$

The temperature rise distribution ($\Delta T$) is finally extracted by combining the reflectance maps obtained from activation to the calibration maps of $C_{TR}$ according to (2.9).

$$C_{TR}(i,j) = \frac{1}{\Delta T_{1,2}} \frac{\Delta R_c(i,j)}{R_c(i,j)}$$ (2.8)

$$\Delta T_{op}(i,j) = T_{op}(i,j) - T_{base} = C_{TR}(i,j) \times \frac{\Delta R_a(i,j)}{R_a(i,j)}$$ (2.9)

Subscript $c$ and $a$ refer to the calibration and activation process respectively, $i$ and $j$ refer to the pixel location index within the field of view. $\Delta T_{1,2}$ is the set temperature during calibration, while operating temperature rise $\Delta T_{op}$ is obtained during activation relative to the fixed chuck temperature $T_{base}$.

2.4. Challenges and Shortcomings of TR imaging

Thermoreflectance imaging is a powerful method for thermal characterization but has some limitations never-the-less. Initially, since the optical method is based on reflectance, the measurements are limited to the surface of exposed regions and to the topmost opaque/reflective layer in a stack. The surface limitation therefore introduces an accessibility challenge where embedded features cannot be directly reached for measuring their temperature. This limits the TR measurement technique especially when recent developments require more invasive techniques, as more technologies move to 2.5D and 3D chip-stacking, and field plates are introduced that cover active regions in high power transistors. This limitation is discussed further in Chapters 4 and 5 and motivates a reverse modeling approach to help estimate the temperature within embedded features.

Another setback for the TR technique is that some materials under-test are transparent to visible light and cannot be measured directly (such as oxides, Gallium Nitride and Silicon Carbide). Some of these material are opaque for shorter wavelengths of light that exceed their bandgap. For instance, Gallium Nitride (GaN) and Gallium oxide (GaO) can reflect near-ultra-violet (UV) illumination (350 nm and 250 nm respectively) and their surfaces can
be measured accordingly. Chapter 4 presents a study where GaN based power devices are thermally mapped using near-UV (above-bandgap) illumination to overcome this limitation.

An optical imaging system’s resolution is limited by the principle limit of light diffraction that determines the smallest distance between features that can be resolved. The diffraction limit of light \( d \) is a function of the wavelength \( \lambda \) and the optical setup, mainly the numerical aperture \( N_A \) as \( d = \lambda / 2N_A \). Most of the available thermoreflectance apparatus use objective lenses with \( N_A \approx 0.5 \), thus the diffraction limit imposed resolution is around 0.3 \( \mu \text{m} \) to 0.6 \( \mu \text{m} \) at best. This would render measurements challenging for features in the sub-micron scales not considering other challenges with instrumentation at that magnification level.

Other concerns arise regarding accuracy especially when measuring large temperature variations where non-linear effects can cause deviations from the linear TR region. Calibration for a constant \( C_{TR} \) field is obtained at set temperatures changes at the order of 10 K and is accurate for measuring temperature rises in the same range. At temperatures on the order of 100 K range, the coefficient may exhibit non-linear variations [100] that, if ignored, can lead to measurement errors. This issue is be addressed by using high temperature calibration in Chapter 4 to more accurately represent the reflectance-temperature relation at elevated temperatures.

Thermoreflectance imaging, even with some of its limitations, remains a versatile and advantageous instrumentation method that provides quick, high resolution, thermal maps. Therefore, thermoreflectance imaging will be the chosen experimental method for the studies presented in this report, with the mentioned challenges discussed in more detail for each of the specific device technologies, and the approaches for addressing those challenges accordingly detailed in the respective Chapters.

2.5. Thermal Property Measurements using TDTR

The property measurements that are presented in this work were conducted with the TMX Scientific Transometer\textsuperscript{TM} TDTR system present in the NETS laboratory at SMU. The system uses the pump and probe technique where a pulsed Nd:YAG (532 nm) is used
Figure 2.4. Schematic representation of pump-and-probe technique for the time-domain thermoreflectance.

as a pump laser to thermally excite the gold transducer layer, and a continuous wave (CW) Argon-ion laser (488 nm) serves as the probing laser (Figure 2.4). To collect sufficient data for a proper statistical analysis, five sets of a minimum of 500 pulsed excitations are performed for each thickness sample. The total thermal resistance is extracted from parameter fitting for each of the different thickness samples. The model theory and parameter fitting are presented in the next Chapter.
Chapter 3
Numerical Thermal Characterization

In any scientific study, experimental investigations provide the actual data relating to real life observations that ground an analysis to the truth and characteristics of a system or device under test. These observations alone, however, would not provide a full investigation and must be complemented by a set of numerical studies which would elevate the acquired observations and help build an understanding and a set of governing relations (i.e., models). Also the case for thermal investigations, numerical analysis is an invaluable tool that also gives an understanding of the effects of different material and geometric properties, enables to predict a device’s electrical and thermal response, and facilitates device design and integration. Moreover, numerical studies provide a fast and cost effective digital/artificial platform for representing a thermal or electrical system, and can be used in design cycles (what-if scenarios) for evaluating a range of different parameters and reporting metrics other than temperature (heat flux, conductance, thermal resistance). Therefore a numerical investigation would reduce the needed number of prototypes and experiments, and by that shortening the design cycle and reducing time and cost.

The benefit of numerical techniques in semiconductor and microelectronic thermal investigations goes beyond the above general benefits and provides additional advantages. Given the device scale and complexity, experimental techniques may be limited in resolution and accessibility (surface measurements, 2D), and thus numerical techniques would help fill in the access gap by using estimates of temperatures (in 3D) based on physical models. Moreover, experimental techniques would report only temperature readings and thus a numerical investigation is needed to calculate and extract important thermal properties and parameters such as thermal conductance and interface resistances.
In the first Section of this Chapter, we present the theoretical background for solving the heat conduction equation in microelectronics, and the governing electro-thermal interactions that result in heat generation (i.e., self-heating) in microelectronics. In the following section, we present the numerical simulation tools and techniques used in this dissertation, from forward modeling, reduced mathematical models and experimentally-driven model optimization (reverse modeling) and discuss their advantages to generic and commercial numerical techniques in dealing with microelectronic devices and structures.

3.1. Full Electro-Thermal Modeling

Microelectronic device operation is a result of complex solid-state physics and coupled electro-thermal interactions that are subject to strong dependencies on operation temperatures. To begin, electrical properties are temperature depend and therefore, any variation in ambient device temperature can cause it to deviate from its designed electrical performance (impedance change, distortion, shifts in frequency response). Moreover the electro-thermal relation is a two way coupled system where the electrical activation also determines the rate of heat generation within a device which then results in the temperature gradients within the structure. These physical interactions increase the complexity of an electro-thermal model and may require direct or indirect coupling of the electrical and thermal models, and by that increase the numerical complexity of a full-scale numerical model.

3.1.1. Electrical Modeling Equations for Semiconductor Operation

The main governing equation for electrical conduction in a semiconductor device is the Poisson equation of electrostatics that relates the electric potential $\phi$ to the charge density $\rho$ through the electric permittivity of the medium $\epsilon$.

$$\nabla \cdot (\epsilon \nabla \phi) = -\rho$$  \hspace{1cm} (3.1)

The charge density in the above equation represents the different charge carriers including electrons, holes, donors and trap charges.
\[ \nabla \cdot (\epsilon \nabla \psi) = -q(p - n + N_D^+ - N_A^-) - \rho_S \]  

(3.2)

Where \( N_D^+ \) and \( N_A^- \) are the doping concentrations, and for constant permittivity \( \epsilon \), the leading term becomes the Poisson relation \( (\epsilon \nabla^2 \psi) \).

Moreover, the **Carrier Continuity Equations** ensure the conservation of the charge carriers contributing to the current flow.

\[ \frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot \vec{J}_n + G - R \]  

(3.3)

\[ \frac{\partial p}{\partial t} = -\frac{1}{q} \nabla \cdot \vec{J}_p + G - R \]  

(3.4)

Where \( n \) is the charge concentration, \( q \) charge unit, \( J \) current density, \( G \) and \( R \) generation and recombination rates. The two continuity equations 3.3 and 3.4 and the Poisson equation 3.2 are solved for \( n, p, \) and \( \psi \). However, the equation terms should be modeled in terms of the independent parameters and the physical constants.

The current density can then be represented as function of the carrier concentrations \( n \) and \( p \) using the **Boltmann Transport Theory**.

\[ \vec{J}_n = -q\mu_n n \vec{\nabla} \phi_n \]  

(3.5)

\[ \vec{J}_p = -q\mu_p p \vec{\nabla} \phi_p \]

Where \( \phi \) is the quasi-fermi potential for electrons and holes. To combining the continuity equations with the poisson’s equation for electrostatics, the **Drift-Diffusion Model** is used to represent the current density in terms of carrier concentrations and electrostatic potential as shown in equation 3.7. The two terms are the drift (electric field induced) and diffusion components (concentration induced) of the electric current.
\[ \vec{J}_n = nq\mu_n \vec{E}_n + qD_n \vec{\nabla}n \] (3.6)
\[ \vec{J}_p = pq\mu_p \vec{E}_p + qD_p \vec{\nabla}p \]

Where \( \mu \) and \( D \) are mobility and diffusivity of the charge carriers (electrons and holes).

Under steady state conditions, the charge concentration rate is balanced by the generation and recombination rates, giving the current density balance:

\[ \nabla \cdot \vec{J} = 0 \] (3.7)

Writing the equation in terms of the potential and charge concentration, and using Boltzmann statistics for charge densities, we can rewrite the continuity equation as:

\[ \nabla \cdot \{qD_n \nabla n - nq\mu \nabla \psi - \mu n[k_B T \nabla \ln(n_i)]\} = 0 \] (3.8)
\[ \nabla \cdot \{qD_p \nabla p - pq\mu \nabla \psi - \mu p[k_B T \nabla \ln(n_i)]\} = 0 \] (3.9)

The above equation is solved for \( \psi, n \) and \( p \), from which the current density can be calculated using the drift-diffusion equation, and the electric field obtained from the definition \( \vec{E} = -\nabla \psi \). The remaining complexity then spans from the modeling of the individual parameters like the diffusivity and mobility of charge carriers.

The diffusivity \( D \) of the charge carriers could be related to the mobility using the Einstein Relationship

\[ D_n = \frac{kT}{q} \mu_n \] (3.10)
\[ D_p = \frac{kT}{q} \mu_p \]
The above relation is a simplification where more accurate representations require using Fermi-Dirac statistics. Moreover, there exists numerous models for the charge carrier mobility with the simplest being a constant mobility model.

The Fermi-Dirac function determines the probability of a charge to occupy an energy level above the conduction (for electrons) or below the valence (for holes).

\[
f(E) = \frac{1}{1 + e^{(E-E_f)/kT}} \approx e^{(E-E_f)/kT} \tag{3.11}
\]

\[
n_0 = f(E_c)N_c = N_ce^{-(E_c-E_f)/kT} \tag{3.12}
\]

\[
p_0 = (1 - f(E_v))N_v = N_ve^{-(E_f-E_v)/kT}
\]

where the effective density of states are given as

\[
N_c = 2 \left( \frac{2\pi m^* n kT}{\hbar^2} \right)^{3/2} \tag{3.13}
\]

\[
N_v = 2 \left( \frac{2\pi m^* p kT}{\hbar^2} \right)^{3/2}
\]

As seen in the above presented equations, the different electrical metrics, from charge distribution, diffusivity, electrical conductivity as well as the governing equations are dependent on the thermal energy \((kT)\), with \(k\) being the Boltzmann constant, and in-turn dependent on the device temperature. As a result, the temperature is a contributing factor that alters a device’s electrical performance, and thus must be accurately characterized for thermally critical applications as to not influence operation.

3.1.2. Heat Generation Mechanisms in Microelectronics

In addition to altering device operation, self-heating effects can result in overheating concerns that further impact reliability and lifetime of a device. The temperature distribution
within a semiconductor device is mainly governed by the heat conduction equations (Table 3.1) stating that heat is either conducted away \( Q_c = \kappa \nabla T \) from the source location or is absorbed by the structure as latent heat and thus increases the temperature locally.

Table 3.1. Source-driven conduction equation governing heating in micro-electronic devices.

<table>
<thead>
<tr>
<th>Within the Heat Source</th>
<th>Outside Heat Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho \cdot C_p \frac{\partial T}{\partial t} = -\nabla \cdot \vec{Q}<em>c + Q</em>{gen} )</td>
<td>( \rho \cdot C_p \frac{\partial T}{\partial t} = -\nabla \cdot \vec{Q}_c )</td>
</tr>
<tr>
<td>( \rho \cdot C_p \frac{\partial T}{\partial t} = -\nabla \cdot (\kappa \nabla T) + Q_{gen} )</td>
<td>( \rho \cdot C_p \frac{\partial T}{\partial t} = -\nabla \cdot (\kappa \nabla T) )</td>
</tr>
</tbody>
</table>

The heat conduction problem for the major front-end devices and components presented in this work is considered to be a source driven thermal system, where heat generation drives the thermal response, in contrast to boundary-driven systems. Since the wafer dimensions are smallest in depth direction, the wafer domain is considered to extend infinitely in the lateral directions and thus edge boundaries are considered to be adiabatic. Only the substrate base boundary condition is considered as the heat sink, and thus thermal resistances and temperature rises will be relevant between the junction (highest device temperature) and the case (substrate base) for our unpackaged wafer devices.

The heat generation (driving term for our thermal system) is the result of electrical interactions and can most simply be defined per unit volume as presented in (3.14) in terms of the current density \( J \) and the electric field \( E \).

\[
Q'''' = (\vec{J}_n + \vec{J}_p) \cdot \vec{E} \tag{3.14}
\]

The heating term in application is more complex and comprises of different electro-thermal heat generation mechanisms that occur when a semiconductor device is activated. The electric field \( E \) is a dependent parameter and thus the heat generation must be presented in terms of current densities \( J \) and charge carrier properties which are often solved for in the problem definition.
Joule heating is the primary generation term in most active devices, especially for analog and high power devices, and consists of a low frequency (DC) and a high frequency (optical) term. The DC term in (3.15) is the by-product of electron and hole flow random collisions with the lattice structure that creates thermal motion. The optical heating is the result of absorption of an optical wave into the lattice where the phonon energy either generates an electron-hole pair, or dissipates as Joule heating.

\[ Q_{\text{Joule,DC}} = \frac{|J_n^2|}{q\mu_n n} + \frac{|J_p^2|}{q\mu_p p} \]  

(3.15)

Another form of heat generation is the Recombination heat which occurs when an electron-hole pair recombines. This term is the primary mechanism for light emitting device operation, like photodiodes and laser diodes. When an electron leaves the valence band to the conduction band, energy is absorbed and an electron-hole pair is generated. The reverse process is Recombination where the electron loses an energy packet and recombines in the hole site within the valence band. The energy lost from the electron is emitted as a photon packet and is either released from the device, as is the case in light emitting devices, or is reabsorbed in the lattice as additional thermal energy. The total heat generated (3.16) from this mechanism depends on the total recombination rate (spontaneous, trap assisted, and Auger) and the difference in energy (Fermi) levels of the conduction and valence band \((E_{fp} \text{ and } E_{fn})\).

\[ Q_{\text{Recomb}} = R_{\text{net}}(E_{fn} - E_{fp}) \]  

(3.16)

\[ = (R_{\text{trap}} + R_{\text{Aug}} + R_{\text{spon}})(E_{fn} - E_{fp}) \]

Thompson and Peltier generated/absorbed heat (3.17) are additional heat generation mechanisms that occur as a result of spatial and temporal variation, respectively, in the thermoelectric power of an electron \((P_n)\) and a hole \((P_p)\) within the semiconductor material. Thompson heat is related to time change in \(P\) when electron-hole pairs are generated or
recombined. Generation of the pair absorbs energy and the recombination releases energy as the form of heat. Peltier is related to electro-thermal heat generation/absorption as the flow of electrons and hole across a spatial thermoelectric power gradient.

\[
Q_{\text{Thompson}} = qR_{\text{total}}T(P_p - P_n)
\]

\[
Q_{\text{Peltier}} = -T(\vec{J}_n \cdot \vec{∇}P_n + \vec{J}_p \cdot \vec{∇}P_p)
\]

The formulation of the different heat generation regimes is important for electro-thermal modeling, however, in some applications such as high power devices, discussed in detail in Chapter 4 and 5, Joule heating is the dominant term and other forms of heat generation may be neglected.

3.1.3. Temperature-Dependent Thermal Properties

Temperature is not only a contributing factor in governing physical (electrical and thermal) equations but also alters material and thermal properties. For instance, a semiconductor material’s band gap drops with temperature \( E_g(T) = E_{g0} - \alpha T^2/(T + \beta) \), which limits its high voltage capabilities by reducing its breakdown voltage. Moreover, elevated temperature levels increase phonon concentrations causing lattice scattering and a reduction in electron mobility. Most mobility models consider temperature for more accurate representation. In a similar manner, temperature can affect most other related parameters such as electron and hole diffusivity, density of states and electrical resistivity.

Temperature also affects thermal properties where increases lattice thermal energy increases scattering and thus reduces thermal conductivity in ordered structures (non-amorphous) as well as heat capacity. Thermal conductivity is seen to decrease with temperature for crystalline and polycrystalline materials used in semiconductor devices (GaN, Si, SiC...) due to the increase scattering at higher lattice thermal motion. The most abundant model for fitting this temperature dependence is the power law (3.18) in reference to ambient conditions
(300 K). The fitting parameter $\alpha$ and $\kappa_0$ are presented for readily used materials in Table 3.2.

$$\kappa(T) = \kappa(300K) \cdot \left( \frac{T}{300} \right)^\alpha$$  \hspace{1cm} (3.18)

Table 3.2. Temperature-dependent Conductivity Parameter for generic Semiconductor compounds (Data compiled by ATLAS SILVACO).

<table>
<thead>
<tr>
<th>Material</th>
<th>$K_{300}$ (W/m K)</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>148</td>
<td>-1.65</td>
</tr>
<tr>
<td>Ge</td>
<td>60</td>
<td>-1.25</td>
</tr>
<tr>
<td>GaAs</td>
<td>46</td>
<td>-1.25</td>
</tr>
<tr>
<td>AlAs</td>
<td>80</td>
<td>-1.37</td>
</tr>
<tr>
<td>InAs</td>
<td>27</td>
<td>-1.1</td>
</tr>
<tr>
<td>InP</td>
<td>68</td>
<td>-1.4</td>
</tr>
<tr>
<td>GaP</td>
<td>77</td>
<td>-1.4</td>
</tr>
<tr>
<td>GaN</td>
<td>265</td>
<td>-1.38</td>
</tr>
<tr>
<td>AlN</td>
<td>295</td>
<td>-1.21</td>
</tr>
</tbody>
</table>

In addition to temperature dependence, two elements may share the same sub-lattice in ternary and quaternary compound semiconductors, which are represented by the percent composition $x$ (ex: $Al_xGa_{1-x}N$, $In_xGa_{1-x}P$, ...) The blend introduces more irregularities in the crystal structure due to varying atomic sizes, and thus causes a decrease in the thermal conductivity and leads to an $x$-dependence in $\kappa$. For example, pure GaN and AlN are more conductive than an $Al_xGa_{1-x}N$ blend where $Al$ and $Ga$ atoms share the same sub-lattice. The conductivity variation with composition thus appears as a basin function determined by the parameter $C$ as showing in the general fit for semiconductor blends in (3.19). The density $\rho$, specific heat $c$, and $\beta$ are interpolated linearly with composition $x$.  

31
\[
\kappa_{300}^{AB} = \frac{1}{\left(\frac{1-x}{\kappa_{300}^A} + \frac{x}{\kappa_{300}^B} + \frac{(1-x)x}{C}\right)}
\]

(3.19)

\[
\alpha^{AB} = (1-x)\alpha^A + \alpha^B
\]

(3.20)

The heat capacity temperature-dependency is also represented in a similar form in (3.21). The fitting parameters and ambient values are shown in Table 3.3 for a selected list of semiconductor materials.

\[
C(T) = \rho \left[ c_{300} + c_1 \frac{T^{\beta}}{300} - \frac{1}{T} \right]
\]

(3.21)

Table 3.3. Heat Capacity Ambient Value and Model Parameters (Data compiled by ATLAS SILVACO).

<table>
<thead>
<tr>
<th>Material</th>
<th>(\rho_B) (gm/cm)</th>
<th>(c_{300}) (J/K kg)</th>
<th>(c_1) (J/K kg)</th>
<th>(\beta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>2.33</td>
<td>711</td>
<td>255</td>
<td>1.85</td>
</tr>
<tr>
<td>Ge</td>
<td>5.327</td>
<td>360</td>
<td>130</td>
<td>1.3</td>
</tr>
<tr>
<td>GaAs</td>
<td>5.32</td>
<td>322</td>
<td>50</td>
<td>1.6</td>
</tr>
<tr>
<td>AlAs</td>
<td>3.76</td>
<td>441</td>
<td>50</td>
<td>1.2</td>
</tr>
<tr>
<td>InAs</td>
<td>5.667</td>
<td>394</td>
<td>50</td>
<td>1.95</td>
</tr>
<tr>
<td>InP</td>
<td>4.81</td>
<td>410</td>
<td>50</td>
<td>2.05</td>
</tr>
<tr>
<td>GaP</td>
<td>4.138</td>
<td>519</td>
<td>50</td>
<td>2.6</td>
</tr>
</tbody>
</table>

The above modeling equations present the idealized temperature and composition dependencies. It is important to note that thermal properties also vary with doping concentrations, fabrication methods, and dislocation density. These numerous factors contribute to the complexity of accurately modeling thin-film properties, and in most cases, the presented literature values are estimates used for bulk-material (Table 3.2).
3.1.4. TDTR Property Measurements

As thermal properties of thin-films exhibit large margins of uncertainty due to fabrication, crystal structures and material doping concentrations and dislocation densities, a thermal property measurement setup is extremely beneficial to obtain accurate thermal properties to be used as input for numerical and simulation models. Time-domain thermoreflectance (TDTR) is a versatile technique developed for measuring the thermal conductivity ($\kappa$) of thin-film materials. The pump-and-probe TDTR method used in this work [26,28], has been implemented for a wide range of thin-film thermal conductivities, between a high $\kappa$ of 2000 W/m.K measured for diamond and graphite [268,269] to a low of 0.05 W/m.K measured for WSe$_2$ [53].

TDTR can be implemented with relatively simple sample preparation and without the need for resistor heater/transducer patterns required for other types of conductivity measurement techniques ($3\omega$). The basic principle of the TDTR approach is that a “pump” laser optically excites a top absorptive layer of a thin-film stack (Fig. 1), inducing a thermal response. The top absorptive layer, gold (Au) in the current work, also serves as a transducer layer from whose surface a CW “probing” laser reflects into a photodetector to monitor changes in reflected light. Based on thermoreflectance physics, the temperature rise and decay at the surface can be inferred from the reflectance response. The temperature response is used in a reverse modeling approach to extract desired information related to the thermal properties of the thin-film stack. The transient temperature response at the surface is governed by the thicknesses and thermal properties of the stack of thin films making up the sample. These properties can be extracted by comparing the experimental response to a computational solution of a corresponding thermal model. The normalized temperature ($\theta$) response is governed by Fourier’s law of thermal conduction given in (3.22). The equation is reduced to 1D through-plane (z-direction) conduction since the heating spot is two orders of magnitude larger than the probing spot, which makes the radial variations negligibly small.

$$\rho c_p \frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left( \kappa \frac{\partial \theta}{\partial z} \right) + Q_{ab}(z, t) \quad (3.22)$$
The temperature response is driven by the heat addition, $Q_{ab}$, of the pump laser which is absorbed optically by the surface transducer layer. Convective and radiative effects are negligible because the heating is localized and the temperature rise is low, and so an adiabatic temperature condition is prescribed at the top boundary ($\partial \theta / \partial z = 0$ at $z = 0$). The condition at the wafer’s base is set to match the controlled chuck temperature ($\theta = \theta_b$ at $z- > \infty$). The heat addition, $Q_{ab}$, given in (3.23) depends on the reflectivity of the material $R$ and the absorption rate $\gamma$ of the incident laser intensity $I$. The laser intensity $I(t)$ is expressed in (3.24) in terms of the fluence $F$ and the pulse width $\tau$ of the pulse laser, while the absorption $\gamma$ is given in terms of the laser wavelength $\lambda$ and the transducer’s extinction coefficient $k$ as $\gamma = 4\pi k / \lambda$.

$$Q_{ab}(z,t) = I(t)(1 - R)\gamma e^{-\gamma z} \quad (3.23)$$

$$I(t) = \frac{2F}{\tau \sqrt{\pi}} e^{-4((t-t_0)/\tau)^2} \quad (3.24)$$

The energy balance equation (3.22) can be solved analytically for simple structures such as the semi-infinite medium. However, for more complex layer stacks, the equation must be solved numerically in space and time. The calculated temperature response at the surface is then normalized and fitted to the normalized measured response by varying the unknown thermal properties that make up the stack.

The total thermal resistance ($R_{th, total}$) of the sample thin-film material as given in (3.25) consists of two terms: a) the intrinsic thermal resistance, which depends on the material intrinsic thermal conductivity ($\kappa_i$), and b) the interface thermal resistances ($R_{int}$) at the interface at the sample film interfaces.

$$R_{th, total} = \frac{h}{\kappa_i} + R_{int} \quad (3.25)$$

The total thermal resistance is linear for a sample layer of thickness $h$. By plotting $R_{th, total}$ against $h$, it becomes possible to extract the $\kappa_i$ and $R_{int}$ values from the slope.
and y-intercept of the linear fit for each of the tested BEoL materials. As a final step, it is helpful for thermal simulations to report an effective conductivity for the thin-film materials. This is because for thin films, the interface resistance becomes a significant contributor to $R_{th,\text{total}}$ and it is impractical to consider the interface within a Fourier continuum approach. An effective conductivity ($\kappa_{eff}$) would comprehend both the interface effects as well as the material’s intrinsic thermal resistivity, and is hence well-suited for computational analysis.

By definition, a thin-film layer of thickness $h$ and thermal conductivity $\kappa_{eff}$ would provide the same thermal resistance as a film of thickness $h$ and conductivity $\kappa_i$ in series with an interface resistance $R_{int}$. This relation (3.26) is used to calculate and plot the $\kappa_{eff,i}$ as a function of $h$, as given in (3.26).

$$R_{th,\text{total}} = \frac{h}{\kappa_i} + R_{int} = \frac{h}{\kappa_{eff}}$$  \hspace{1cm} (3.26)

$$\kappa_{eff,i} = \left(\frac{R_{int}}{h} + \frac{1}{\kappa_i}\right)^{-1}$$  \hspace{1cm} (3.27)

### 3.2. Self-Adaptive Multi-Scale Thermal Modeling for Ultra-fast Computational Capabilities

Even with accurate thermal properties and a set of complete governing equation models for the electro-thermal system, the numerical complexity of solving the heat conduction equations in microelectronic structures remains and the major challenges arise from several factors. To begin with, heat is generated in localized region in the active devices and dissipates through the electronic stack to the substrate, transitioning across multiple physical scales in the three spatial dimensions and through different materials. The temporal scales are also irregular where the period of heat generation is orders of magnitude that of the heat dissipation. Undertaking full thermal modeling with traditional numerical techniques would require tremendous computational efforts. Consequently, iterative and models optimizations, which may be required given the large uncertainties and complexities present, are
infeasible with full scale simulations.

To overcome the above challenges, Wilson and Raad [262] presented a numerical approach tailored specifically for microelectronic thermal modeling applications, in which a multi-grid method is utilized to handle large variations in scale, and where grid meshing is physics-driven. The grid generation is automatically refined to ensure grid invariance which makes the numerical methods easy to setup and independent of any user experience. The transient 3D modeling engine refines the domain where required and thus increases the efficiency of computations, and as a result, reduced the solution time by two orders of magnitude [262]. This drastic improvement in computational efficiency opens the possibilities for ultra-fast computations to be implemented and thus iterative methods and optimizations of the models become feasible.

The developed thermal modeling software TMX T°Solver® based on the above described numerical approach is the simulation software used at NETS lab and will be used for the microelectronic device thermal modeling studies of this work. T°Solver® solves the heat transfer conduction problem within a defined model structure. The temperature dependent thermal conductivity and heat capacity is modeled within T°Solver according to (3.18) and (3.21). The boundary conditions are considered adiabatic at the lateral boundaries, since the wafer is considered infinitely wide in comparison to the depth in the through-plane direction. The top boundary is also considered adiabatic since heat transfer from the top surface is neglected. This condition is justified since convection effects are minimal for the studied activation levels and the temperature rise level is minimal for any significant radiation heat transfer. Moreover, the surface of microelectronic wafers are passivated with a dielectric (Silicon Nitride) adding to the thermal insulation to the top surface and blocking heat transfer. The heat sink boundary condition (isothermal) is defined at the substrate base to match experimental conditions where the unpackaged wafer is mounted during testing on a thermally controlled stage of a large thermal capacity, and for which the temperature is monitored and maintained at a fixed base temperature value ($T_b$). Other model details specific to each device technology will be presented accordingly in their respective Chapters.
3.3. Experimentally-Driven Model Optimizations for Full Thermal Characterization

As discussed earlier, challenges in numerical thermal modeling arise in complex devices and in the uncertainties in physical parameters that may lead to inaccurate results. The presented self-adaptive multi-scale thermal modeling capabilities are most suited for transient simulations of 3D complex electronic structures, [208, 209, 211] and may offer a solution to the mentioned challenges. The fast simulations allow the possibility of conducting iterative studies to converge on uncertain input parameters, such as the heat source profile definition or the process-dependent thermal properties (thin film conductivity, interface resistances...). By iterating for the model input parameter files, the thermal model can be driven to match experimentally obtained temperature boundary values measured at device surface by methods described in Chapter 2. The obtained optimized model would thus be experimentally validated and representative of the actual thermal response of the device under test. This would offer a digital representation of the device for further design scenarios and parametric studies. The details and benefits of the reverse modeling approach is presented in the upcoming Chapters where it will be applied to develop experimentally-valid thermal simulation models of complex high-power GaN devices to investigate their thermal performance and extract important thermal properties of the device structures.
Chapter 4
High-Resolution Thermoreflectance Imaging Investigation of Self-Heating in AlGaN/GaN HEMTs on Si, SiC, and Diamond Substrates

Perhaps the most stringent thermal considerations arise for microelectronics operating under high power applications. Microelectronic devices in development for power and switching applications, such as Gallium Nitride based devices, have been thermally-limited due to the extensive self-heating effects and the excessive operation temperatures. With this limitation to microelectronic performance and reliability, novel and innovative thermal management strategies are required to drive thermal design along with advanced characterization techniques especially suited to meet the challenges in device scale, measurement accessibility, and numerical modeling.

Gallium Nitride high electron-mobility transistors (HEMTs) offer considerable high-power operation but suffer in reliability due to potentially damaging self-heating. In this study, self-heating in AlGaN/GaN HEMTs on high conductivity substrates is assessed using a high-resolution thermoreflectance (TR) imaging technique, to compare the thermal response between GaN-on-Si, GaN-on-Diamond, and GaN-on-4H-SiC. The TR method accuracy at high power density is verified by using a non-linear coefficient of thermoreflectance ($C_{TR}$) as function of temperature. The acquired steady-state thermal maps give a thermal resistance of 11.5 mm.K/W for GaN-on-Si (based on peak channel temperature), compared to 2.7 and 3.3 mm.K/W for GaN-on-SiC and GaN-on-diamond substrates respectively. The tested GaN-on-diamond HEMT exhibits similar heating rates to those seen on a SiC substrate, with a slightly higher peak temperature, which indicates a higher thermal boundary resistance that could offset the benefits of using a higher conductivity substrate and lead to faster thermally-enhanced degradation. The analysis reveals the importance using high resolution imaging to detect hot-spots and areas of peak temperature that largely affect failure
initiation and device reliability and which may not be otherwise observable.

4.1. Introduction

The III-V compound semiconductor Gallium Nitride (GaN) has found widespread application in advanced high frequency and high voltage switching applications due to GaN’s high breakdown electric field and electron mobility [171]. This exceptional electrical performance, however, has been documented to suffer from undesirable self-heating effects under high power applications [5,240]. Therefore, the need of improvement in thermal performance via reduced self-heating has necessitated research into HEMTs on high thermal conductivity substrates such as diamond [256].

GaN HEMTs grown on Silicon substrates have been taking advantage of a mature substrate-manufacturing infrastructure that has eased these devices into the power electronics marketplace. However, this approach suffers from substantial self-heating effects mainly because of buffer layers of AlGaN, that are usually thick, highly defective and thermally-resistive, but are necessary to bridge the lattice and thermal expansion mismatch between Si and GaN (lattice constants of 5.43 and 3.18 Å respectively [191]). 4H-SiC substrates not only offer higher thermal conductivity ($\approx 400$W/m.K), but they also alleviate the need for AlN or AlGaN buffer layers thanks to a better lattice match to GaN. Moreover, GaN can be grown directly on SiC because the two materials have similar wurtzite crystal structure with 3.5% mismatch [41]. While diamond has very high thermal conductivity, its integration with GaN HEMTs has been extremely challenging for a number of reasons. To name a few, it is an unavailability of single crystal diamond wafers of large device processing size on the current market, the high cost of single crystal substrates of existing size, as well as a challenge of bonding GaN to single-crystal diamond wafers. Recent studies demonstrated that the use of polycrystalline diamond could remarkably improve HEMT electrical and thermal performance for GaN-on-diamond HEMTs [9,73,74,239].

Polycrystalline Diamond substrate, grown with chemical vapor deposition (CVD-diamond) in-lieu of the etched Si substrate, necessitates an adhesion/bonding layer of SiN. The result-
ing GaN-on-diamond devices are observed to be structurally stable [149] with a three fold increase in power density as reported by Pomeroy et al. [198].

4.1.1. Thermal Mapping Challenges for GaN HEMTs

The different substrate configurations and interfaces in GaN HEMTs produce variation in thermal performance that require some experimental validation to observe the actual device thermal response. Experimental analysis is essential since simulation results alone may not be sufficient or valid due to inherent uncertainties in device and material properties [138]. There are various experimental thermal measurement techniques, however each of them present some limitations in measuring the maximum channel temperature of GaN HEMTs, mostly related to accessibility, resolution, and accuracy of the measurements [13, 164, 193, 198]. Optical methods of thermal imaging with sub-infrared (IR) wavelength are proven to provide higher resolutions with sufficient access to active regions of HEMT devices. Recent advances in optical thermoreflectance (TR) technique that utilizes visible and near-UV based thermal imaging shown to provide resolutions that can reach 10× that of IR.

4.1.1.1. Accessibility

To measure the operating temperature of GaN HEMTs, optical techniques are preferred for their non-invasive and non-destructive measurement approach. A natural limitation for optical measurements, such as infrared (IR), Raman and Thermoreflectance (TR), is that they are limited to measuring surface features and cannot reach embedded regions of a device, such as a Source-Drain channel covered by a field-plate, or the channel region under the gate electrode. One accessibility solution proposed by Martin-Horcajo et al. [164] is measuring the back-side gate electrode of a GaN-on-SiC HEMT. This approach is able to give a localized temperature measurement but is not applicable for other devices with opaque substrates. Moreover, some distortion may occur due to the (a) signal passing through multiple transparent and semi-transparent layers, (b) diffusion of the spot size due to those layers, and (c) not knowing exactly where one is measuring from. In addition,
optical measurements are limited to measuring opaque films and cannot measure transparent thin films, often requiring additional sample coatings, and leading to an uncertainty as to the measurement location. This is the case with emerging materials such as Gallium Oxide ($Ga_2O_3$) which are transparent to IR and visible light, but may reflect higher energy illumination.

4.1.1.2. Accuracy

In addition to the accessibility limitation, measuring the maximum temperature using available methods can undeniably suffer in accuracy, due to depth and region averaging, among other factors. Surface temperatures obtained by thermoreflectance (TR) is lower than the electron channel temperature at the GaN/AlGaN interface beneath the surface. Raman spectroscopy averages throughout the depth of the GaN medium, which is transparent to visible light, giving lower-than-actual readings [14]. Electrical approximation methods’ [232] accuracy can even be worse as they consider the effect of the device temperature on its electrical characteristics such as the gate resistance measurements [165, 193]. In that manner, electrical and photoluminescence measurements report on the average temperature of the entire device which may lead to an underestimation of the peak channel temperature, and which also can be influenced by charge trapping effects [180, 228]. Killat found that when compared to Raman, pulsed-electrical (PIV) measurements underestimate the 2DEG channel temperature by around 7%, with DC measurements underestimating the maximum temperature by up to 50% [128]. Other invasive/contact measurements utilizing embedded temperature sensors in the channel region may add heat conduction paths and electrical interference with the device operation, and also requires additional sample fabrication and electrical access to the sensing probes [13].

4.1.1.3. Resolution

A third challenging aspect that faces thermal mapping of GaN devices, and microscale electronic devices in general, is the deep imaging resolutions required to detect temperature
peaks. Infrared (IR) thermography has a diffraction-limited lateral resolution of around $3 - 10\mu m$, with no depth resolution since GaN is transparent to IR and visible light. Novel approaches have been aiding improvements in spatial resolution, such as the fabrication of solid immersion lenses (SIL) on the back side of transparent substrates [199]. However, the costly sample preparation method is still in its early developments and is not applicable to opaque substrates.

Thermoreflectance (TR) thermal imaging for temperature mapping of the active GaN HEMT devices used in this study offers advance capabilities most suited for microelectronic components. TR is a non-invasive technique based on detecting surface reflectance changes with temperature [98,243], and therefore it is more suited than IR for the high reflectivity and low emissivity materials present in microelectronic devices [129,177]. By employing a shorter wavelength illumination than IR (visible and near-ultraviolet (NUV)), the TR method can image at higher resolutions [99].

Given that GaN is transparent to IR and visible light, TR can measure reflectance of the GaN surface by using NUV illumination, which is higher than the bandgap of GaN [194,243,272]. Above-bandgap TR provides surface measurements from the GaN layer.

In comparison, $\mu$-Raman reports a thickness average temperature within the GaN layer, and visible TR [160,193] illumination reflects from multiple interfaces within the stack structure, with both effects leading to an underestimation in GaN surface temperature.

In this study, above-bandgap TR is used to experimentally assess and compare the passive cooling effectiveness of GaN HEMT devices on three different substrates. The sub-micron resolution thermal maps are obtained using TR imaging with enhanced accuracy provided by a higher order calibration for a temperature-dependent TR coefficient ($C_{TR}$).

In this study, GaN HEMTs device structures on three different substrates are experimentally assessed for their passive cooling effectiveness. The sub-micron resolution thermal maps are obtained through TR imaging with enhanced accuracy and suitable device accessibility. A higher order calibration approach is performed to improve the accuracy of the measurements and to compare the thermal performance of the different GaN HEMT devices.
4.2. Methodology

4.2.1. Thermal Mapping Procedure

4.2.1.1. Thermoreflectance Setup

The TR measurements are conducted using the TMX T°Imager® thermal metrology system that is capable of steady-state as well as transient temperature measurements [241]. This study presents the quasi-steady temperature maps obtained from LED illumination at 365 nm wavelength, while the transient analysis will be addressed in the future work.

![HEMT device structure and top view image](image)

Figure 4.1. a) Schematic of HEMT device structure. The substrate is either Si, SiC, Diamond, or is etched b) Top view image showing Source, Drain and Gate contacts.

In order to minimize subsurface reflections, the measurements were performed with above-bandgap illumination wavelength (WL) of 365 nm, which corresponds to the bandgap of GaN [194, 272]. The thermal imaging results were verified with additional measurements
made first with a 370 nm and then with 360 nm bandpass optical filters to block all below-bandgap wavelengths.

Light entering the optical setup is first sampled using a photodetector (PD) to monitor and corrects for any fluctuations in the source intensity. The light beam is directed down to the sample through a 100× objective lens, and reflects back up to a mounted camera that acquires a 1024 × 1024 intensity map of the light reflection with a pixel size of 70 nm (Figure 4.1). The thermal chuck is mounted on an XYZ nano-positioning stage in order to correct for both lateral motion (drift and jitter) and vertical motion (focus) that may result from thermal expansion at higher magnification levels.

![Figure 4.2](image)

**Figure 4.2.** Sample a) calibration map for $C_{TR}$ showing a global reflectance change response from the GaN material, b) reflectance change map $\Delta R/R$ showing the localized response within the GaN Channel Region, c) the resulting temperature rise map $\Delta T$ for the active GaN-membrane and GaN-on-Si device showing self-heating within the channel.
4.2.1.2. Activation

The acquisition of the thermal maps is performed in two main steps: activation and calibration (Figure 4.2). In activation, the device-under-test (DUT) is placed on a thermally controlled stage set at 20°C, and activated using a Keithley 2410 power source at a given power level. The drain voltage is modulated while the reflected intensity ($R$) is measured by the camera. A total number of 50 pairs of "cold" and "hot" frames are acquired for the OFF and ON device activation state respectively. The reflected intensity change map $(\Delta R/R)_{\text{Act.}}$ is then extracted. The procedure is repeated for various activation levels and biasing conditions (Table 4.1).

Table 4.1. Activation conditions for the different GaN HEMTs. Source contact is grounded.

<table>
<thead>
<tr>
<th>Device</th>
<th>Gate Voltage (V)</th>
<th>Drain Voltage (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN-on-Si</td>
<td>-1</td>
<td>5 to 25</td>
</tr>
<tr>
<td>GaN-on-SiC</td>
<td>0</td>
<td>5 to 25</td>
</tr>
<tr>
<td>GaN-on-diamond</td>
<td>-1</td>
<td>5 to 25</td>
</tr>
<tr>
<td>di-GaN-on-SiC</td>
<td>-1</td>
<td>10 to 50</td>
</tr>
<tr>
<td>GaN-membrane</td>
<td>-0.8 to -1.8</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 4.2. Gate (G), Source (S) and Drain (D) dimensions (in $\mu$m) of GaN HEMTs on different substrates.

<table>
<thead>
<tr>
<th>Device</th>
<th>Gate Length</th>
<th>Gate Width</th>
<th>S-G spacing</th>
<th>G-D spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN-on-Si</td>
<td>2.1</td>
<td>75</td>
<td>3.0</td>
<td>10.2</td>
</tr>
<tr>
<td>GaN-on-SiC</td>
<td>3.6</td>
<td>75</td>
<td>1.4</td>
<td>11.4</td>
</tr>
<tr>
<td>GaN-on-diamond</td>
<td>1.9</td>
<td>75</td>
<td>3.5</td>
<td>9.6</td>
</tr>
<tr>
<td>di-GaN-on-SiC</td>
<td>2.2</td>
<td>75</td>
<td>1.4</td>
<td>11.4</td>
</tr>
<tr>
<td>GaN-membrane</td>
<td>3.3</td>
<td>75</td>
<td>4.0</td>
<td>8.6</td>
</tr>
</tbody>
</table>
4.2.1.3. **Calibration**

Calibration is necessary to provide the T-R relation that translates the reflectance change maps ($\Delta R/R$) into temperature rise maps ($\Delta T$). As TR response depends on several parameters such as illumination, passivation, surface material, roughness, and optical setup (numerical aperture), in-situ calibration is needed for every device. A known temperature change $\Delta T_{1,2}$ is set using the thermal stage and another set of reflectance intensity maps are obtained at the low and high temperature. The TR coefficient $C_{TR}$, which represents the reflectance change per degree change in temperature, is then calculated according (4.1) at every pixel location $(i, j)$.

$$C_{TR}(i, j) = \frac{1}{\Delta T_{1,2}} \left( \frac{\Delta R}{R}(i, j) \right)_{cal.}$$ (4.1)

Since the TR response has been observed to weaken at higher temperatures \cite{100, 194}, a systematic set of calibrations were performed at successfully higher temperatures to obtain the relationship between $C_{TR}$ and $\Delta T$, which was observed in this study to be linear ($C_{TR}(\Delta T) = a\Delta T + b$, where $a$ and $b$ are linear coefficients). This would ensure an improved accuracy over the wider ranges of temperature experienced by the tested GaN devices ($\approx$50-100K); otherwise, the temperature rise would be underestimated.

4.2.1.4. **Thermal Map**

Finally, the temperature rise map $\Delta T_{\text{activ.}}$, which is relative to the base temperature $T_{\text{base}}$ was calculated by combining the calibration maps and the activation maps at each power level according to (4.3), and solving for $\Delta T$, where $a$ and $b$ are obtained from calibration, and $\Delta R/R$ from activation.

$$\Delta T = \frac{1}{C_{TR}} \frac{\Delta R}{R} = \frac{1}{a\Delta T + b} \frac{\Delta R}{R}$$ (4.2)

$$a\Delta T^2 + b\Delta T = \Delta R/R$$ (4.3)
The thermal maps report temperature rise results within a maximum error of 5% (displayed in color), while higher error data points are eliminated to reveal the DUT (grayscale) (Figure 4.2). To evaluate the actual thermal performance of each of the device structures, the heating profiles are obtained for the different substrate configurations and at increasingly higher power levels.

4.2.2. Sample Description

The studied devices are commercial GaN HEMTs on (111) Si, 4H-SiC, and CVD-diamond substrates with an additional GaN-membrane sample (i.e., etched Si substrate) [241] and a commercial GaN-on-SiC capped with a 500 nm thick layer of nano-crystalline diamond (NCD).

The GaN-on-Si HEMT consisted of a 428 nm AlN seeding layer, a 936 nm of AlGaN transition layers, a 737 nm thick GaN buffer layer, and a 20 nm thick AlGaN barrier layer to provide the 2-DEG channel. The ohmic contacts for the Source and Drain were patterned to the dimensions shown in the device images (Figure 4.1). Ohmic metallization consisted of 20 nm Ti/120 nm Al/40 nm Ni/50 nm Au, rapid annealed for 30 seconds at 850 °C in N₂ atmosphere [239]. The Schottky gate electrode is composed of 20 nm Ni followed by 200 nm Au. A 100 nm thick PE-CVD SiN layer was deposited for the purpose of passivation. To maximize exposed channel surface region for optical measurements no field plates were implemented. The channel width is 75 µm and the gate length and gate-drain spacing vary among the devices tested (Table 4.2).

For GaN-on-SiC HEMTs, commercial heterostructures consisting of 25 nm Alₐ₂₋₂Gaₐ·₇₈N / 1.5 µm GaN on 4H-SiC were fabricated concurrently [241, 243]. Typically, about 100 nm thick AlN nucleation layers are employed to manage lattice/thermal mismatch between GaN and SiC [163].

The GaN-on-Diamond samples were built using a proprietary double-flip method [81] using material provided by Element Six, inc. The Si substrate is replaced with thick CVD diamond using a 30 nm SiN barrier layer deposited to the N-polar side of the heterostructure,
as described by Tadjer et al [239]. On the GaN/SiC and GaN/diamond epitaxial heterostructures (epi), an identical HEMT device fabrication process was employed [241,243].

4.3. Thermoreflectance Imaging Results

4.3.1. TR Activation

Sample activation and calibration maps for GaN-on-Si and GaN-membrane (Figure 4.2) show a negative TR response from the GaN material within the channel [242,243].

The average temperature rise to activation power relation (Figure 4.3) appears to be non-linear, with the heating rate decreasing at higher temperatures. "The Arrhenius Principle" states that the temperature rise must be linear with power for constant material properties. Non-linearities at higher temperatures are due to a change (often decrease) in thermal conductivity at elevated temperatures. For the present materials, the reduction in thermal conductivity creates a positive feedback that further increases self-heating, leading to an upward curving Power-Temperature (P-T) relation, and not the observed downward curving slope. Therefore, the slowing heating rate maybe caused by a drop in the TR signal, which is seen when multiple temperature calibrations are performed (Figure 4.3). To obtain the thermal maps with high accuracy, an existence of $C_{TR} - T$ dependency at elevated temperatures is assumed and is investigated for all the devices through a multiple temperature calibration, while the TR relation is assumed as non-linear as presented in (4.4).

$$\left( \frac{\Delta R}{R} (T) \right)_{act.} = \int C_{TR}(T) \cdot dT$$

(4.4)

4.3.2. TR Calibration

Multiple calibrations are performed between the fixed base temperature of 20°C and a high temperature ranging between 40°C and 100°C. The obtained $C_{TR}$ is then plotted against the high temperature to observe and correct for any $C_{TR}$ temperature variations. From the temperature-dependent $C_{TR}$ relation, the reflectance change can be extracted according to
Figure 4.3. a) The calibration curve of CTR vs T showing a decrease in TR response at higher temperatures. b) Temperature rise curve as a function of applied power for GaN-on-Si device before and after correction for CTR-T dependence. The correction recovers the linearity of P-T relation for the GaN-on-Si devices.
Figure 4.4. Calibration results for GaN-on-SiC and GaN-on-diamond showing $C_{TR}$ linear decrease with temperature.

(4.4), and reveals that for GaN-on-Si, the coefficient decreases with temperature (Figure 4.3). A similar $C_{TR} - T$ dependency was also observed for the GaN channel in the remaining substrate configurations (Figure 4.4).

Evidently, the calibration curves show that the thermoreflectance response decreases linearly at higher temperatures, and the correction will result in an increase in the measured temperature rise. Assuming a uniform decrease of $C_{TR}$, a single calibration map is used, and a pixel-by-pixel correction is performed for the temperature dependent $C_{TR}$. The temperature rise data is obtained by combining the activation and calibration maps and then correcting for non-linear TR response. Temperature rise results within a maximum error of 5% are displayed in color, while higher error data points are eliminated to reveal the DUT (grayscale) (Fig 4.2).

This correction improves the accuracy of the thermal measurements that otherwise would underestimate the self-heating in the device by up to 35% for GaN-on-Si (Figure 4.3).
4.3.3. TR Thermal Mapping Results

Thermal imaging is performed for the activated device samples with different substrates at increasing power levels. The same procedure is performed for the other GaN HEMT devices, where activation is performed first, multiple temperature calibrations next, and temperature corrections last.

A summary of measured average channel temperatures is presented in Fig. 4.5 and the heating rate for the measured HEMTs is summarized in Table 4.3.

Figure 4.5. A summary of the measured Average (open markers) and Peak (solid markers) experimental temperature rise (in °C) within the active channel of the measured GaN HEMT devices on the different substrates. The linear experimental fit are shown for the measurement data for the average (in dotted lines) and peak temperature (in long dashed lines). The heating rate (total thermal resistance) is reported from the linear regression analysis.
Table 4.3. Summary of the extracted thermal resistances (in mm.K/W) of the GaN HEMTs on Different Substrates based on average and peak temperature rise in the active channel region.

<table>
<thead>
<tr>
<th>Device and Substrate</th>
<th>Channel Average Heating Rate (mm.K/W)</th>
<th>Channel Peak Heating Rate (mm.K/W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN-Membrane</td>
<td>121</td>
<td>-</td>
</tr>
<tr>
<td>GaN-on-Silicon</td>
<td>7.79</td>
<td>11.53</td>
</tr>
<tr>
<td>GaN-on-Silicon-Carbide</td>
<td>1.77</td>
<td>2.70</td>
</tr>
<tr>
<td>Diamond-GaN-on-SiC</td>
<td>1.82</td>
<td>2.46</td>
</tr>
<tr>
<td>GaN-on-diamond</td>
<td>1.14</td>
<td>3.25</td>
</tr>
</tbody>
</table>

4.3.3.1. GaN-on-Si and GaN-Membrane

The thermal mapping results (Figure 4.5) show that the highest average heating rate occurs for the GaN-membrane devices where the average temperature rise in the channel is observed to be an order of magnitude higher than that of GaN-on-Si devices. This is expected since the Si substrate is etched and replaced with a low thermal-conductivity polymer compound for structural support which drastically reduces thermal conductances. This excessive heating in GaN-membrane HEMTs would severely limit the activation level which render GaN-membrane HEMTs infeasible for high power applications. Nevertheless, GaN-membrane HEMTs found applicability in several micro-sensing applications, where high power is not a requirement, but thermal management and temperature control are important for maintaining a good accuracy.

The GaN-on-Si device shows an average thermal resistance of about 7.8 mm.K/W and about 11.5 mm.K/W based on peak channel temperature, which is close to the values reported in [242]. The ratio of thermal resistances between the membrane and the GaN-on-Si devices (121:7.8 or 15.5:1) is also comparable to the ratio of 11.6:1 reported in [192] and 9.8:1 reported in [242] for GaN HEMTs with an etched Si substrate.
4.3.3.2. GaN-on-SiC and GaN-on-diamond

The self-heating in GaN-on-Si is substantial when compared to GaN on a higher thermal conductivity ($\kappa$) substrates like CVD-diamond and 4H-SiC. The limiting factor in the GaN-on-Si heterostructure is thick nucleation layers of AlGaN with high defect density and low thermal conductivity [170].

The lowest average thermal resistance is observed for both GaN-on-SiC and GaN-on-diamond (1.77 and 1.14 mm.K/W compared to 7.79 mm.K/W for Si). The improvement in heat dissipation is due to: a) the high thermal conductivity of the substrates ($\kappa_{4H-SiC} \approx 330$ to 400 W/m.K [156], $\kappa_{diamond} \approx 1000$ W/m.K [269]), and b) the absence of the thick AlGaN buffer layers required to grow the GaN/Si epi, thus improving the thermal conductance between the GaN layers and the substrate.

The thermal performance of the two high-$\kappa$ substrates are comparable (Table 4.3), with GaN-on-diamond HEMT (with 30 nm SiN) exhibiting a lower average channel temperature ($R_{th} = 1.14$ compared to 1.77 mm $\cdot$ K/W), but a slightly higher peak temperature ($R_{th} = 3.25$ compared to 2.70 mm $\cdot$ K/W) which could lead to faster device degradation. A side-by-side comparison for the thermal resistance distribution along the channel length, $R_{th}(x) = \Delta T(x)/Q'$, (Figure 4.6) distinctly shows the difference in temperature rise profiles for these two devices.

With the standard thermal conductivity of diamond being three times higher than that of SiC, one would expect a higher thermal conductance from GaN-on-diamond, on the order of 20-40% for similar TBR values [237]. And so, the initially unexpected closeness in the observed thermal performance indicates the presence of a higher TBR in the GaN-on-diamond devices, which is attributable to a more thermally resistive SiN adhesion layer (measured for various SiN thicknesses between 18 and 40 $m^2K/GW$ [58, 237, 239], as compared to that in the GaN-on-SiC heterostructure (less than 10 $m^2K/GW$ [56]) resulting from a higher-$\kappa$ AlN seeding layer at the interface.

Another probable reason of reduced CVD-dimaond thermal performance could be a near-interface disorder of the diamond crystals at the seeding layer, which reduces the effective
Figure 4.6. a) Measured thermal maps showing temperature rise in GaN HEMTs active region on SiC (left) and diamond (right) substrate ($P \approx 0.5$ W, 6.6 W/mm). (b) The local thermal resistance across the Gate-Drain channel shows peak temperature rise at the drain-side edge of the gate.

substrate conductivity. More accurate estimates of the thermal properties of the tested devices can be extracted from a reverse modeling approach, which will be the topic of future study by the authors.

It is also important to note that the measured peak temperature may not be the highest device temperature, which may occur under the gate electrode, or a gate-connected field plate. Inaccessible to direct optical thermal imaging, these peak device temperatures would require a thermal modeling approach to estimate. The thermoreflectance imaging system used in this work provides deep-submicron resolution and high temperature accuracy, which enables the measurement of the temperature profiles as close as possible to the gate electrode. Such data are invaluable in validating thermal models and in understanding electrothermal interactions at the nanoscales [38].

Another main takeaway from the experimental investigation is the importance of using the peak temperature for reliability assessments, since average temperatures may depend on device dimensions and the averaging region. This could thus offer misleading information and
underestimate hot-spot temperatures leading to optimistic thermal and reliability models.

An exhibit of this discrepancy is seen when comparing measured GaN-on-diamond device in this study (D1) with another GaN-on-diamond device (D2) measured previously by the co-authors [239]. The devices D1 and D2 had slightly different channel length (19 and 11.5 µm), however, the averaging region is what resulted in the different heating rates of 2.95 mm.K/W for D2 compared to 1.33 mm.K/W for D1 (Figure 4.7), mainly due to a smaller averaging region around the hot-pot for D1. The peak temperatures and the temperature decay profile are similar in both devices since the same GaN/Di epi stack governs the through-plane and in-plane heat dissipation.

Figure 4.7. a) Measured thermal maps for GaN-on-diamond device 2 [239] with a different gate geometry than our measured device b) Comparison of temperature profile for both measured GaN-on-diamond devices with different drain channel length shows similar temperature rise.
While the average device temperature may be dependent on the device geometry and user defined parameters, the maximum temperature, which occurs at the drain side of the gate edge, appears to be independent of mentioned parameters, and is a more relevant parameter for reliability assessment for the following reasons. First, the device degradation is seen to initiate at location of a peak temperature, which is also the site of electron trap formation [169]. Moreover, the peak temperature location is where the maximum thermal mechanical stresses occurs, due to the presence of materials with different coefficients of thermal expansion (CTE), as well as due to the increase in inverse piezo-electric stresses [190], factors which eventually lead to pit or crack formation [95].

For the above reasons, the peak temperature is a more relevant parameter to report in reliability studies since its location is where failure initiates [136]. This also presents the self-heating as a more realistic thermal stressing method than global heating for reliability assessment studies, because it subjects devices to non-uniform temperature gradients and stress gradients observed in actual device operation. Therefore, for the most accurate device thermal investigations, high-resolution thermal imaging techniques, with enhanced accuracies are required to detect hot-spots that result from the self-heating in electrically active HEMT devices. Other thermal investigation methods that acquire the average device temperature (electrical measurements [79]) or localized temperature readings (i.e. with a µ-thermocouple [13]) based on global heating would fail to detect localized hot-spots and temperature gradients within the active device region.

4.3.3.3 Diamond-GaN-on-SiC

A slight improvement in the thermal response of the Diamond-GaN-on-SiC is observed in comparison to the GaN-on-SiC (Figure 4.5). The addition of the nano-crystalline diamond (NCD) served as an additional heat spreading in close proximity to the active region that lead to a slight reduction in the peak temperature (about 8%), far less than the 20% reported by Anderson et al for a gate after diamond approach [11]. For the device in this study, the diamond cap is added on top of SiN passivation layer, and AlGaN barrier layer, which
increased thermal resistance from device active region to NCD capping layer and reduced its thermal cooling effectiveness. Another numerical studies by Wang et. al reported an improvement of about 10% in drain current associated with device cooling [256]. The average channel temperature remained relatively unchanged (within 3%) most likely because the direction of main heat dissipation path from the GaN to the substrate was not altered by the addition of NCD capping layer.

4.4. Conclusions

An experimental investigation into the thermal performance of different substrates for GaN HEMT devices is presented using a high-resolution TR imaging method with correction for non-linearities at high temperatures. This results in a more accurate hot-spot temperature measurement, that would otherwise be underestimated by up to 20%, leading to overconfident thermal models of device lifetime and degradation. The importance of accurate detection of hot-spots is also argued for representative reliability assessment studies, adding to the benefits of high-resolution thermal imaging techniques.

The self-heating in HEMT devices on different substrates is compared between GaN-Membrane and GaN-on-Si, and between the highest conductivity substrates, GaN-on-diamond and GaN-on-SiC. Even with the diamond having a much higher thermal conductivity than SiC, a comparable thermal response between GaN HEMTs on the two substrates indicates a higher TBR could be contributing to the total thermal resistance of GaN-on-diamond.

A deeper investigation into the thermal characteristics of both substrate configurations is presented in the next Chapter to assess the observed thermal responses and provide a physical basis for comparison. Combining the thermoreflectance imaging data presented with a fast thermal simulation would provide reverse modeling capabilities to extract and compare important thermal parameters of the substrates and the GaN/substrate interface layers (Chapter 5).
Chapter 5

Full Thermal Characterization of AlGaN/GaN HEMTs on Silicon, 4H-Silicon Carbide, and Diamond Using a Reverse Modeling Approach

The experimental mapping of the tested devices has been presented in the previous Chapter, while in this Chapter, a full numerical approach is presented. Three dimensional (3D) thermal models are developed and optimized to match and accurately represent the actual device response obtained by thermal mapping. The approach of coupling the experimental results presented in Chapter 3 within the numerical analysis provides several advantages, that includes helping overcome the serious challenge of the unavoidable uncertainty in the input data that faces numerical modeling of complex microelectronic devices.

Gallium Nitride (GaN) high electron mobility transistors (HEMTs) operate at high power levels and are thus especially thermally-critical devices. Not only do they require innovative thermal management strategies, but can also benefit from advanced thermal characterization, both experimental and numerical, in their design and system integration stages. The thermal numerical analysis of microelectronic devices faces challenges of complex physics and uncertain thermophysical properties which leads to numerically expensive models that are prone to error. By the use of an innovative reverse modeling approach to mitigate the above challenges, this work presents the full thermal characterization of GaN power devices with different substrates aimed at managing performance-limiting self-heating. The approach develops and optimizes a thermal simulation model to match the numerical results to experimentally-obtained thermal maps of the devices under test. The experimentally-optimized simulation model can then be used to extract full 3D temperature distributions, infer in-situ thermal properties, and provide a numerical platform that can be used to conduct further parametric studies and design iterations. The presented analysis provides a full thermal characterization of different GaN HEMT devices and compares the thermal perfor-
mance of different substrates on the basis of thermal properties. The extracted properties for HEMTs on Si, SiC and diamond substrates are compared and a set of conclusions are presented to guide further developments in GaN HEMT thermal management strategies.

5.1. Introduction

The wide bandgap semiconductor materials such as Gallium Nitride (GaN) and Silicon Carbide (SiC) have been replacing traditional Silicon (Si) technology in advanced power switching applications due to their superior frequency and power performance [112,114,218]. GaN is a compound semiconductor material from the III-Nitride group that can withstand higher breakdown electric fields (3.3 MV/cm compared to 0.25 MV/cm for Si [217]), has a wide band gap (3.4 eV compared to 1.1 eV for Si), and provides a dense two-dimensional electron gas (2DEG) of high electron mobility (1500 – 2000 cm²/V.s at room temperature). These material properties make GaN highly suitable for high-voltage DC and high-frequency switching operation [95].

A major vulnerability for GaN device reliability stems from the extreme heat flux densities and destructive self-heating effects that the HEMT device encounters as a result high-power operation [257]. Elevated operating temperature in GaN devices increases phonon scattering and reduces the electron mobility in the GaN channel, causing potentially irreversible electrical degradation and eventual failure [65,108]. Consequently, effective thermal management strategies must be implemented to avoid device operational failure at high power and effectively exploit the full promise of GaN material properties.

Past mitigation strategies to suppress self-heating have been reported using active cooling techniques such as micro-channel cooling [2,44,92] and jet impingement cooling [141], or passive cooling techniques such as using higher thermal-conductivity substrates [105,237,268], or adding a heat-spreading material closer to the active regions to more effectively dissipate the generated heat [11,76,240,256]. The assessment of any proposed thermal management solution, such as different substrates used in this study, requires both physical modeling and complementary experimental thermal mapping for validation.
5.1.1. Thermal Modeling Challenges

Experimental thermoreflectance-based thermal imaging provides informative quantification of the temperature rise in active regions of GaN HEMTs near the surface of the GaN layer. Advanced measurements have been conducted to measure the transverse temperature distribution for cross-sectioned HEMT devices, but these types of measurements remain highly complex, may be highly uncertain, and require extensive sample preparation [103,104].

The measured temperature distribution provides a good insight into the thermal response of HEMTs but does not provide complete thermal property characterization for the complex devices. To offer a deeper understanding into the coupled interactions of electrical, thermal, and mechanical physics in microelectronic devices, full electro-thermal computational models are needed [184]. Numerical studies and multi-physical modeling help assess the contributions of different physical parameters and design layouts, but can prove to be vastly extensive, time consuming, and could prove challenging to develop and validate for several reasons, including those discussed next.

5.1.1.1. Coupled Physics

The self-heating in semiconductor devices is governed by heat generation that occurs under different electro-thermal mechanisms such as Joule, recombination, and thermo-electric heating. Even though Joule heating is the dominant self-heating mechanism in GaN HEMTs, the determination of the rate and distribution of heat generation in GaN HEMTs is complex and hard to predict without solving for the full electric field distribution. Moreover, the distribution of heat generation within the active device is also dynamic, depends on the electrical activation, and is susceptible to potentially irreversible changes in the physical properties of a device (e.g., aging) [95,159].

5.1.1.2. Uncertain Thermal Properties

Another challenge for thermal modeling is validating any obtained numerical results, as the defined input parameters may contain significant uncertainties that lead to an unfaithful
representation of the actual device response [138]. An example is the uncertainty that exists for thin-film thermal conductivities such as that of GaN. GaN thermal conductivity depends on thickness [277], crystal structure [119], impurity and dislocation density [133,278], doping concentrations [261], and fabrication method and quality and can vary between 100-400 W/m.K [157, 170, 227]. Temperature dependence, described by the power law in (5.1), can also result in a thermal conductivity decrease by more than 40% for GaN within a 100 K change [55,231].

\[
\kappa(T) = \kappa_{300\text{K}} \cdot \left( \frac{T}{300\text{K}} \right)^{\beta}
\]  

(5.1)

Some of these variables, such as the temperature-dependence of \( \kappa_{\text{GaN}} \), can be modeled but there remains other uncertainties that cannot be accurately predicted for all possible cases and processing factors, such as high-temperature contact annealing, and other electromechanical and thermo-mechanical interactions with potential bonded and seeded layers used in the particular thermal management approach. Such deviations from idealized properties are difficult to predict, and would thus require additional in-situ property measurements.

5.1.1.3. Large Variation in Physical and Time-Scales

The most challenging aspect in conducting a numerical analysis of complex microelectronic devices is the large variation in spatial and temporal scales. Finite element electrical and solid-state simulations can be performed in a localized domain within the active region of the device, which can be computed rather efficiently in 2D and even in 3D. By comparison, the thermal modeling of microelectronic devices must incorporate the full device structure that spans several spatial scales from the heat generation region (microscale) to the substrate heat sink (mesoscale). Moreover, the full 3D device structure presents further geometric complexities and different material adjacencies in three spatial dimensions. Without a specialized numerical engine that uses transient multi-grid meshing in both time
and space [205, 262], solving the full 3D heat transfer model would be extremely expensive and even impractical, especially if one is interested in transient behavior and/or parametric optimization.

For the above reasons, the validity of computational thermal analysis results depends largely on the accuracy of specified input parameters [256] and cannot be trusted without some form of experimental validation [138]. The study presented in this Chapter couples both experimental and numerical approaches to overcome their respective limitations and to offer a more complete thermal characterization of complex 3D devices, eventually providing a full 3D temperature rise profile (quantitative) as well as inferring important thermal properties of the device components (qualitative).

5.2. Methodology

5.2.1. Coupled Experimental and Numerical Approach

5.2.1.1. Method Description

The method described in this work combines high-resolution thermal maps obtained by the use of thermoreflectance imaging [129,177,178], with multi-grid numerical simulations to develop and optimize an experimentally-valid simulation model that is representative of the actual tested device (Section 5.2.3). The optimization is performed by the use of an iterative approach that refines uncertain model parameters, such as material thermal conductivities and heat generation peak location (within acceptable ranges) ultimately yielding a numerical thermal field that matches in the RMS sense the experimentally observed response. The outlined method of experimentally-driven thermal modeling has been previously presented [206] and demonstrated for a CMOS device [208].
5.2.1.2. Outcomes from Optimized Model

The main outcome from this approach is an optimized thermal model that provides a valid representation of the device’s actual thermal response. The optimized thermal model can provide several advantages summarized as follows to overcome thermal characterization challenges mentioned in Section 5.1.1.

1. The numerical 3D temperature distribution can be extracted for the full GaN HEMT structure. In this manner, the physics of the simulation model can be used to extend beyond any access limitations to the experimental optical methods and infer temperatures of embedded or optically-inaccessible regions.

2. Important uncertain thermal properties can be inferred from the optimized input parameters, such as effective conductivities of GaN, substrate, and interface layers.

3. Since the extracted temperature maps and thermal properties pertain to the actual thermal-mechanical state of the device measured at the particular instance in time when the device is experimentally characterized, the method can be used as a diagnostic tool to monitor time varying parameters and observe long-term effects such as aging and degradation [174,187]. These transient device dynamics are particularly important for HEMT devices which experience gradual degradation instead of sudden failure and are largely influenced by thermal loading [45, 245] and thermal interactions that are difficult to model in the design stage.

4. The optimized thermal model can be used as a platform to conduct parametric studies and ”what-if” scenarios, which would aid and fast track the design cycle and test different device configurations without the need for fabricating additional samples.

5.2.1.3. Selected Parameters for Optimization

The selection of the main variables for optimization is user-defined and is based on two criteria. First, the parameters chosen must have a non-negligible influence on the thermal response of the device, such as the properties of materials closest to the heating region,
Figure 5.1.  a) Schematic of HEMT device structure. The substrate is either Si, SiC, Diamond, or is etched b) intensity image (1024x1024 pixels) of the device under 100× magnification showing the Source, Drain and Gate contacts. Purple regions indicate the GaN drain and source channel where the thermal measurements are conducted c) device model built to scale in T°Solver®.
materials separating the heat source from the substrate, or low conductivity media that have a large contribution to the thermal resistance of the device. If the temperature rise is insensitive to other parameters, then the minor variations and uncertainties can be neglected, and a literature value can be assigned to those parameters excluded from the optimization method. For instance, the specific heat and density are not considered in this initial study since the device is activated in quasi-steady operation and is insensitive to transient effects.

The second criterion for a successful outcome is that the selected parameters for optimization must be known to exhibit some degree of uncertainty, either from a dynamic behavior or from a physical complexity. For that reason, device dimensions will not be optimized since they are available from the device design dimensions, from top surface reflectance images, and from cross-sectional electron microscopy (Figure 5.2).

Figure 5.2. x-SEM measurements showing thickness and composition of GaN/AlGaN stack.
5.2.2. Descriptions of Sample Devices

The studied devices are Gallium Nitride (GaN) high electron-mobility transistors (HEMTs) grown on different substrate materials: (111) Si, 4H SiC and CVD diamond, with an additional sample of commercial-grade GaN-on-SiC capped with a 0.5 \( \mu \text{m} \) thick nano-crystalline diamond (NCD) film, similarly to the devices reported previously by the authors [11, 243]. The device dimensions and spacings for the Source (S), Drain (D) and Gate (G) are extracted from the optical images and are presented in Table 5.1.

Additional details on the fabrication methods were presented in a previous study by the authors [239, 243]. The experimental thermal profiles and heating rates presented in the previous Chapter will be used to drive the model optimizations performed in this work.

5.2.3. Numerical Thermal Modeling

The representative thermal models of the different GaN HEMTs are built in TMX T°Solver®, which is a 3D thermal modeling engine that implements an automated self-adaptive multi-grid meshing, is therefore capable of handling large variations in scale in both space and time. This renders the engine extremely fast (at least two orders of magnitude faster than commercial packages) in solving the finite-volume heat transfer computations and allows the possibility of conducting iterative refinements and experimentally-guided optimizations on uncertain input parameters with prescribed uncertainties [206].

The epitaxial structure and the HEMT dimensions are built accurately to match the respective device (Figure 5.1c). Initially, literature values are defined for the input thermal values are defined for the input thermal

<table>
<thead>
<tr>
<th>Device</th>
<th>Gate Length</th>
<th>Gate Width</th>
<th>S-G spacing</th>
<th>G-D spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>GaN-on-Si</td>
<td>2.1</td>
<td>75</td>
<td>3.0</td>
<td>10.2</td>
</tr>
<tr>
<td>GaN-on-SiC</td>
<td>3.6</td>
<td>75</td>
<td>1.4</td>
<td>11.4</td>
</tr>
<tr>
<td>GaN-on-diamond</td>
<td>1.9</td>
<td>75</td>
<td>3.5</td>
<td>9.6</td>
</tr>
</tbody>
</table>

Table 5.1. Gate (G), Source (S) and Drain (D) dimensions (in \( \mu \text{m} \)) of GaN HEMTs on different substrates.
parameters of the simulation model, such as conductivity, density and specific heat [170]. The heat source is defined at the same rate as the corresponding power level measured experimentally. The source profile is set as an exponential distribution underneath the gate electrode, with a peak location at the drain-side edge. The top surface of the device is assumed to be adiabatic where no convection or radiation heat transfer occurs. The substrate base temperature is set at a uniform 20°C to match the controlled temperature of the sample-holding chuck. The device model is built on a thermally wide substrate domain to avoid any near-boundary effects in the lateral direction.

The numerical temperature results are extracted from the GaN channel regions within the Gate-Source and Gate-Drain spacing to correspond to the TR mapped regions. The same number of data points is extracted from the simulation results as the pixel count of the thermal images to be compared and matched within the optimization method. The iterative approach is stopped whenever the RMS error is minimized, indicated by a convergence of the parameters when $\Delta RMS$ between two consecutive iterations drops below a certain tolerance. The thermal properties estimates are finally inferred from the converged input parameters.

5.3. Reverse Modeling Results

5.3.1. GaN-on-Si Optimized Model

The material properties most significant in determining the thermal response of GaN-on-Si HEMTs are the thermal conductivity of thin-film GaN, which as discussed earlier, exhibits large uncertainties, and that of AlGaN, which in particular is hard to characterize since only thin, hetero-epitaxial (i.e., defective) AlGaN layers are possible to grow. The GaN proximity to the heat generation region makes it a determining factor in the lateral heat conduction, while the less conductive underlying AlGaN layers, separating the active region from the heat sink at the substrate, dominate the through-plane conduction.

An initial optimization run was performed for the GaN-on-Si device to optimize for the temperature dependent conductivity of GaN ($\kappa_{0,GaN}$ and $\beta_{GaN}$) and the effective conductivity
of AlGaN ($\kappa_{\text{eff},\text{AlGaN}}$), which directly relates to the thermal boundary resistance (TBR) between GaN and Si. The TBR between GaN and the different substrates can be obtained from the thickness ($h$) and effective conductivity ($\kappa_{\text{eff}}$) of the interface layer according to (5.2). The heat generation profile peak location was also optimized for to observe any peak migration effects [95].

$$TBR_{\text{GaN-subs}} = \frac{h_{\text{interface}}}{\kappa_{\text{eff,interface}}} \ (m^2K/GW) \quad (5.2)$$

The simulations were iteratively conducted until the difference between the numerical and experimental thermal maps was minimized in the RMS sense. The procedure was repeated for the different power levels to obtain a statistical distribution. The temperature profiles shown as carpet plots in Figure 5.3 indicate that the converged numerical solution matches closely the measured experimental map. The optimized thermal parameters and the RMS error between the experimental and simulation results are presented in Table 5.2 for the different activation power levels and summarized below:

$$\kappa_{0,\text{GaN}} = 213 \pm 12 \quad (W/m.K)$$
$$\beta_{\text{GaN}} = -1.28 \pm 0.1$$
$$\kappa_{\text{eff,AlGaN}} = 26.1 \pm 4.9 \quad (W/m.K)$$
$$TBR_{\text{GaN-Si}} = 26.8 \quad (m^2K/GW)$$

The obtained GaN thermal conductivity ($\kappa_{0,\text{GaN}}$) is close to the literature values reported at around 220 to 240 W/m.K for undoped GaN and close to reported power ratios ($\beta$) between 1.22 and 1.43 [170, 173, 231]. Moreover, the heat generation peak shift is not seen to have a drastic effect on the temperature profile, probably since the peak generation is already localized to the drain edge of the gate at the applied drain voltages. For that reason, the peak location was excluded from the subsequent optimizations.
Figure 5.3. Converged temperature profiles for the matched numerical and experimental thermal maps for the gate-drain access region of a GaN-on-Si device for a DC power level ($I_{DS} \times V_{DS}$) of 6.4 W/mm.

Table 5.2. Results of Converged Thermal Parameters of the Optimizations conducted for the GaN-on-Si HEMTs.

<table>
<thead>
<tr>
<th>Power (mW)</th>
<th>GaN</th>
<th>GaN</th>
<th>AlGaN</th>
<th>$Z_{Q, peak - Z_{gate}}$</th>
<th>$\epsilon_{RMS}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_{0, eff}$ ($\frac{W}{m \cdot K}$)</td>
<td>$\beta$</td>
<td>$\kappa_{eff}$ ($\frac{W}{m \cdot K}$)</td>
<td>(\mu m)</td>
<td>(°C)</td>
<td></td>
</tr>
<tr>
<td>226</td>
<td>209</td>
<td>-1.31</td>
<td>19.6</td>
<td>0.84</td>
<td>1.42</td>
</tr>
<tr>
<td>353</td>
<td>193</td>
<td>-1.29</td>
<td>21.6</td>
<td>0.60</td>
<td>1.24</td>
</tr>
<tr>
<td>482</td>
<td>210</td>
<td>-1.35</td>
<td>23.4</td>
<td>0.40</td>
<td>1.65</td>
</tr>
<tr>
<td>618</td>
<td>209</td>
<td>-1.11</td>
<td>28.2</td>
<td>1.20</td>
<td>2.41</td>
</tr>
<tr>
<td>772</td>
<td>226</td>
<td>-1.22</td>
<td>33.2</td>
<td>1.10</td>
<td>3.30</td>
</tr>
<tr>
<td>921</td>
<td>231</td>
<td>-1.41</td>
<td>30.5</td>
<td>-0.2</td>
<td>4.73</td>
</tr>
</tbody>
</table>
5.3.2. GaN-on-SiC Optimized Model

For the higher conductivity substrates (4H-SiC and CVD diamond), the chosen optimization parameters relate to the substrate thermal conductivity ($\kappa$) and the interface thermal boundary resistance (TBR) between GaN and the substrate. The intermediate layers such as AlN or AlGaN nucleation layers or SiN barrier layers have a considerable contribution to the overall thermal performance of the device. Any added resistances between the GaN active region and the substrate, either from intermediate layers or their interfaces, can drastically reduce the thermal conductance, and as a result potentially offset the added benefits of higher conductivity substrates [58].

Figure 5.4. Comparison of matched experimental and simulation results for GaN-on-SiC for an activation power level of 15.7 W/mm.

Optimizations were conducted for the GaN-on-SiC HEMT thermal model at each of the mapped power levels. The GaN thermal parameters were set to the values obtained from the GaN-on-Si device, while the conductivity of the 4H-SiC substrate ($\kappa_{4H-SiC}$) and the seeding layer of AlN ($\kappa_{eff,AlN}$) were chosen for optimization. A sample of the thermal maps for GaN-on-SiC (Figure 5.4) reveals a close matching to the observed experimental results. The converged thermal parameters resulting from the optimization runs (Table 5.3) summarized below, show a converged RMS error of less than 10% of the average channel temperature and 5% of the peak.
Table 5.3. Results of Converged Thermal Parameters of the Optimizations conducted for the GaN-on-SiC HEMTs.

<table>
<thead>
<tr>
<th>Power (mW)</th>
<th>4H-SiC κ (W/m.K)</th>
<th>AlN κeff (W/m.K)</th>
<th>εRMS (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>512</td>
<td>487</td>
<td>44.3</td>
<td>0.95</td>
</tr>
<tr>
<td>750</td>
<td>449</td>
<td>52.2</td>
<td>1.55</td>
</tr>
<tr>
<td>1030</td>
<td>448</td>
<td>47.0</td>
<td>1.81</td>
</tr>
<tr>
<td>1244</td>
<td>440</td>
<td>49.5</td>
<td>2.04</td>
</tr>
</tbody>
</table>

κ_{4H-SiC} = 456 ± 18 (W/m.K)
κ_{eff, AlN} = 48.3 ± 2.9 (W/m.K)
TBR_{GaN-SiC} = 2.07 (m².K/GW)

The thermal conductivity of the substrate was obtained at 456 W/m.K which in close agreement with the literature-reported values for 4H-SiC of around 450 W/m.K (measured [172, 203, 261]) and 423 W/m.K (calculated [127, 202]).

For the GaN-on-SiC HEMT, the two associated materials have a Wurtzite crystal structure with similar lattice constants \( a \) (3.07 Å for 4H-SiC and 3.16 Å for GaN), allowing GaN to be grown directly on SiC with minimal mismatch (3.5% compared to 17% for GaN-on-Si and 11% for GaN-on-diamond [264]). By eliminating the need for the AlGaN high-defect layers, the thermal conductance of the GaN-substrate interface is improved resulting in a low TBR when compared to GaN-on-Si.

The obtained AlN effective thermal conductivity value is also in agreement with reported literature values. Bulk AlN films have a conductivity \( \kappa_{eff, AlN} \) of around 280-320 W/m.K [115, 231], while thin-film AlN are typically highly defective and thus have a much lower \( \kappa_{eff, AlN} \) due to boundary phonon scattering, estimated at around 50 W/m.K [115].
experimental studies have measured a lower value between 10-20 W/m.K (15% and 25% Al) and attribute this decrease in conductivity to other phonon-scattering mechanisms such as oxygen impurity scattering [23, 115]. The large margin of uncertainty in $\kappa_{\text{eff, AlN}}$ with film thickness, Al fraction and grain size [115] further justifies the need for the reverse modeling approach for in-situ thermal property estimation presented in this work. The associated TBR is close in scale to the values reported in the literature (about 4-5 m²K/GW [55], and less than 10 m²K/GW [56]). For an initial estimate, the reached TBR indicates a high thermal conductance at the GaN/SiC interface, which is an order of magnitude larger than that at the GaN-on-Si interface.

5.3.3. GaN-on-diamond Optimized Model

For GaN-on-diamond HEMTs, the SiN adhesion layer presents a low-$\kappa$ barrier to heat dissipation from the active GaN region. In addition, the CVD diamond substrate grows in a granular pillar structure of increasing crystal size. This results in a low conductivity region near the SiN interface (near-interface disorder), as well as an anisotropic diamond conductivity [117] that is thickness dependent, which further adds to the uncertainty in the thermal conductivity of the substrate.

The optimized coupled approach was used on the GaN-on-diamond HEMT to extract the thermal resistance contribution of the substrate and the SiN adhesion layer. The model optimization was performed for both the in-plane and through-plane thermal conductivities of the CVD diamond substrate ($\kappa_{y,\text{eff,di}}$ and $\kappa_{xz,\text{eff,di}}$) and the thermal conductivity of the SiN layer ($\kappa_{\text{eff, SiN}}$). An effective conductivity was assumed for diamond to account for the depth variation of $\kappa_{\text{eff,di}}$, and for SiN to account for the interface thermal effects at the SiN interfaces.

The matched results for the temperature distributions (Figure 5.5) show close matching between the numerical and experimental results. The converged results for the GaN-on-diamond device (Table 5.4) are summarized as follows.
Figure 5.5. Sample temperature rise profile for the matched numerical and experimental results for GaN-on-diamond device 1 and device 2.

Table 5.4. Results of Converged Thermal Parameters of the Optimizations conducted for the GaN-on-diamond HEMTs.

<table>
<thead>
<tr>
<th>Power (W)</th>
<th>Diamond $\kappa_{y,\text{eff}}$ (W/m.K)</th>
<th>Diamond $\kappa_{zz,\text{eff}}$ (W/m.K)</th>
<th>SiN $\kappa_{\text{eff}}$ (W/m.K)</th>
<th>$\epsilon_{RMS}$ ($^\circ$C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>1445</td>
<td>1013</td>
<td>1.42</td>
<td>1.19</td>
</tr>
<tr>
<td>562</td>
<td>1406</td>
<td>975</td>
<td>1.40</td>
<td>1.61</td>
</tr>
<tr>
<td>873</td>
<td>1510</td>
<td>1039</td>
<td>1.45</td>
<td>2.39</td>
</tr>
<tr>
<td>1228</td>
<td>1520</td>
<td>1045</td>
<td>1.55</td>
<td>2.84</td>
</tr>
</tbody>
</table>
\[ \kappa_{y, \text{eff,di}} = 1470 \pm 47 \text{ (W/m.K)} \]
\[ \kappa_{xz, \text{eff,di}} = 1018 \pm 28 \text{ (W/m.K)} \]
\[ \kappa_{\text{eff, SiN}} = 1.46 \pm 0.06 \text{ (W/m.K)} \]
\[ TBR_{\text{GaN-di}} = 20.6 \text{ (m}^2\text{K/GW)} \]

From Table 5.4, one can observe that the converged effective through-plane and in-plane conductivity of the CVD diamond substrate are \( \kappa_{y, \text{eff}} = 1470 \text{ W/m.K} \) and \( \kappa_{xz, \text{eff}} = 1018 \text{ W/m.K} \), respectively. The value of \( \kappa_{xz, \text{eff}} \) is around 30\% lower than \( \kappa_{y, \text{eff}} \) due to the asymmetry and orientation of the columnar structure of the CVD diamond crystal. The effective thermal conductivity of the diamond substrate is observed to be lower than the approximate 2000 W/m.K estimate for crystalline diamond due to the near-interface crystal disorder [50, 117, 269] and increased grain boundary scattering [50, 263].

As to the interface effects, the converged value of 1.46 W/m.K for \( \kappa_{\text{eff, SiN}} \) corresponds to a TBR estimate between GaN and diamond of 21 m\(^2\)K/GW, which is close to the reported literature values between 12 and 25 for \( \approx 30 \text{ nm SiN by Sun et al [237]} \). The reported TBR is an order of magnitude larger than that of the AlN seeding layer in GaN-on-SiC, which is expected given the lower \( \kappa_{\text{SiN}} \).

5.3.4. Comparison between Different Substrates

By comparing the converged thermal parameters (TBR and \( \kappa_{\text{eff, subs.}} \)) for the three substrate configurations (Table 5.5), it is evident that the diamond substrate is thermally superior with 3× the thermal conductivity of 4H-SiC (1470 compared to 450 W/m.K). However, the TBR estimate between GaN and CVD diamond \( \mathcal{O}(10 \text{ m}^2\text{K/GW)} \) is an order of magnitude larger than that between GaN and 4H-SiC \( \mathcal{O}(1 \text{ m}^2\text{K/GW)} \). Even with a thinner layer of SiN (30 nm) compared to AlN (100 nm), the low conductivity SiN contributes
significantly to the total thermal resistance. Whereas for the case of GaN-on-SiC, the AlN seeding layer is of a higher $\kappa_{\text{eff}}$ and thus results in a lower TBR [175,264].

The TBR of GaN-on-Si remains the largest due to the large thickness (700 nm) and highly-defective (low-$\kappa$) AlGaN layers (26.1 W/m.K), both of which contribute to a high thermal resistance.

Even with the higher thermal conductivity of diamond, the GaN-on-diamond devices exhibited a higher peak thermal resistance than their GaN-on-SiC counterparts (3.25 vs. 2.70 mm.K/W, respectively). Diamond undeniably is a high thermal conductivity substrate material has a great cooling potential and is a giant leap from the thermal performance of GaN-on-Si, but improvements to the GaN-diamond epi are still needed to avoid faster degradation and justify the extra cost of GaN-on-diamond technology via a reduction in overall thermal management solution cost at the system level. Fortuitously, the reverse modeling approach and its results presented herein point to two main regions that can benefit from thermal improvements.

The first most notable setback is the high TBR of SiN which acts as a bottleneck to heat dissipation and may offset the added substrate thermal conductivity [57]. Reducing the thickness of the SiN layer [51], or using a higher conductivity material to substitute for SiN would improve the heat conductance within the device.

Another improvement strategy for the GaN-on-diamond structure would be to enhance Table 5.5. Comparing Substrate Conductivity and TBR for GaN on Si, SiC, and CVD diamond.

<table>
<thead>
<tr>
<th>Substrate Matl.</th>
<th>$\kappa_{\text{eff}}$ (W/m.K)</th>
<th>Interface Matl. h (nm)</th>
<th>$\kappa_{\text{eff}}$ (W/m.K)</th>
<th>TBR ($\frac{m^2 K}{GW}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>-</td>
<td>AlGaN 700</td>
<td>26.1</td>
<td>27</td>
</tr>
<tr>
<td>SiC</td>
<td>456</td>
<td>AlN 100</td>
<td>48.3</td>
<td>2</td>
</tr>
<tr>
<td>di.</td>
<td>1470$^a$</td>
<td>SiN 30</td>
<td>1.46</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>1018$^b$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$^a$ is the value for $\kappa_{y,\text{eff}}$ while $^b$ is for $\kappa_{xz,\text{eff}}$. 

75
the thermal conductivity of CVD diamond (larger crystals at the interface, lower boundary scattering rate). Improved thermal conductivity of nano-crystalline diamond (NCD) has been recently demonstrated via a graphoepitaxial approach on nanopatterned substrates [49, 106] Any approach that leads to an increase in $\kappa_{\text{eff,di}}$ closer to the value for single-crystalline diamond (around 2000 W/m.K) would directly translate to a measurable improvement in GaN HEMT power performance.

The two strategies aim at reducing different thermal resistance contributions and thus have a different effect on reducing the total $R_{th}$. To understand the benefit of the two above-mentioned thermal management strategies, a parametric study is conducted based on the optimized GaN-on-diamond thermal model. The reduction in average and maximum temperature rise in the GaN channel is observed for a set of TBR and $k_{di}$ input parameters. Separately, the effect of varying TBR and $k_{di}$ on the row-averaged temperature profile shows a larger sensitivity to changes in TBR (Figure 5.6), indicating that the TBR is the major contributor to the total thermal resistance.

A sensitivity analysis is finally conducted to assess the reduction in temperature rise (and $R_{th}$), and the percent reduction in self-heating reported for the studied range in TBR and $k_{di}$. The results show a noticeable reduction in $R_{th}$ when the boundary conductance and the substrate conductivity is increased (Figure 5.7a). The percent temperature reduction observed exceeds 30% for the lowest TBR value of 5 m².K/GW (Figure 5.7b). Since that TBR value may not be entirely achievable with the current state of technology, a sensitivity analysis helps provide some guidance on the effectiveness of the two thermal mitigation strategies. The reported sensitivity values validate the observations of Figure 5.6 and reveal a higher sensitivity of temperature rise and $R_{th}$ to TBR (0.3 – 0.39) when compared to the sensitivity to $k_{di}$ (0.13 – 0.18), as can be seen in (Figure 5.7b). This indicates that, given the current state of GaN-on-diamond device technology, a strategy of reducing the TBR would yield a better improvement in thermal management than the same percent increase in $\kappa_{\text{eff,di}}$ from enhancements to the CVD-diamond crystal structure.
Figure 5.6. Variation of temperature profile in GaN-on-Diamond HEMT for (a) the TBR range between 25 and 5 m²K/GW and (b) for the conductivity range $k_{di,xx}$ between 1000 and 1400 W/m.K with $k_{di,yy}=1.45 \times k_{xz,di}$. 
Figure 5.7.  

a) Results of parametric study showing the total thermal resistance for the studies values of diamond conductivities and TBR values, based on maximum and average GaN channel temperatures, and power level of 16.4 W/mm. The through-plane conductivity is taken as $1.45 \times k_{xz,di}$.  

b) The corresponding percent change in self-heating temperature rise with reported sensitivity values to changes in $k_{di}$ and TBR.
5.4. Conclusions

The study presented an advanced, numerically and experimentally coupled investigation into the thermal characteristics of GaN HEMT devices on different substrates intended to improve thermal performance. The unique approach utilizes experimentally-observed thermal maps of tested devices to optimize a corresponding simulation model and overcome both experimental and numerical thermal characterization challenges. Deep-submicron resolution thermoreflectance imaging provides accurate thermal maps in surface regions of the HEMT, with the ability to detect microscale hot-spots, while the advanced numerical engine provides the fast thermal modeling capabilities that render iterative methods feasible.

The reverse-modeling approach couples the experimental maps with fast simulation capabilities, allowing for the development of an experimentally-valid simulation model that can provide estimates for uncertain thermal parameters, expand 2D experimental optical imaging limitations to 3D numerical solutions, and serve as a platform for additional parametric studies. This feasibility of this approach was demonstrated by extracting and comparing estimates of important thermal parameters of GaN HEMT devices, such as the thermal conductivity of the substrate and the TBR of the interface layers, as well as to observe weak thermal effects such as the temperature dependent thermal conductivity of GaN and the anisotropic thermal conductivity of CVD-diamond.

By comparing the extracted thermal parameters of the GaN HEMTs, the Diamond substrate showed a superior thermal conductivity to SiC. However, as observed in the experimental study in the previous Chapter, the associated GaN-on-diamond HEMTs exhibit similar heating rates to GaN-on-SiC and can still benefit from several thermal improvement strategies. On the basis of the reverse modeling approach and its results presented in this paper, two main areas for potential thermal improvements can be identified: the SiN bonding layer with high TBR, and the near-interface crystal disorder which reduces the effective thermal conductivity of the CVD-diamond substrate.

In addition to providing experimental thermal maps, the combined approach provides a more complete thermal characterization of the GaN HEMT devices, especially in important
areas that are inaccessible to direct measurement. The main advantage of the combined approach is the ability to build an understanding of the thermal response based on the thermal properties of the components in the different GaN HEMT configurations, from which one can offer guidelines for future developments in GaN HEMT thermal management strategies.
Chapter 6

Compact 3D Thermal Model for VLSI and ULSI Interconnect Network Reliability Verification

In addition to advance power devices at the front-end of advanced integrated-circuits (ICs), the integration stage (back-end) that accompanies IC development also encounters thermal challenges that may accelerate network degradation and compromise reliability. The application possibilities for the presented advanced thermal investigative techniques are numerous and beneficial in many advance technologies. However, before delving into emerging technologies, it would serve a great benefit to use the developed techniques and offer an assessment for present investigative models and techniques that may require some update to meet the stricter thermal design criteria. Presented in this Chapter, the experimental and numerical techniques are applied to assess the thermal models for the integration level (BEoL) of ICs, expand the understanding into its thermal characteristics, and develop less conservative thermal model to meet the thermal challenges in device integration (VLSI and ULSI) imposed by device scaling and increasing processing power demands.

In this Chapter, the thermal characteristics of interconnects within the BEoL are investigated using thermoreflectance imaging, and the influence of different metrics (interconnect shape factor, insulation thickness, passivation thickness) is investigated. The available reduced models for thermal conductance of BEoL interconnects are assessed for their accuracy and a more accurate compact thermal model is developed accordingly for the quick and accurate thermal characterization of self-heating in Back-End-of-Line (BEoL) interconnect network for reliability assessment. The model offers a more accurate thermal tool compared to the currently used conservative 2D reduced models, and is then expanded to a three dimensional model (3D) using a resistive network analysis. Both 2D and 3D reduced models are developed and verified against numerical simulations of different configurations and further
validated against thermoreflectance measurements of specially fabricated samples devices.

The compact model ultimately provides quick thermal characterizations for vast ICT networks at a fraction of the analysis time and computational cost required for full numeric simulations. When integrated with electrical layout designs, the model would serve as a thermal check on VLSI and ULSI networks to identify thermally critical regions where overheating could lead to degradation and reliability issues.

6.1. Introduction

The Back-End-of-Line (BEoL) is the stage comprising all of the interconnects (ICTs) that are responsible for signal relay between, and power delivery to all devices within an Integrated Circuit (IC). As electrical design dictates the performance of microdevices in terms of frequency, power rating, and efficiency, the reliability and expected lifetime of the network remains largely dependent on its thermal performance [17, 247]. In fact, operating temperatures play a major role in accelerating several degradation mechanisms, especially Electromigration (EM) [21, 22, 146, 236]. For that reason, the thermal performance of a device is a factor that limits the electrical activation levels, and a crucial part of IC design is ensuring that the vast BEoL network, which may consist of millions of ICTs, can tolerate the required activation levels with reliable operation and minimal thermal degradation [7, 183, 222].

6.1.1. Electromigration Failures and the Degradation of IC BEoL Networks

Material transport in conducting media can occur under different conditions such as concentration, temperature, and stress gradients. When caused by a high electric current, the atomic transport mechanism is called Electromigration. The exchange in momentum causing the material transport occurs when the flowing electrons collide with the metallic ions in the conducting medium, carrying the loosened ions in the direction of electron flow. Consequently, the atoms diffuse under EM conditions and are depleted from the cathode and accumulate at the anode, creating voids and hillocks at the opposite conducting lead ends (Figure 6.1, and thus leading to a resistance increase as the network degrades or eventually
fails from short-circuits or circuit-breaks. One of the major concerns in IC ICT networks is the degradation that occurs due to Electromigration (EM).

![Image of interconnect showing hillock and void formation due to material transport under Electromigration. (Source: Kyung-Hoae 2011)](image)

Electromigration failure analysis was first described by J. R. Black who presented an empirical model in 1960 as shown in (6.1). The equation relates the mean-time-to-failure ($MTTF$) to a geometric parameter ($A$ cross-sectional area), an activation parameter ($J$ current density), and to material properties ($E_a$ activation energy) and finally to the average lead temperature $T$. The square term is later replaced by $J^n$ where $n$ depends on the different depletion or growth failure modes. $k_B$ is the Boltzmann constant.

$$MTTF = \frac{A}{J^2} \exp \left( \frac{E_a}{k_B T} \right)$$  \hspace{1cm} (6.1)

The empirical model of Black has been proven as an oversimplified model where the fitting parameters for activation energy $E_a$ and $n$ obtained at stress conditions differ from those at normal operation. More accurate models built on Black’s equation, adding the effect of hydrostatic stress, current density and temperature [146]. Other reliability models offer physics-based EM analysis where the void initiation and growth time are modeled under the effects of electrical activation, material properties, layout design and residual stresses [112]. The model determines the void initiation time based on the stress distribution in the ICT network. The void growth then determines the power-ground (p/g) network deterioration where the voids grow and the resistance ratings increase and is updated at each subsequent
time step. The failure consideration criteria is taken when the IR (voltage) drop exceeds a certain tolerance. The effects of in-chip temperature distributions are studied in a later work [111] and are seen to drastically vary the failure time, and the failure location, as local ICT temperatures can be higher or lower than the average die temperature depending on the MET layer and the location of the ICT. Higher temperatures can accelerate EM, while overestimated temperature can wrongly predict void initiations [111]. Thermal stresses are reduced at higher temperatures since the material contains residual stresses from the fabrication at higher zero-stress temperature $T_{zs}$. Higher temperatures reduces the temperature difference ($T - T_{zs}$) thus reducing thermal stresses and retarding EM failures. The observed effects stress the importance of within-die temperature distributions on lifetime estimations and the crucial need for accurate reduced thermal models for reliability assessment of IC’s.

Different materials require different activation energies for the onset of Electromigration, with Copper (Cu) having higher diffusion resistance than Aluminum (Al), with $E_a$(Cu) = 0.8 – 1.2eV for dominant surface diffusion, as compared to $E_a$(Al) = 0.7eV. Refractory metals such as Tantalum (Ta), Titanium (Ti), and Tungsten (W) are used for vias to connect ICTs at different levels of the BEOI since they have high atomic bonding energies and can thus serve as barriers to EM diffusion. Another technology strategies that improves the atomic bonding of conducting metals that reduces EM vulnerability is introducing a 0.5% Cu blend into Al ICTs.

Geometry parameters such as the cross-sectional area ($A$) and length ($L$) are also key factors in EM reliability models. Increasing the area of a conductor or reducing the current intensity reduces $J$ and thus extends lifetime. Interconnect length, if below a certain limit denoted as Blech limit, can render segments immune to EM effects [22]. Other design strategies can also be implemented to reduce EM effects. For example, reducing sharp corners may reduce current crowding effects at the corners, thus reducing localized $J$. Moreover, reversing the current stress reduces EM effects and can even result in self-healing [72], which is why connectivity ICTs carrying bidirectional electric current are less susceptible than their power deliver counterparts that carry unidirectional currents.
The above conditions can be determined in the choice of technology and design for the BEoL stage. Temperature, on the other hand, is a crucial factor that is difficult to model and that can accelerate EM since thermal energy contributes to part of the activation energy needed for the onset of EM, causes a positive feedback as resistance increase due to degradation further increases self-heating and feeds back EM and accelerated void growth. This exponentially reduces EM lifetime where a 5 Kelvin temperature rise could reduce the lifetime by 35% and a 10 K by 58% for Aluminum as estimated by Black’s equation [21].

The ICTs in the BEoL are separated from the substrate by an electrically insulating inter-layer dielectric (ILD) medium, which is required for signal isolation to reduce RC delay, cross talk and dynamic power consumption [15, 251]. The low electrical conductivity ($\sigma$) material also exhibits a poor thermal conductivity ($\kappa$) that hampers heat dissipation within the BEoL, rendering active ICTs susceptible to considerable self-heating [17]. This thermal side-effect is becoming a major concern for IC reliability, especially with the ongoing trend in electronic miniaturization and ever-increasing power requirements. For these reasons, accurate and quick thermal characterization, as well as innovative thermal management strategies are required for the BEoL stage to ensure reliable operation [247]. With the continuing trend in device down-scaling and increasing power demand, and with transition from monolithic ICs to 2.5 and 3D integration, the undesired thermal heating effects are further exacerbated leading to more constrained device performance. This raises the need for more advance thermal characterization tools, both experimental and numerical, to ensure reliable device operation for an acceptable lifetime. The in-chip temperature distribution is challenging to determine and requires either time-consuming extensive 3D modeling, or over-simplified reduced thermal models that prove to be inaccurate or overly conservative.

In this chapter, a reduced three-dimensional (3D) thermal model is developed to estimate the temperature rise of active ICTs in BEoL networks. The model is assessed for its feasibility for large scale numerical thermal analysis and for its accuracy through experimental and numerical validations. Initially, available 2D models are presented and assessed for accuracy in Section 6.2. Aiming to improve on accuracy of the oversimplified 2D models, a combined
2D model is developed to estimate the 2D conductance from a thermally long interconnect. In Section 6.5, the studied reduced models and the newly developed models are compared to numerical simulations and validated with experimental thermal mapping of representative device samples. In Section 6.3, the 2D reduced model is extended to a 3D model that considers entrance effects, conduction through the vias, and mutual heating between multiple ICTs. The developed 3D reduced model makes it possible to obtain fast and accurate thermal assessments of a large-scale ICT network that would otherwise require vast and impractical computational efforts. A case verification study with numerical and experimental results is presented in Section 6.6.

6.2. 2D Model for Thermal Conductance from an Isolated Interconnect

Given the long and narrow shape factor of ICTs, it is logical to begin with an initial two-dimensional (2D) analysis of the heat transfer from a thermally long heat source. In simple terms, neglecting axial heat conduction, the temperature rise in an active ICT under Joule heating mainly depends on the heater/resistor material and geometry, the activation level, and the surrounding dielectric material. The ICTs considered in this study are thin conductive lines of width \( w \) and thickness \( t \), and are located in a BEoL with an insulation thickness of \( t_i \) above the substrate and an upper-passivation thickness \( t_p \) (Figure 6.2).

The temperature rise upon activation is determined from the heat generation rate \( Q'_{gen} \) and the total thermal resistance between the ICT and the substrate’s heat sink. The effective thermal resistance is directly related to the effective conductivity \( \kappa_{eff} \) of the dielectric \( (R_{th,eff} = t_i/(\kappa_{eff}w)) \). \( Q'_{gen} \) is determined from the electrical activation, based on the sheet resistivity of the ICT material \( (\rho'_{0}) \) and the activation current density \( (J) \). The temperature rise can be obtained from both the electrical and thermal subsystems (Figure 6.2) according to (6.2), and an allowable activation level, given an allowable temperature rise limit, \( \Delta T_{allow} \), can be obtained using (6.3).

\[
\Delta T = R_{th} \cdot Q'_{gen} = \rho'_{0}J^2 \frac{t_i}{\kappa_{d,eff}} \quad (6.2)
\]
6.2.1. Reduced Mathematical Models for Effective Conductivity

6.2.1.1. 1D Model

The simplest model to calculate $\kappa_{\text{eff}}$ is to consider only a rudimentary one-dimensional (1D) heat flow across a rectangular region of the dielectric underneath the ICT. The dielectric thermal resistance to the substrate is calculated as $R'_{\text{th,1D}} = \frac{t_i}{\kappa_d w}$. This is the most conservative model as it does not consider heat spreading in the dielectric layer, an effect that can improve conductance and reduce the operating temperature.

A better representative heat conduction models would consider 2D heat flow and account for lateral heat spreading beneath the ICT, and can be derived by the use of mathematical domain transformation models [20, 88, 222]. The 2D models are less conservative, yielding a higher conductance and a lower $R'_{\text{th,2D}} = \frac{t_i}{\kappa_{d,\text{eff}} w}$ with an effective dielectric conductivity ratio $\frac{\kappa_{d,\text{eff}}}{\kappa_d} > 1$. 

$$J_{\text{allow}} = \sqrt{\frac{\Delta T_{\text{allow}} \kappa_{d,\text{eff}}}{\rho'_0 t_i}}$$ (6.3)
6.2.1.2. Bilotti 2D Model

The 2D heat spreading effect was studied by Bilotti [20] and is based on the mathematically derived shape factor for conduction between a rectangular heat source located in an insulating medium at a distance from the substrate surface plane (Figure 6.2.a with \( t_p = 0 \)). Bilotti’s equation shown in (6.4) can accurately represent heat conduction between a thermally long ICT and the substrate surface, both considered at uniform temperatures. The model however, does not consider heat flow from the sides and the top of the ICT when it is embedded under some thickness of passivation, as is the case for BEOl ICTs which are present in a multi-layered stack. Considering only the bottom region for heat dissipation overestimates the resistance of the dielectric and thus leads to a higher estimated temperature rise and a more conservative model.

\[
\frac{\kappa_{d,eff}}{\kappa_d} \bigg|_B = \left(1 + B^* \frac{t_i}{w}\right)
\]

(6.4)

6.2.1.3. Andrews 2D Model

The mathematical model of Andrews’ [12] described by (6.5) more accurately represents the case of a passivated ICT since it considers conduction from a rectangular source placed in a semi-infinite medium (Figure 6.2.a with \( t_p = \infty \)). Due to the semi-infinite passivation, the top surface is modeled as adiabatic and thus the model provides the effective conduction to the substrate surface. This adiabatic assumption is valid since the convection and radiation heat losses at the top surface are negligible due to the typically moderate temperature rises.

\[
\frac{\kappa_{d,eff}}{\kappa_d} \bigg|_A = \alpha \frac{t_i}{w} \left[ \log_{10} \left(1 + \frac{t_i}{w}\right) \right]^{\beta_1} \left(\frac{w}{t}\right)^{\beta_2}
\]

(6.5)

Both equations are based on geometric parameters of the system, width and thickness of the resistor (\( w \) and \( t \)) and on the thickness of the dielectric insulation (\( t_i \)) with \( \kappa_{d,eff}/\kappa_d = f(w, t, t_i) \). The equations can be expressed in non-dimensional form as \( \kappa_{d,eff}/\kappa_d = f(w', t') \) where \( w' = w/t_i \) and \( t' = t/t_i \). The Bilotti parameter (\( B^* \)) is obtained for large excess width.
of insulation, which is applicable for our assumption of a narrow resistor being much smaller than the width of the oxide layer. The coefficients in the Andrews equation \((\alpha, \beta_1, \beta_2)\) are determined to fit our simulations.

### 6.3. Reduced 3D Thermal Model for an Interconnect Network

As full thermal modeling may require extensive computational effort and expense, reduced 3D models provide quicker and more effective thermal analysis method for microelectronic devices that contain standardized dimensions and structures, such as Silicon-on-Insulator (SOI) devices \([47, 88, 235, 271]\), FinFET structures \([46]\), and some BEOl layouts \([111, 246, 251]\).

However, as the complexity and variety of device designs increase, custom reduced modeled are required to be tailored to each individual structure; otherwise, more rigorous extensive computational modeling would be required. In the case of BEOl networks, the sheer number of ICTs renders computational methods infeasibly expensive. For that reason, reduced model have been developed in an effort to conduct a quick thermal analysis of active power grid (P/G) networks for reliability assessment.

Teng et al. \([246]\) developed an electromigration reliability tool (iTEM) that first calculated the chip temperature using a finite difference method and then estimated ICT temperature using a reduced model based on a series solution for \(\kappa_{d,\text{eff}}\) \([88]\). For their EM assessment tool, Huang et al. \([111, 112]\) estimated the temperature distribution within the BEOl by utilizing a thermal model based on cuboidal thermal cells with effective local thermal properties. However, they then report the temperature as an average in a region around the ICT, which does not capture actual temperature rise of the interconnects themselves. For more densely packed ICTs, this method would fail to represent individual ICT, and if one opts to then use finer numerical grids, the method would become similar to a full FEA, which is numerically infeasible for the previously stated reasons.

Chen et al. \([40]\) conducted computational studies and used reduced thermal resistance networks to model ICT heating between isolated ICTs (2D model), parallel coupled ICTs,
and cross-coupling of power and signal lines, with results being verified with numerical results. Chiang et al. [251] analyzed the length effect and the effect of the separation of vias by the use of the ”lossy wire model” and a 2D reduced model that considers a new expression for the 2D heat spreading $\kappa_{d, eff}$. The preceding two studies presented a reduced model for ICT configurations at two BEoL levels but are not general enough to address more complex configurations. Moreover, no experimental work was performed to ascertain the validity of the proposed models.

In the analysis of the previous section, only thermally-long ICTs were addressed, with no consideration of longitudinal heat conduction. In addition to through-plane and transverse heat dissipation, axial conduction also contributes to the thermal response of active ICTs and results in an axial temperature distribution, mainly observed at the edges of ICTs. Moreover, the presence of connective vias at the ICT ends also provides additional thermal paths between ICTs at different BEoL levels, and results in additional heat transport effects. As a result, in the following subsection, a reduced 3D model is developed based on the ”lossy wire model” to consider the additional thermal effects in the axial direction, namely entrance effects and thermal coupling through vias and cross configurations.

6.3.1. 2+1 D Heat Conduction: The Lossy Wire Model

The lossy wire model (LWM) is widely accepted for the thermal design and analysis of structures with narrow shape factors, such as heat fins and extended surfaces [4, 134] that experience some heat loss to the surroundings. The temperature profile along a lossy wire can be derived by evaluating the balance between heat lost at a differential length ($dx$) and the axial heat conducted axially. This steady energy balance governs the temperature distribution along any ICT of index $j$ and is given in (6.6).

$$\frac{d^2T}{dx^2} - \frac{T - T_0}{\chi_H^2} = -\frac{\rho_m J^2}{\kappa_m}$$

(6.6)
The heat conduction along the wire is expressed in the first term, the heat lost to the surroundings in the second, and the heat generation rate in the third. The second term is the effective heat lost to the surroundings and is expressed in terms of $\lambda_H$, a characteristic thermal length. The equation is subject to a boundary condition of fixed temperatures at both ends of the ICT, $\Delta T_{x=0}$ and $\Delta T_{x=L}$. The third term represents the Joule-heating rate upon electrical activation at $J_{RMS}$ and is used to solve the non-homogeneous problem with zero boundary conditions. The homogeneous solution is obtained by assuming the third term to be zero, which gives the boundary-driven conduction as a function of boundary conditions at both ends. Superposing both solutions gives the temperature distribution along an ICT of length $L$ as presented in (6.7).

$$\Delta T(x) = \Delta T_{\text{max}} \left( 1 - \frac{\cosh\left(\frac{x - L/2}{\lambda_H}\right)}{\cosh\left(\frac{L/2}{\lambda_H}\right)} \right) + \Delta T_0 \frac{\sinh\left(\frac{L-x}{\lambda_H}\right)}{\sinh\left(\frac{L}{\lambda_H}\right)} + \Delta T_L \frac{\sinh\left(\frac{x}{\lambda_H}\right)}{\sinh\left(\frac{L}{\lambda_H}\right)}$$ (6.7)

The temperature profile of any active lossy wire is defined in (6.7) in terms of only the three following parameters:

6.3.1.1. The Thermal Entrance Length

($\lambda_H$) is a characteristic dimension that determines the distance of the entrance regions that is affected by a node temperature boundary condition. The thermal entrance length depends mainly on the ratio of axial heat conduction (represented by the metal thermal conductivity $\kappa_m$) and the heat lost to the surrounding dielectric (represented by the effective conductivity $\kappa_{d,\text{eff}}$ obtained from the reduced 2D model of Section 6.2):

$$\lambda_H = \left( \frac{\kappa_m T L}{\kappa_{d,\text{eff}}} \right)^{1/2}$$ (6.8)

The thermal entrance length $\lambda_H$ combines the balance of axial heat conducted and the heat lost to the dielectric and describes the end effects within an ICT. The effective conductivity ($\kappa_{d,\text{eff}}$) of the dielectric medium is obtained from the reduced models for 2D heat
conduction presented in the previous section.

6.3.1.2. The Maximum Attainable Temperature

\((\Delta T_{max})\) is the temperature attained by a thermally long ICT at a location away from the edge \((\approx 5\lambda_H)\) (Figure 6.3). \(\Delta T_{max}\) is obtained by solving the non-homogeneous energy balance equation with zero boundary conditions (6.6):

\[
\Delta T_{max} = \frac{p_0 J_{RMS}^2 \lambda^2}{\kappa m}
\]

6.3.1.3. The End Temperatures

\((\Delta T_0\) and \(\Delta T_L\)) are boundary conditions for the 1D energy balance equation. In the present model, as neighboring ICTs connect to each other and to vias and other features, a set of common nodes will arise whose temperatures \((\Delta T_n)\) are the unknowns that will be solved for by the thermal resistance network analysis described in the following section.

6.3.2. Thermal Resistance Network Analysis

Constructing the temperature distributions for several ICTs of index \(j\) requires first to solve for \(\lambda_{H,j}\), \(\Delta T_{max,j}\), and the temperature \(\Delta T_n\) of the connecting nodes. \(\Delta T_{max,j}\) and \(\lambda_{H,j}\) can be obtained from technology and activation parameters by the use of the combined 2D model of Section 6.2. However, the values for \(\Delta T_n\) are unknown and result from heat transfer between different connected features such as ICTs, vias, and/or cross configurations. Therefore, to model the heat transfer between the different components and configurations, a resistive network analysis is used in this study to calculate the node temperatures \(\Delta T_n\), which requires a resistance representation of heat transfer between the nodes to all other connected components, such as ICTs, vias, and cross configurations.
6.3.2.1. ICT Entrance Thermal Resistance

The first resistance that is required is the resistance to heat transfer from a node \(n\) to its connected ICT \(i\) (Figure 6.3), denoted by "entrance thermal resistance" \(R_{\text{th,n,j}}\). This resistance can be obtained from the temperature gradient \(dT/dx\) as follows.

\[
R_{\text{th,n,j}} = \left( \frac{\kappa_{m,j} A_{cs,j}}{\Delta T_n - \Delta T_k} \frac{\partial T_j}{\partial x} \right)_n^{-1}
\]  

(6.10)

The temperature gradient \((\partial T/\partial x)\) can be obtained by differentiating (6.7) at \(x = 0\) to yield:

\[
\frac{\partial T_j}{\partial x} \bigg|_n = \Delta T_{\text{max},j} \left( \frac{1}{\lambda_{H,j}} \tanh \frac{L_j}{2\lambda_{H,j}} \right) - \Delta T_n \left( \frac{1}{\lambda_{H,j}} \coth \frac{L_j}{\lambda_{H,j}} \right) + \Delta T_k \left( \frac{1}{\lambda_{H,j}} \operatorname{csch} \frac{L_i}{\lambda_{H,j}} \right)
\]  

(6.11)

In general, the resistance is expressed between the two node temperatures \(\Delta T_n\) and \(\Delta T_k\) of an ICT of index \(i\) using (6.10) and (6.11). For thermally-long ICTs \((L > 5\lambda_H)\), the \(R_{\text{th,n}}\) reduces to the form given in (6.12), and heat conduction is considered between the node temperature \(\Delta T_n\) and \(\Delta T_{\text{max},i}\) (Figure 6.3).

\[
R_{\text{th,j,long}} = \left( \frac{\kappa_{m,j} A_{cs,j}}{\lambda_{H,j} \tanh \frac{L_j}{\lambda_{H,j}}} \right)^{-1}
\]  

(6.12)

6.3.2.2. Via Thermal Resistance

The heat conduction through a via is considered as 1D flow with a thermal resistance calculated by the use of the via material and dimensions as follows.

\[
R_{\text{th,via}} = \frac{h_{AB}}{k_{\text{via}} A_{\text{via}}}
\]  

(6.13)

\(h_{AB}\) is the height of the via extending between the two ICTs, \(k_{\text{via}}\) is the thermal conductivity of the via material, and \(A_{\text{via}}\) is the cross-sectional area of the via which depends on the
Figure 6.3. Schematic showing a) thermally long \((L/\lambda_h > 5)\) and b) thermally short \((L/\lambda_h < 5)\) ICT with the considered entrance thermal resistance at the nodes and the associated temperature distribution for a given \(\lambda_H\). Zero boundary condition is considered \(\Delta T_n = \Delta T_0 = 0\) and \(\Delta T_k = \Delta T_L = 0\) at both ends \(x = 0\) and \(x = L\). Note that the temperature distribution in a thermally short ICT does not reach \(\Delta T_{max}\).

dimensions of the vias, and the number of vias present for a given ICT width, all of which are predefined technological (i.e. fabrication) parameters.

The heat conduction through the via material is considered as the primary heat transfer path since the medium is of higher conductivity than the dielectric. The secondary heat conduction through the dielectric is considered by evaluating the effective conductance \(G_{d,\text{via}}\) between two ICT ends as given in (6.14).

\[
R_{th,\text{via},d} = G_{d,\text{via}}^{-1} = \frac{3h_{AB}}{k_d w_A w_B} \tag{6.14}
\]

Since the two described heat conduction paths are parallel, the resistances in (6.13) and (6.14) are handled in a parallel configuration to obtain the overall resistance across the via between the two ICTs.
6.3.2.3. Cross Configuration Thermal Resistance

Another configuration that can influence the temperature of ICTs is the cross-coupled mutual heating that occurs between two transversely running ICTs. The cross-coupling can lead to additional local heating load at the cross location when both ICTs are active, or possibly to a desirable cooling effect if one of the ICTs is inactive.

The heat exchange between two ICTs in a cross configuration is modeled and added as a thermal resistive connection in the present model. The cross resistance $R_{\text{th,cross}}$ given by (6.15) was analytically obtained by Cheng et al. [48] by integrating the conductance as a function of the view factor along both ICTs.

\[ R_{\text{th,cross}} = \frac{3}{2\pi \kappa_d} \frac{h_{AB}}{(w_A + t_A)(w_B + t_B)} \]  

(6.15)

$h_{AB}$ is the vertical separation between the two ICTs $A$ and $B$, $w$ and $t$ are the width and thickness of each ICT, respectively, and $\kappa_d$ is the thermal conductivity of the separating dielectric.

The resistive connection is placed at the cross junction between the two ICTs. Another configuration presented by [48] is seen to give more accurate results is a six-resistor model (Figure 6.4). The six-resistor network is equivalent to the derived resistance of the single resistor model, such that $R = 1.5R_{\text{cross}}$, where $R_m = l_c/\kappa_m \cdot w \cdot t$ is the axial thermal resistance in the conducting ICT segment of dimensions $(l_c \times w \times t)$ within the cross region. $R = 1.5R_{\text{cross}}$ is the cross resistance, and $R_j$ is the ICT entrance thermal resistance that connects the cross configuration to the full resistance network.

The heat generation in both ICT segments located in the cross region are also included in the thermal resistance analysis. The heat generation is calculated from the Joule heating per unit length ($Q'_x$), scaled by each segment’s length, and then distributed equally upon the three nodes of each ICT. The ground resistance connects the lower ICT to the substrate heat sink and is calculated by the use of Bilotti’s model since only conduction from the bottom surface is considered.
Figure 6.4. Schematic showing the resistive network configuration for a) Single Resistive Cross model b) Six Resistors Cross Model based on [40]. The value of R is such that the equivalent network resistance is equal for the two ICTs in both configurations. c) Resistance network schematic showing the six-resistor model [48] connected to the ICT segment ends.

The cross configuration thermal network (Figure 6.4) is solved for the unknown temperatures (nodes 2, 4, 5, and 7) given a) the temperatures of the four connected segments in ICTs A and B, and b) the heat generation in the active regions of A and B. The temperature profile within the cross is then reconstructed from the three nodes within each ICT: nodes 2, 3, 4) in B, and nodes 5, 6, 7 in A (Figure 6.4).

6.3.3. Iterative Approach to Solve Resistance Network

In general, the entry thermal resistance $R_{th,i}$ and node temperature rise $\Delta T_{n,i}$ are coupled (6.11) and cannot be solved separately. Therefore, to calculate $\Delta T_n$, an iterative method is implemented where ICTs are initially assumed as being thermally-long, in which case $R_{th,long}$ can be determined independently of node temperatures. This allows an initial thermal resistance network to be built with $R_{th,long}$ and $\Delta T_{max}$, all of which are known beforehand.

The iterative method then proceeds as follows:

1. Assuming thermally-long ICTs for the first iteration, the resistance network is built with $R_{th,j,long}$ (6.12), $\Delta T_{max,j}$ (6.9), and the corresponding $R_{th}$ for vias (6.13) or cross
2. The thermal network is solved for an estimate of $\Delta T_n$.

3. The resistive network ($R_{th,n,j}$) is updated based on the newly updated $\Delta T_n$.

4. In the next iteration, the updated resistive network is solved for $\Delta T_n$, and the process is repeated by updating $R_{th,n,j}$ and solving for $\Delta T_n$ until convergence (small change in $\Delta T_n$).

After the above iteration process is complete, the node temperatures $\Delta T_n$, the maximum temperature rise $\Delta T_{max,j}$ and the thermal entry length $\lambda_{H,j}$ are used to construct the full axial temperature profile $\Delta T_j(x)$ in all the ICTs.

6.3.4. Model Validation with Simulations and Thermal Imaging

The developed 2D and 3D reduced models are verified against thermal simulations to ensure that the results from reduced modeling can closely estimate those from full 3D thermal simulations. The software used to conduct the thermal simulations in this study is T°C Solver® presented in Chapter 7.2. The sample devices are build to scale in the software with material properties are defined as literature reported values. The heat generation rates are set to match the experimental heat generation which corresponds to the electrical activation power.

The conducting material is considered to be of a higher thermal conductivity than the surrounding dielectric material [255], and thus any lateral temperature variation is considered to be negligible. The simulation model faithfully represents the location and dimensions of the ICTs within the stack, the electrical parameters and the thermal parameters of the ICT material. The effect of the chemical mechanical planarization (CMP) dummy fill pattern is shown in a previous study to affect the heat conductance within the BEOl stack and result in a passive cooling effect [77]. The CMP fill is therefore modeled with the same effective thermal properties as used in [77].

Experimental thermal maps of activated sample devices of similar test structures are initially used to validate the numerical simulation model. thermoreflectance (TR) imaging
system [207, 212], which provides deep-sub-micron resolution thermal maps. The optical imaging technique provides accurate thermal maps at sub-micron resolution scales from which the temperature distribution along the ICT features can be extracted. This approach is taken, in lieu of using the experimental maps to validate the reduced model, for the following reasons: 1) the actual thermal properties may contain uncertainties and unknown interface effects that may affect the agreement between the reduced model and the experimental results, and 2) the same input parameters can be defined for both the simulation and the reduced model where any discrepancy would be the result of reduced modeling and not the result of mismatch in defined model input parameters. The methodology and samples will be presented in the respective Sections for the 2D and 3D results.

6.4. TDTR Conductivity Measurements of BEoL Thin-film Materials

Before presenting the results of the thermal investigation into the thermal characteristics of the integration stage (i.e., BEoL) of ICs, a property measurement study is presented to characterize the materials that are present in the BEoL. This will provide the thermal properties that will be utilized in the developed models and in the simulations to follow.

6.4.1. Studied Materials

The materials chosen for this study are widely used in standard components in the fabrication and operation of the BEoL stage. The main component of the P/G network is the current carrying interconnect network patterned at the multiple metallization (MET) levels of the BEoL. A high electrical conductivity material is ideal for that objective to efficiently conduct electric signals with minimal resistive and IC power losses.

Aluminum (Al) is a prime candidate for an interconnect material since it has one of the highest electrical conductivities among metals. However, since Al can suffer from electromigration (EM), which is the material transport that occurs under high electrical current densities, Aluminum Copper has emerged as an improved replacement with better resistance to EM since the addition of Copper in small concentrations (0.5 to 5 weight%) improves the
Al grain adhesion. EM is accelerated at elevated temperatures, and so a combination of improved interconnect material design and strict thermal considerations are required to reduce EM induced void formations and P/G network degradation.

Polysilicon (hereafter referred to as Poly or PolySi) is another electrically conductive material that is used for gates, local interconnections, and microresistors in the BEoL network. The Poly layer is deposited in the first BEoL layer and since it has a higher thermal resistivity than AlCu, it contributes more to Joule heating that could elevate local BEoL temperatures when subjected to high current densities. The high electrical resistivity and the proximity to active devices at the front-end demands adequate thermal consideration for Poly components.

Electrical isolation between the interconnects of the BEoL is required to reduce resistive-capacitive (RC) delays, crosstalk, and dynamic power consumption [15]. For that reason, interconnects are placed within an insulating medium. Silicon dioxide (SiO$_2$) is the most versatile and easy to process insulating material as it is based on the ubiquitous Silicon and is compatible with the metal–oxide–semiconductor (MOS) fabrication process. Tetraethyl orthosilicate based silicon oxide (TEOS) is a highly insulating form of SiO$_2$ with very high conformity. An advantage of using TEOS SiO$_2$ is that it can be processed at lower temperatures compared to standard SiO$_2$ without requiring the use of toxic silane, and provides improved step coverage and film stress control. TEOS is the studied insulating material in the present work since the enveloping thermal insulating material is the determining factor in the thermal response of IC interconnects.

The final studied material, Tungsten (W), is used in the fabrication of vias that establish connections between interconnects at different MET levels. Vias are placed in the inter-MET TEOS layers and are generally made of barrier metals. These materials have a higher atomic bond energy and thus offer a high resistance to EM. Titanium (Ti), Tantalum (Ta), and Tungsten (W) are some barrier metals with high electrical conductivity that also can mitigate EM diffusion.

The above-mentioned materials (AlCu, PolySi, TEOS SiO2, and W) will be considered
for the TDTR conductivity measurements. Other novel materials or compounds may be used for BEoL fabrication, but this study focused on the above materials since they are the most widely implemented ones in the BEoL stage and to also provide thermal properties of the materials used in the numerical studies in this Chapter.

6.4.2. TDTR Results

The total thermal resistance \( R_{th,\text{total}} \) is extracted by fitting the results of the numerical simulation to the experimentally-obtained temperature decay profile for each sample structure included in this work. The extracted data are plotted against sample layer thickness and the corresponding curve fit equations are shown for each material (Figure 6.5).

![Graphs showing total thermal resistance vs. thickness for different materials](image)

Figure 6.5. Total thermal resistances obtained for the thin-film samples as function of film thickness. The slope and intercept of each data fit are used to extract the intrinsic thermal conductivity of the film and its interface thermal resistance.

With the linear trends in hand, it becomes possible to extract the intrinsic thermal con-
ductivity ($\kappa_i$) of each thin-film material as well as its associated interface thermal resistance ($R_{int}$).

But the question arises as to how to properly assess the uncertainty error associated with this data since not only are the number of samples and the trials for each of them relatively small, but the linear fitting equation, and therefore the intercept, are strongly dependent on even small vertical shifts in any of the few data points per curve (represented by the solid symbols in Figure 6.5). A worthwhile approach is to produce bounding linear fits, one upper and one lower, respectively based on considering the upper and lower bounds of the uncertainty for each of the data points. This way, one would respond to the question of what would the error be if all the uncertainties were either additive (upper band) or subtractive (lower band). The upper and lower linear fits would then bracket the “best-fit” plotted in Figure 6.5. The resulting uncertainty errors in the thermal conductivity, based on the maximum and minimum deviations in $R_{th,total}$, are less than 5% for Poly, TEOS, and W, and around 15% for AlCu. The latter is higher because its shallower slope magnifies its sensitivity to variations in $R_{th,total}$.

The properties summarized in Figure 6.6 indicate that the highest thermal conductivities are observed for AlCu (199 W/m.K) and W (115 W/m.K). Conductive metals are expected to have higher intrinsic conductivities since, in addition to the lattice contribution to heat conduction, the electronic contribution of loosely-held electrons in the lattice further increases the total thermal conductivity [110]. The obtained $\kappa$ values are close to those reported in the literature. For example, the conductivity of pure Al is reported to be around 230 W/m.K [1,110], and the effect of adding a certain weight percent (%wt) of copper to the aluminum tends to decrease the thermal conductivity of the resulting alloy. A study by Zhang et al.14 measured $\kappa_{AlCu}$ by a laser flash method and reported a value of around 220 W/m.K (for AlCu blends with 1%-wt Cu), 200 W/m.K (for 3%), and 190 W/m.K (for 5%) [274], which also correlate well with the results reported by Akerman and Havill14. The bulk thermal conductivity of Tungsten is reported in the literature to be between 159 W/m.K 15 and 165 W/m.K 16.
Figure 6.6. Summary of the measured intrinsic thermal conductivities of the BEoL materials and their interface thermal resistances.

The thin-film samples considered in this investigation exhibit a lower thermal conductivity, which could be the result of disordered crystal structures and thermal-mechanical stresses that are typically associated with the high temperatures and pressures used in the deposition process.

The thermal conductivity of Polycrystalline Si (Poly) is measured at 23.3 W/m.K, which is lower than that of single-crystalline Si (≈140 W/m.K) but higher than amorphous Si (≈5-6 W/m.K [254]) due to the polycrystalline nature of the material. The increased scattering present at the grain boundaries adds more resistance to phonon transport and thus reduces the lattice thermal conductivity. The literature reports similar values of $\kappa_{Poly}$ between 13.8 and 22 W/m.K as measured using joule heating and resistance thermometry [33,166,254].

The TEOS SiO$_2$ is a highly electrically insulating material and thus shows the lowest thermal conductivity measured at 1.12 W/m.K. The low thermal conductivity is expected for amorphous silicon dioxide and is measured in the literature to be in the range of 1 to 1.2 W/m.K [31,32,143,276] and verified by molecular dynamics simulations [113,167].

The effective conductivities $\kappa_{eff}$ given by (3.27) are calculated for the different samples.
and plotted against layer thickness (Figure 6.5). Since additional interface effects are accounted for, the value of $\kappa_{\text{eff}}$ is always less than $\kappa_i$, and by definition approaches $\kappa_i$ at the bulk limit.

![Graphs showing effective thermal conductivities for different materials](image)

Figure 6.7. The calculated effective thermal conductivities of the sample films as reconstructed from the extracted intrinsic thermal conductivities and interface resistances.

The difference between $\kappa_{\text{eff}}$ and $\kappa_i$ depends on the intrinsic conductivity of the material and the interface resistance. For low conductivity materials, the bulk thermal resistance is the major contributor to TTR, where films below a thickness of 1000 nm, the interface effects lead to a drastically reduced $\kappa_{\text{eff}}$. The bulk resistance $t/\kappa_i$ is larger for thicker samples and for that reason, $\kappa_{\text{eff}}$ approaches $\kappa_i$ as $h$ increases. For TEOS with a thickness of 100 nm, $\kappa_{\text{eff}}=1$ W/m.K is very close to the intrinsic value, and Polysilicon $\kappa_{\text{eff}}$ is seen to approach $\kappa_i$ for layers of thickness 1700 nm.

For high conductivity materials, the bulk resistance is low and thus the interface resis-
stances dominate the TTR, which leads to a low $\kappa_{eff}$. Even at a thickness of 1000 nm, the $\kappa_{eff}$ of AlCu is still at 40 W/m.K which is a small fraction of the bulk conductivity (200 W/m.K). A similar effect is seen with W, where $\kappa_{eff}$ for the thickest sample layer is around 55 W/m.K compared to a $\kappa_i$ of 115 W/m.K. Similar values are reported in the literature where the thin-film conductivity is reported for of a 273 nm thick W sample at 25 W/m.K and at 50 W/m.K after being heat treated [214]. In another study, $\kappa_W$ is reported at 50-60 W/m.K a 260 nm sample [107]. The reported bulk and effective conductivity is close to our measured results, however, higher interface resistance may have resulted in a lower value than what is reported in other studies.

6.4.3. TDTR Study Conclusions

The thermal conductivity and interface resistances of thin-film materials used in the fabrication of IC BEoL structures are measured using TDTR. The characterization of the thermal properties is essential to accurately model the thermal response and self-heating in IC structures for reliability analysis. The reported thermal conductivity measurements reveal a thermally resistive TEOS medium which might be the most restrictive to heat dissipation. Moreover, the interface effects observed for high-conductivity metals are seen to add to the total thermal resistance of thin-film metals such as AlCu and W and reduce their effective conductivity. The accurate measurements of material properties would provide more representative thermal models for reliability assessment of IC BEoL networks.

6.5. 2D Combined Model Results

The simulation models used for verifying the 2D reduced model is first validated with experimental thermal mapping of the test samples. Specially-fabricated devices of different widths and at two levels in the BEoL are thermally mapped with the T°Imager® while being electrically activated with a constant current source, resulting in Joule-heating induces temperature rise profiles. Once the simulation model is validated, it is used to verify the reduced 2D for the presented case of an isolated ICT.
6.5.1. Validation Simulations with Thermal Maps

The temperature rise of the test structures are extracted from the acquired thermal maps (Figure 6.8) at a region away from the edges is extracted from the both the measurements and simulation results and is used to calculate the total thermal resistance per unit length (2D) as as $TTR' = \Delta T/Q'_x$, where $Q'_x$ is the activation power level per unit ICT length. The allowable heat flux can then be calculated for any desired temperature limit $\Delta T_{allow}$ as $Q''_{allow} = \Delta T_{allow}/(TTR' \cdot w)$ (Figure 6.9). The results show close agreement between the simulation models and the observed temperature rise for the tested devices. Two types of Poly resistors of different electrical resistivities are tested and hence the two data sets present for Poly.

A thermal model of the tested devices is built in T°Solver® with the corresponding dimensions, thermal properties and heat generation rates. The effect of the chemical mechanical planarization (CMP) dummy fill pattern is shown in a previous study to affect the heat conductance within the BEoL stack and result in a passive cooling effect [77]. The CMP fill is therefore modeled with the same effective thermal properties as will be presented in the next Chapter [77]. The validation results (Figure 6.9) show that the simulation models results closely approximate the observed temperature rise in the six test devices at the two BEoL levels. For a more general comparison to the previously developed 2D model, the effective conductance in the stack is presented in Figure 6.9(b) in terms of $\kappa_{eff}/\kappa_d = Q'_x \cdot t_i/(w \cdot \kappa_d \cdot \Delta T)$. Close result agreement is observed, and the two curves of level 1 (Poly) and level 2 (MET) are almost confounded since a relatively thick passivation ($t_p/t_i > 5$) is present. Minor discrepancies in the experimental validation could be the attributed to uncertainties in the defined thermal parameters for the simulation model, such as the effective thermal parameters of the present materials. Moreover, the obtained $\kappa_{d,eff}/\kappa_d$ results in Figure 6.9 are higher than those presented in Figure 6.11 due to the increased conductance from the CMP dummy fill.

From the experimental thermal maps (Figure 6.8), it is also observed that heat dissipation in the axial direction can largely affect the temperature distribution especially at the ends of an active ICT. This is mainly due to the factors relating to thermal entry, heat conduction
Figure 6.8. Thermal maps of the sample BEOl devices micro-resistors of widths 1, 3 and 10 $\mu$m at the different locations: Level 1 of Polysilicon (Poly) shown in the left column and level 2 of AlCu metalization (MET1) shown in the right column.
Figure 6.9. Comparison of the heating rate for measured ICT samples as compared with their estimates from the reduced model, with $\kappa_d = 1.1W/m.K$. 
and heat generation at the vias. The combined model given in (6.16) and depicted in Figure 6.11 can accurately estimate the 2D effective heat dissipation from a thermally-long active ICT, and is used to develop a more ambitious 3D reduced model that would be capable of considering the above-mentioned effects.

6.5.2. Verifying 2D Reduced model with Simulations

The 2D reduced models for $\kappa_{\text{eff}}/\kappa_d$ are now compared to full 3D simulations for a set of device structures and dimensions. The mathematical models of Bilotti and Andrews are seen to inaccurately represent the effective conductivity for the case of an ICT embedded in finite passivation (Figure 6.10). $\kappa_{\text{eff}}$ is underestimated in Bilotti’s model and overestimated in Andrews’ model, which is expected since the models represent two extreme cases of no passivation (Bilotti) and infinite passivation (Andrews), with the former removing heat conduction paths (reducing thermal conductance), and the latter doing the exact opposite.

For the purpose of obtaining an improved conductance model, this study derives a semi-empirical model that combines the two aforementioned mathematical models and develops a combined model that accurately represents conductance from ICTs with finite passivation thickness, $t_p$). The combined model bridges the range of $t_p$ between the Andrews and Bilotti models and presents a more accurate effective conductivity for a real ICT that has a finite passivation thickness. Given that the model should converge to Bilotti at $t_p = 0$ and to Andrews’ at $t_p = \infty$, and as observed from an initial set of numerical simulations, an exponential variation has been used to model the variation of $\kappa_{\text{eff}}$ with $t_p$. Presented in (6.16) is an exponential relation of $\kappa_{\text{eff}}/\kappa_d$, referred to in this text as the combined 2D model. The exponential variation is seen to depend on the shape factors $t'_p$, $w'$ and the fitting exponential coefficient $\zeta$.

$$\frac{\kappa_{d,\text{eff}}}{\kappa_d} \bigg|_C = \frac{\kappa_{d,\text{eff}}}{\kappa_d} \bigg|_A \cdot \left(1 - e^{-\zeta \frac{t'_p}{w'}}\right) + \frac{\kappa_{d,\text{eff}}}{\kappa_d} \bigg|_B \cdot e^{-\zeta \frac{t'_p}{w'}} \quad (6.16)$$

Adding the passivation thickness dependency improves the accuracy of the 2D model.
appreciably (Figure 6.10). A set of thermal simulations is conducted for a thermally long interconnect with dimensions $w$, $t$, $t_p$ and compared with the combined reduced model. The results of the reduced model are seen to closely estimate the simulation results, and a final tuning is performed for the fitting parameters from Bilotti, Andrews, and the exponential $\zeta$ (Figure 6.10). The optimal values are obtained as follows and are close to literature values (reported for the Andrews model) [12, 40]. The results of the reduced model (red symbols in Figure 6.10) provide an estimate of the simulation results with an accuracy of within 2%, compared to as much as 20-40% error for the Bilotti and Andrews reduced models.

$$B^* = 0.944$$

$$\alpha = 1.553$$

$$\beta_1 = -0.726$$

$$\beta_2 = -0.067$$

$$\zeta = -1.43$$

Using the optimized fitting parameters the finalized model (6.16) can be presented graphically (Figure 6.11) for different widths and passivation thickness shape factors ($w' = w/t_i$ and $t'_p = t_p/t_i$).

As observed in Figure 6.11, $\kappa_{eff}/\kappa_d$ is always higher than 1, since 1D heat conduction is the most conservative conductance model. The heat spreading effect increases conductance by at least 20% within ($w' < 5$), and could exceed 100% for narrower ICTs ($w/t_i < 1.5$). For wider ICTs with $w/t_i < 1.5$, the effect of passivation is no longer significant and $\kappa_{eff}/\kappa_d$ gradually approaches 1 for thermally wide resistors. Moreover, the Andrews model is seen to accurately model the cases for which $t_p/t_i > 8$, whereas the Bilotti model does better for cases with $t_p/t_i < 0.25$.

The implications of this result for BEoL designers is significant in that it points to the facts (a) that there is now a fast estimate for the effective thermal conductance from an active ICT which is more closely related to the physics and hence much less conservative (b) that
Figure 6.10. Comparison between the newly developed combined model and the simulation results, the models of Bilotti and Andrews are also shown for each case to highlight the improvement of using the combined model.

narrower ICTs’ temperature rise is not as significant as would have been thought previously, and (c) accounting for the presence of additional passivation introduces an improvement in the total thermal conductance. Understanding how the thermal conductance can be improved in relation to other physical parameters the allowable electrical activation levels can be safely—from a thermal perspective—extended without requiring additional costs or bulkier ICTs, and without the need for extensive computations or conservative models.

6.6. 3D Compact Model Results

A similar approach to the previous section is presented to verify the developed 3D compact model (Section 6.3). The reduced modeling results will be verified against numerical simulations to ensure that results agree given the same defined input parameters for the simulations and the reduced model. Experimental thermal maps of activated sample devices of similar test structures were initially used to validate the numerical simulation model that will in turn be used to verify the reduced 3D model.
Figure 6.11. Combined model chart showing effective conductance $\kappa_{d,\text{eff}}/\kappa_d$ from an ICT of width $w$, insulation thickness $t_i$ and passivation thickness $t_p$, all of which are denoted in the added schematic. The Bilotti and Andrews cases show the extreme effects of, respectively, ignoring passivation or making it infinite.

6.6.1. Validating Simulations with Thermal Maps

To initially validate the simulations, test device structures are thermally mapped using the T°Imager thermal imaging system described in Section 6.5.1. Simulation models that are representative of the test structures are validated for three different device configurations, consisting of a suspended ICT at level 1 (Poly) with via connections to two contact leads at level 2 (MET Level) (Figure 6.12 a). A transversely-running "sense" lead when present is located at level 2. The "sense" is inactive in one case, and active in the other (Figure 6.12 b). The devices are activated at a certain current density $J$ and thus are Joule heated, resulting in an axial temperature rise profile.

From each of the performed thermal maps, the temperature rise is obtained by averaging over a central region in the resistor away from the edges to avoid the via cooling/heating effects. The temperature variation across the width of the micro-resistor is neglected and
the lumped approach is assumed because the thermal conductivity of the resistor material is much larger than the neighboring oxide medium.

The heating or cooling of the resistor’s entrance region around the via is dependent on their materials’ relative electrical resistivity $\rho_0$. For Poly resistors, most of the heat is generated in the higher resistivity Poly. The Tungsten vias conduct some of that heat to the contact leads, cooling the edge regions of the resistor. On the other hand, Tungsten is more resistive than AlCu and thus volumetrically generates more heat for the same current density and thus heats up the entrance regions around the vias. As a last observation, the contact leads are also of AlCu composition and thus experience similar heating density as the AlCu resistor, showing a temperature rise comparable to that of the micro-resistor. These cooling/heating effects can bias our measurements, therefore the edge regions are avoided when measuring the Joule heating temperature rise of the resistors.

It is important to point out that simulation accuracy is always as good as that of the input
parameters for the simulation model, such as the effective thermal conductivity values, which consider the interface resistance effects. For thin films, these parameters often differ from the bulk values and require in-situ measurements with special sample fabrication. A small variation from the actual values could lead to error in experimental validation. Given that ballpark values reported in the literature are used, the experimental validation discrepancies are satisfactory and show good agreement. In the next section, the reduced model will be compared to the simulation models with identical defined input thermal parameters. In this manner, any discrepancies that would arise in comparison would be the result of using reduced models, and not due to mismatch in the input parameters.

The test structures are thermally mapped to validate their representative simulation models. The resistor and sense lead are activated with a current of 30 mA and 250 mA, resulting in a Joule heating rate $Q'_x$ of 0.945 and 0.625 mW/µm, respectively. The actual device dimensions and experimental power levels are faithfully represented in the simulations.

The thermal simulations are conducted for the three cases of $a)$ an active resistor with no sense lead, $b)$ an active resistor with an inactive sense lead, $c)$ and an active sense lead with an inactive resistor. The temperature is extracted from a central axial trace along the resistor-contact leads averaged across a 5 µm centered width. The results from both experimental and simulation temperature profiles are extracted and compared in Table 6.6.1. Row (1) shows the thermal maps obtained using thermoreflectance imaging, row (2) shows the thermal profiles obtained from the simulation models, and row (3) presents a comparison between the temperature traces of the experimental measurements and of the simulation model. When the sense lead is present (case (B) and (C)) a 10 µm portion of the resistor lead lies beneath the sense lead and is inaccessible to the optical measurement method. For that reason, the central 10 µm portion of the trace in subfigure (3.B) and (3.C) corresponds to the sense lead. Moreover, the chemical-mechanical planarization (CMP) process utilized a dummy fill pattern, clearly seen in (1.C) which covers most of the wafer surface, but is intentionally designed to keep a small region around the device exposed for optical access.

The simulation results for the three cases are compared to the experimental measure-
Temperature axial distribution along the resistor lead showing heating in the activated region and the cross coupling between the resistor and the sense lead. The three graphs show the single resistor configuration (Left), the active resistor with a cross coupled sense lead (Middle), and the active sense lead atop an inactive resistor (Right). Row (1) showing experimental thermal map, row (2) showing simulation thermal map, and row (3) showing a comparison between temperature distributions along the axial direction from the experimental data (Symbols) and the numerical data (Solid lines).

<table>
<thead>
<tr>
<th>(A) Active Resistor</th>
<th>(B) Active Resistor</th>
<th>(C) Active Sense</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inactive Sense</td>
<td>Inactive Sense</td>
<td>Inactive Resistor</td>
</tr>
</tbody>
</table>

(A) Active Resistor

(B) Active Resistor

(C) Active Sense

Inactive Sense

Inactive Sense

Inactive Resistor

(1) Experimental Maps

(2) Simulation Results

(3) Temperature Traces
ments in row (3). The simulation results are in close agreement with the observed thermal maps. The oxide effective conductivity that matches the observed temperature distributions is obtained at $\kappa_{d,\text{eff}} = 1.07 \, \text{W/m.K}$, which is within an expected range for thermally grown oxide (TEOS). In (3.A), the central region shows some distortion due to the presence of a very thin sense lead that is neglected, but however, results to edge distortion effects in the experimental trace. Regarding the contact lead, only a small section is exposed from underneath the CMP fill pattern. The measurement setup is optimized to detect the temperature rise from the resistor material and thus has a low sensitivity to the contact leads heating, which, along with a small $\Delta T$ signal, lead to high uncertainty in the contact lead measurements. In (3.B) the temperature is seen to closely estimate the heating in the resistor as well as the mutual heating to the sense lead, which is extracted from the exposed sense portion that is not covered by the CMP fill pattern.

6.6.2. Verifying 3D Reduced Model with Simulations

The reduced 3D model is compared to the experimentally-validated simulation model. The tested structures consist of sample ICT components in the configurations of Figure 6.12 (b) under three different activation setups. The iterative error converges within a few iterations, where the residual error drops below 1.5% after 5 iterations, below 0.01% after 10 iterations, and below 0.002% after 20 iterations. The results presented in Table 6.6.2 are obtained with high computational efficiency where the model is solved and plotted in less than one second.

The results of row (1) initially show that the reduced model closely estimates $\Delta T_{\text{max}}$ in the thermally developed regions, which is expected after validating the 2D model in Section 6.2. Moreover, the model succeeds in representing entrance effects, as well as the conduction through the vias and the cross configuration with the respective $\lambda_{H,i}$, as well as succeeds in calculating the node temperatures $\Delta T_n$ to build the temperature distributions in each of the leads. However, some discrepancies are observed, where secondary effects contribute to deviations from the simulation results, seen clearly in regions denoted by (M), (N), and
Table 6.2. Comparison between temperature trace of simulation results (Dashed) and approximate model (Solid) for Resistor lead (Blue), Sense Lead (Green), and Contact Leads (Red). (1) shows the results for the initial model without global heating effects. (2) results with global heating (3) results with global heating and corrected cross model The three rows present the original model and two later modifications that are implemented based on presented observations and discussions. The first column (A) shows the case of an active resistor, the second column (B) shows the case of an active sense lead with an inactive resistor, while the third shows the case for active sense and resistor leads.

<table>
<thead>
<tr>
<th>Case</th>
<th>(1) Original Model</th>
<th>(2) with Global Heating</th>
<th>(3) with Corrected Cross Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A)</td>
<td>Original Model</td>
<td>With Global Heating</td>
<td>With Corrected Cross Model</td>
</tr>
<tr>
<td>(B)</td>
<td>Active Sense</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C)</td>
<td>Both Active</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
These discrepancies are observed in two main regions and are thus associated with two aspects of the model.

The first observation can be clearly inferred from the contact leads in (row 1) where there appears to be a temperature bias due to an additional uniform/global heating in the simulations. This could be due to the thermal resistance of the substrate which is not modeled by the 3D reduced model, and thus the simulation temperature does not decay to zero. Note that the contact leads are active whenever the resistor leads are active, but due to a much lower resistivity, exhibit negligible heating in comparison. By adding the global heating effect, the mutual heating between the different leads can be modeled, resulting in closer temperature estimates as shown in Table 6.6.2 row (2). A quick simulation investigation shows that the temperature decays to a global temperature between 3 to 7% of the $\Delta T_{\text{max}}$ depending on the lead width. More complicated models can be used to estimate mutual heating between neighboring active ICTs, however, the temperature is seen to sharply decay a few microns away from an active ICT and only the global heating effect persists.

The second discrepancy in the reduced model is observed in the cross region where the model used from [40] is seen to inaccurately represent the heating in our studied configuration. The temperature is seen to be underestimated in the model, which could be due to inaccuracies in the cross resistance ($R_{\text{th,cross}}$) or the ground resistance ($R_{\text{Bil.}}$) of the lower ICT. To improve the result agreement, the cross model is corrected where the coefficient of $R_{\text{th,cross}}$ in 6.15 is changed from 3/2 to 15/13, and $R_{\text{Bil.}}$ is corrected by a factor of 2/3. The improved model results are presented in row (3) of Table 6.6.2, with more accurate approximations of the cross temperature distribution for the three instances of activations.

By implementing the above improvements, the model is seen to give closer estimates to the temperature profiles of the studied configuration while accounting for edge conduction and cross coupling effects. The model could be improved further by considering more of the secondary heat conduction paths in the model to get more accurate representations of $\lambda_h$ and $\Delta T_{\text{max}}$, which would increase the model’s complexity.

It is crucial to note that the reduced model is able to reach a solution in less than one
second for our test structure. This is at least two orders of magnitude faster than the ultra-fast computations conducted by T°Solver, which in turn is another two orders faster than traditional/commercial thermal software engines. To model complex 3D microelectronic devices, the T°Solver multi-grid self-adaptive engine is highly effective and practical, and would give highly accurate representation of the thermal response. However, the engine may not be as convenient for a BEoL stage with large scale integration (VLSI and ULSI) consisting of millions of individual ICTs. For that reason, the reduced model would be the only feasible solution for a fast thermal tool for reliability assessment of IC BEoL grids. The 2D and 3D models are implemented in an easy to use GUI for initial assessments (Figure ) and can be integrated in future work to other simulation softwares

6.7. Conclusions

This study presented a reduced thermal model based on mathematical derivation for the 2D heat conduction for ICTs in BEoL stack. The 2D is validated against thermal measurements and numerical simulations, and is later used with the lossy wire model to develop a reduced 3D (2D+1D) model to account for edge conduction effects to vias, and mutual heating effects in cross configurations. The numerical model is based on resistive circuits that are solved with fast analytical computations and few iterations to reach a steady state thermal response of an ICT network. The presented method is more accurate than the currently applied reduced models based on Andrews or Bilotti, and can provide a full 3D thermal characterization with a fraction of the time and computational cost of full thermal models. The 3D reduced model is verified and compared to simulation results for a simple geometry of BEoL device samples, which is in turn validated with experimental thermal maps of specially fabricated BEoL device samples. The reduced thermal model finally can be expanded and integrated into already existing BEoL layout softwares to serves as a fast thermal check for grid networks to ensure reliable operation within thermally safe limits.
Figure 6.13. Graphical user interface for 2D reduced model built in MATLAB GUIDE. The interface shows input electrical and thermal parameters for an isolated interconnect, and provides the allowable current density levels. The model also reports on the effective thermal conductivity of the surrounding dielectric.
Figure 6.14. Graphical user interface for the 3D reduced model built in MATLAB GUIDE. The interface shows input electrical and thermal parameters for an isolated interconnect as well as the dimensions of the interconnect and its location within the BEOI stack. The model provides the allowable current density levels as well a temperature distribution along the length of each interconnect segment. The insert graph shows the overlaid temperature of the test samples used in this study.
Chapter 7

Standardized Heat Spreader Design for Passive Cooling of Interconnects in the BEOL of ICs

The ability of accessible thermal investigation at the sub-micron scale not only allows to improve on the available models and technologies, but also opens the possibility to develop advance thermal management strategies that were infeasible to validate and at scales that were previously inaccessible. With advanced capabilities in experimental and numerical methodologies, thermal solutions should no longer be limited to the system level and can be designed and integrated at the device (front-end) and integration level (back-end) to maximize their thermal cooling benefit.

The Back-End components of integrated circuits (ICs), being susceptible to self-heating deteriorating effects due to the microscale interconnects carrying large current densities, require innovative passive cooling strategies. This Chapter investigates the cooling effect of the high conductivity metalization fill that is used in conjunction with the Chemical Mechanical Planarization (CMP) process. Thermal simulation models are developed for a three-level BEOL stack and then validated experimentally using high resolution thermal imaging. A coupled approach is used to optimize the numerical model and fully characterize the BEOL embedded passive cooling solution. Results indicate that stricter thermal constraints are present for devices higher in the BEOL stack, and exhibit the highest cooling potential from the use of CMP fill in the underlying metal (MET) layers. As a thermal improvement strategy for such thermally-critical interconnects, a CMP fill based cooling solution is presented by modifying the fill pattern in order to maximize heat dissipation within the BEOL. The resulting heat spreader design achieved up to 30% reduction in temperature for 10 \( \mu \text{m} \)-wide interconnects, and can reach 43% for narrower interconnects. The gained cooling makes it possible to extend the activation limits of interconnects by 15% and 25%, respectively. The
experimentally validated HS simulation model is used to conduct parametric analysis for a range of interconnect dimensions with an eye on standardizing the HS design within the BEOL stage of ICs.

7.1. Introduction

As presented in Chapter 6, the BEOL stage of an Integrated Circuit (IC) microchip consists of interconnects that relay signals between, and distribute power among the front-end devices. These interconnects are activated at high current densities and surrounded by a highly thermally resistive dielectric medium (Inter-layer ILD, or Pre-metal PMD dielectric), which causes critical self-heating [16]. This makes BEoL interconnects more prone to temperature-induced exponential reduction in lifetime as a result of Electro-Migration (EM) failures.

Fortuitously, a procedure used during the BEOL fabrication stages replaces a portion of the mentioned insulating medium with high thermal conductivity metalization (MET) fill, and can hold vast cooling potential in reducing operating temperatures in the BEoL. When fabricating the multiple MET levels of the BEOL, a Chemical-Mechanical Planarization (CMP) process is applied to ensure that a level surface is available for the deposition of the next higher dielectric and MET layers [120,135,273] (Figure 7.1). A generic CMP process introduces two major challenges in the form of surface defects associated with thickness loss in the metal and in the dielectric regions (referred to as "dishing" and "erosion", respectively) [186,275].

As a proposed solution to reduce the above surface defects, a dummy MET fill pattern incorporated within the dielectric region of the MET layer helps improves surface homogeneity and increases fabrication yield [36]. This pattern of metallic fill has been shown to add some parasitic electrical effects such as capacitance coupling and signal delay [94,124] as well as a visible increase in thermally-induced mechanical stresses [123]. However, from a heat transfer perspective, the fill material is of high thermal conductivity, which compensates for the
poorly conductive dielectric and in turn enhances the effective thermal conductance within the BEOL stack. As a result, the CMP fill could serve as a local heat-spreading medium within the BEOL which reduces operating temperatures of BEOl active components and thus improves their reliability.

Thermal studies on the cooling effect of CMP fill are lacking by comparison to those on the thermal characterization of BEOL interconnect [246,251]. Chiang et al. [252] and Nguyen et al. [183] studied the cooling effects of vias in enhancing heat transport from interconnects, and using a similar approach, they studied the impact of dummy thermal vias. The effect of dummy vias is similar to that of a CMP dummy fill, except that the fill density, spacing, and locations of the vias are different from those of the CMP pattern. Datta et al. [68] performed simulations to assess what thermal improvements the CMP fill pattern might have on the heat conduction of the metal layers. The authors reported an improvement in the in-plane conductivity and in temperature planarity. Their study considered different fill pattern densities and different inter-layer dielectric (ILD) materials (air, polymer, SiO₂). However, only a single level of interconnects was considered and no validations with actual
device measurements were presented.

In the first part of this study, we present the results of a numerical investigation of the thermal effects of incorporating CMP dummy fill into the BEOL stack for different interconnect dimensions and locations in the BEOL (Section 7.2). The model is an extension of the one presented in Chapter 6 and thus the experimental validation holds for the samples with the CMP fill. After this first numerical assessment of the cooling potential of the high-conductivity MET fill, we present in the second part an investigation of an improved fill pattern designed to serve as a heat spreader (HS) solution in the vicinity of thermally-critical interconnects. The HS design aims at effectively dissipating the generated heat locally from the active interconnects, and thus relieving some of the deteriorating self-heating effects (Section 7.3). Afterwards, we present a coupled approach that uses experimentally observed thermal maps of specially-fabricated devices featuring the proposed HS in order to drive an optimization method that validates the the thermal simulation model we sought to develop. This coupled approach makes it possible to overcome uncertainties in the standard thermal properties used as numerical input parameters and thus yields an experimentally-validated thermal model that can be finally used with confidence for conducting wider-ranging parametric studies. In Section 7.3.4, we use the validated and experimentally-calibrated thermal model to design a candidate HS for optimal cooling solution for interconnects of different dimensions in thermally critical locations in the BEOL. The final outcome is a general BEOL cooling solution that could be standardized and integrated in the initial design process of the BEOL of ICs.

7.2. Thermal Simulation Model

To first assess the added cooling effect of the metallic CMP fill in the BEOL stack, a numerical model representative of a three MET layered BEOL stack (Figure 7.2) is developed. The simulation model is built in T-Solver®, which is a thermal modeling software specifically designed for IC structures with advanced capabilities that provide ultra-fast simulations with a self-adaptive, multi-grid meshing approach essential in handling large variations in spatial
The modeled device structure consists of three microresistor devices embedded at three different levels in the BEOL. The interconnects are 1, 3, and 10 µm wide and located at each of the three metalization levels in the stack. Level 1 is made of Polysilicon (Poly) material and is at lowest level, closest to the substrate surface. Levels 2 and 3 are the two higher metalization layers and are both made of an Aluminum Copper (AlCu) alloy.

The CMP dummy fill patterns are fabricated through the same metalization process as the remaining features at each MET layer (Figure 7.3). Consequently, the fill is located at the same levels as the interconnects, are of the same thickness as the MET features, and are composed of the same material as the layer that they are located within (i.e., Poly in Layer 1, and AlCu in Layers 2 and 3). The patterns are blocked within a “keep-off” distance around device components such as interconnects, micro-resistors, contact leads, and vias to provide optical access for thermal imaging.

Since modeling the complex patterned fill directly would require a large computational effort, and since the multiple interface resistances between the fill material and the oxide
Figure 7.3.  

a) Top view intensity map of a sample device under test (DUT). The center resistor is at level 1 and surrounded by the CMP fill square pattern. Two contact leads are seen at both resistor ends a central sensing lead is running transversely above the resistor. More CMP fill is also seen at the higher levels of the BEOL stack;  
b) Top view showing fill pattern unit cell dimensions and in-plane heat transfer path.  
c) Equivalent thermal resistance circuit model for calculating in-plane effective thermal conductivity.  
d) Side view of CMP fill pattern unit cell and heat conduction path  
e) Equivalent thermal resistance circuit model for through-plane thermal conduction. density $w_f^2/w_d^2$.  

126
medium are unknown, an initial thermal study is conducted assuming a homogeneous fill medium with equivalent effective thermal properties. A representative thermal resistance circuit analysis (Figure 7.3) is used to obtain the effective in-plane thermal conductivity ($\kappa_{ip,eff}$) and through-plane conductivity ($\kappa_{tp,eff}$) of the fill lattice structure. The effective conductivities are calculated using (7.1) and (7.2) for a square pattern fill of side width $w_f$ and conductivity $\kappa_f$ in a dielectric unit cell of width $w_d$ and conductivity $\kappa_d$. A detailed derivation is given in the Appendix.

$$\frac{\kappa_{ip,eff}}{\kappa_d} = 1 - \frac{w_f}{w_d} + \left( \frac{w_d}{w_f} - 1 \right) \left( \frac{\kappa_d}{\kappa_f} \right)^{-1} \quad (7.1)$$

$$\frac{\kappa_{tp,eff}}{\kappa_d} = 1 - \left( 1 + \frac{\kappa_f}{\kappa_d} \right) \left( \frac{w_f}{w_d} \right)^2 \quad (7.2)$$

The effective material density and specific heat are obtained from a volume average for the square pattern fill and defined in (7.3) and (7.4). As shown in Figure 7.3, the CMP fill occupies a square region of width $w_f = 2\mu m$ in the square dielectric lattice of width $w_d = 3\mu m$, resulting in a fill volumetric ratio of $w_f^2/w_d^2 = 44\%$.

$$\frac{c_{p,eff}}{c_{p,d}} = 1 - \frac{w_f^2}{w_d^2} \left( 1 - \frac{c_{p,f}}{c_{p,d}} \right) \quad (7.3)$$

$$\frac{\rho_{eff}}{\rho_d} = 1 - \frac{w_f^2}{w_d^2} \left( 1 - \frac{\rho_f}{\rho_d} \right) \quad (7.4)$$

Three device models are built to represent the active interconnects at the three BEOL levels. The temperature rise is extracted from the simulation results and the allowable activation rate is calculated for each of the devices of the varying widths and at the different levels in the BEOL. The simulation results reported here for comparison are in the form of the allowable heat generation flux $Q'(mW/\mu m)$ instead of the allowable electrical activation since the different materials have distinct electrical conductivities and the heat generated rate is not equal for a given current density $J_{RMS}$. 

127
Figure 7.4. Simulation results of heat generation limit for an allowable temperature rise ($\Delta T$) of 15 °C for interconnects at the three MET levels as function of device (resistor) width.

The simulation results (Figure 7.4) are presented for the allowable heat generation flux for the different device widths, and at the three BEOL levels. A first observation reveals a reduction in the allowable activation level for wider interconnects. As the width increases, the generation region widens and heat initiating in regions farther away from the interconnect edges become more challenging to dissipate laterally. Lateral heat dissipation adds more cooling to narrower interconnects and thus increases their allowable activation level [216].

Moreover, the location in the BEOL also affects self-heating of interconnects. At higher locations (higher in the stack) in the BEOL, activated interconnects are separated from the substrate by a thicker layer of oxide, increasing the thermal resistance and thus elevating the operating temperatures for a given electrical activation. This further limits the allowable activation levels for devices higher in the stack if a maximum temperature is not to be exceeded. This is evident from the lower allowable heat flux for Level 3 as compared to Level 2 and Level 1. This hierarchy is an unfortunate direct opposition to the actual BEOL electrical design requirements, where more power is delivered at the higher levels. The limitation in
the allowable activation levels would thus require more critical thermal management or bulkier conductive materials at the higher BEOL levels to reduce electrical resistance and self-heating, which then adds to the fabrication costs.

Table 7.1. Percent reduction in self-heating temperature rise within the BEoL interconnects, and the percent increase in allowable current density \((J)\) and heat flux \((Q')\) showing the cooling effect of the CMP Dummy Fill (*) at different levels of the BEoL.

<table>
<thead>
<tr>
<th>Level</th>
<th>(\frac{\Delta T - \Delta T^<em>}{\Delta T} \times \frac{Q^</em> - Q'}{Q'}) (%)</th>
<th>(\frac{J_{RMS}^* - J_{RMS}}{J_{RMS}}) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 1</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>Level 2</td>
<td>16.5</td>
<td>9.5</td>
</tr>
<tr>
<td>Level 3</td>
<td>20</td>
<td>13</td>
</tr>
</tbody>
</table>

The CMP fill pattern has a visible cooling effect as seen from the dashed lines in Figure 7.4. The generated Joule heat in the interconnects has to dissipate from the generation region in the microresistor to the heat sink in the substrate through the highly resistive dielectric material. The fill pattern in the intermediate MET levels improves the thermal conductivity of the dielectric medium and results in a higher cooling effectiveness for Levels 2 and 3, reducing their temperature rise during operation by around 16.5% and 20% respectively. For the case of Level 1, no CMP fill is present in the dielectric medium separating the micro-resistor from the substrate, thus resulting in only 3% cooling and is attributable to the surrounding CMP fill. The cooling rates and the consequent increase in allowable activation levels are summarized in Table 7.1. The reported values from the simulations assume ideal thermal contact between the fill pattern and the dielectric medium and thus may be overestimated. In the actual case, interface effects would add more thermal resistances thus reducing the cooling effectiveness of the fill pattern. This first study is to assess possible cooling effects of the CMP fill in ideal conditions. A more accurate approach is presented in Section 7.3 where in-situ testing results are matched with the simulations to provide an experimentally-validated numerical model.
7.3. Heat Spreader Design

The CMP fill pattern is seen to improve the heat conductance, and shows promising cooling potential for thermally critical interconnects at the highest location in the BEOL stack. In the following part of the study, this observed thermal cooling effect is utilized for assessing the design of a heat spreader to cool devices at MET Level 3. The choice of Level 3 is justified since the highest level is most susceptible the self-heating, and since the highest cooling efficiency of the CMP pattern is observed at that level. The CMP fill pattern is thus modified and a heat spreader (HS) design is proposed for the effective passive cooling of BEOL interconnects.

In Section 7.3.1, sample devices of the proposed HS design are specially fabricated and their temperature rise is measured through thermoreflectance microscopy. The cooling effect is observed experimentally and used for the validation of a simulation model in Section 7.3.2. An iterative approach is used to iterate for uncertain model input parameters and to match the observed experimental and numerical results, which are presented and discussed in Section 7.3.3. As a result, the model obtained faithfully represents the actual device’s thermal behavior and can then be used for additional parametric design studies. As an example, a parametric study for determining the optimal heat spreader dimensions is presented in Section 7.3.4.

7.3.1. Temperature Measurements of Sample Devices

The proposed pattern design for the heat spreader shown in Figure 7.5 consists of an array of thin, closely-packed, and transversely-running metal lines that would improve the through-plane and lateral heat dissipation. Referred to as heat spreading, this secondary heat flow distributes the generated heat in a direction perpendicular to that of the primary heat flow running from source (resistor) to sink (substrate). This allows the heat to be spread out to a wider region and thus reduces the operation temperature of the device.

The width of the HS metal lines are 0.5 \( \mu m \) and spaced apart at 0.5 \( \mu m \). The array patterns are located at the first (Poly) and second (AlCu) MET levels and will serve as
Figure 7.5. Schematics showing the design configurations of the microresistor BEOL devices: 
a) without the heat spreader and b) with the heat spreader. c) The actual device under test 
intensity image as obtained by the camera. The top row shows the narrow HS configuration 
while lower part shows the wide HS configuration. d) Side view showing the schematic of 
the HS features at the different levels of the BEOL. e) Top view showing the staggered array 
structure of the Poly and AlCu lines beneath the resistor.
a heat spreader for cooling the devices in the highest and most thermally critical region of Level 3. The interconnect representative device is a 10 \( \mu m \) wide micro-resistor, chosen since the widest interconnects are the most thermally constrained. Two variations of the HS design are tested and shown in Figure 7.5. The first is a 10 \( \mu m \) wide HS configuration (narrow) same as that of the activated microresistor, while the second is a \( w = 52.6 \mu m \) wide HS configuration extending almost twice the width beyond the microresistor in each direction. The two designs will assess the heat spreading effectiveness of the HS in reducing the operating temperature rise, both through enhancing the transverse and the through-plane heat conduction.

7.3.2. Experimentally Validated Thermal Model

A thermal model is built to extract important device thermal characteristics of the measured devices, such as the effective thermal conductivity of the HS array and the separating dielectric layer. The HS arrays are added within the underlying oxide layer and modeled as a uniform block with an effective thermal conductivity, density and specific heat as obtained from a representative thermal resistive network analysis. An intermediate oxide layer is defined between the two HS MET layers with an effective conductivity to also consider the added interface resistances due to stacking. The CMP pattern surrounding the device and the HS is modeled as a uniform fill region with the effective properties as presented in Section 7.2.

In an iterative approach, the input parameters for the simulation model are varied and tuned until the numerical results match the observed experimental thermal response. The simulation model will iteratively run until the RMS error between the observed and computed temperature maps is minimized [208, 211]. In this manner, the model is validated with the experimental temperature maps of the tested devices, and the model parameters are accurate and faithful to the actual device thermal characteristics.

To drive the iterative method, the thermal maps for the activated devices are obtained using the same thermoreflectance methodology detailed in Chapter 2. The temperature rise
of the resistor is then compared to the simulation results for the two configurations with and without the embedded heat spreader. The heating rates used for the experimental activation are matched with the heat source definitions in the simulation model.

7.3.3. Simulation and Experimental Results

Figure 7.6. Thermal simulations of activated devices (Right) compared to experimental temperature maps (Left) of device with narrow heat sink (Top) and wide heat sink (Bottom) for a heating rate of 1.54 mW/µm.

The measured temperature maps for the devices with narrow and wide HS configurations are shown in Figure 7.6 and compared to the corresponding temperature distributions obtained from the numerical model. The temperature rise $\Delta T$ averaged in a central region
of the resistor is extracted from the thermal maps and the simulation results and plotted against the heat generation rate $Q'$ in Figure 7.7.

![Thermal Maps](image)

**Figure 7.7.** Temperature rise profile showing cooling effect and the reduction in overall thermal resistance experienced by the 10 $\mu$m interconnect in Level 3 for device with a) narrow HS and b) with wide HS.

The plot shows a close match between the experimental and simulation results which validates the developed numerical model. By comparing the heating rates $\Delta T/Q'$, the cooling effect can be described for the different configurations. Initial iterative matching is performed for the oxide conductivity for the bare model with no heat sink. The matching results in a conductivity of $\kappa_{d,eff} = 1.4$ W/m.K, which is close to reported values for silicon dioxide ($SiO_2$). Next, an iterative approach on the narrow HS determines the effective thermal conductivity for the inter-layer dielectric between the two HS MET layers. The matching for the narrow heat sink shows an effective through-plane oxide thermal conductivity drop to about $\kappa_{d,eff} = 0.7$ W/m.K. This effective conductivity considers the contact thermal resistances that are added by the multiple interfaces of the HS structure. The matching results are then verified for the wide HS, where the same thermal properties provides matching results with the experimentally observed thermal response of the wide HS device. This validates the numerical model and verifies that the converged input thermal properties are accurate.

The heating rate in Figure 7.7 of the narrow HS design shows an 8.5% reduction in the overall thermal resistance of the stack from 126.8 °C.$\mu$m/mW to 115.3 °C.$\mu$m/mW,
while that of the wide HS design results in a 30.2% reduction from 117 °C.µm/mW to 87.9 °C.µm/mW. The narrow HS case mostly improves the through-plane conductance since the HS is located just beneath the resistor, and the added benefit could be offset by the added interface resistances from stacking the multiple MET layers. The extended HS shows better cooling effectiveness since it improves both the in-plane, as well as the through-plane heat conductance, thus spreading the generated heat to a wider span thus reducing the operating temperature. It is important to note that the cooling is calculated with respect to the base case where the square CMP fill is present. The additional temperature reduction is the added effect of the heat spreader.

7.3.4. Parametric Study

The developed thermal model representing the structures shown in Figure 7.5 is experimentally validated and thus accurately represents the observed thermal behavior of the actual device. The model can thus be used to extract the full 3D temperature distribution and infer the thermal response of embedded features that are otherwise inaccessible with optical surface measurement techniques [211]. Furthermore, the model can be used to conduct more general parametric studies and what-if scenarios to understand the contribution of different parameters in the design, such as geometry, layout, and material. This experimentally-driven iterative modeling thus allows device researchers to perform optimization studies that may otherwise require extensive experimental testing.

As an exhibit of this capability, and since only samples of two HS widths are fabricated for testing, the parametric study chosen is to determine the optimal HS width for cooling of the active interconnect. Using the validated model with the optimized input parameter file, the model is run for the different HS width cases and for cooling a device of different widths. The reduction in operation temperature ∆T, shown in Figure 7.8 (Top), and the increase in allowable activation current density $J_{RMS}$ (Bottom), are reported for the wider HS in comparison to the narrow HS design. The reduction in self-heating by the HS cooling allows the relief of this activation limit by a certain percentage which is presented in the
bottom graph of Figure 7.8. The increase in $J_{RMS}$ allowance in this study considers a max temperature rise of 15°C. Since wide HS configurations ($w_{HS} > w_{dev}$) are seen to improve heat spreading, the results are plotted against the excess HS width $w_{ex} = 0.5 \times (w_{HS} - w_{dev})$ which allows to determine what is the optimal HS dimension for a given interconnect width.

Figure 7.8. Cooling effectiveness of different widths of HS for different micro-resistor widths (Top) and the associated percent increase in current density $J_{RMS}$ allowance (Bottom).

As a first observation, any added HS width improves heat spreading, by that reducing self-heating in the interconnects and increasing the allowable activation limits. For our 10 $\mu m$ resistor, an additional width of 10 $\mu m$ at each side is sufficient to obtain most of the cooling effectiveness of the HS. In fact, the 30 $\mu m$ wide HS is sufficient for achieving optimal cooling for all the other narrower microresistors. Another observation is that the HS is
less effective for the wider resistors since the HS has to dissipate the heat from a wider generation region. The maximum cooling rates of the HS configurations are summarized for the different interconnect widths in Table 7.2. The above results are reported for the actual thermal response obtained from the experimentally-validated numerical model, and thus offer accurate rates of cooling as would be expected from the actual studied devices.

Table 7.2. Maximum Added Cooling and Increase in Allowable Activation Level for the different width Interconnects, as Compared to the Case with Narrow (N) HS Configuration.

<table>
<thead>
<tr>
<th>Width (µm)</th>
<th>( \frac{(\Delta T-\Delta T_N)}{\Delta T_N} ) (%)</th>
<th>( \frac{(J_{RMS}-J_{RMS,N})}{J_{RMS,N}} ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>35.5</td>
<td>24.5</td>
</tr>
<tr>
<td>4</td>
<td>33.9</td>
<td>23.0</td>
</tr>
<tr>
<td>6</td>
<td>30.2</td>
<td>19.7</td>
</tr>
<tr>
<td>8</td>
<td>25.9</td>
<td>16.2</td>
</tr>
<tr>
<td>10</td>
<td>25.2</td>
<td>15.7</td>
</tr>
</tbody>
</table>

After careful validation of the model, other parametrization studies are possible such as determining the location of the HS within the stack height, the spacing of the two MET layers, or even the pattern layout shape to eventually attain an optimal cooling for the more thermally-critical higher level devices of the BEOL.

7.4. Conclusions

This study presents an investigation into the added thermal benefit of incorporating CMP dummy fill into MET layers of the BEOL of ICs. A numerical study and experimental investigation proves that the high fill density improves the heat conductance within the stack, and can offer a passive cooling solution for thermally critical interconnects. By enhancing heat dissipation from active BEOL regions, the operation temperature and thus the degradation rates can be reduced, which improves the BEOL network reliability and lifetime. Moreover, with enhanced thermal management, designers can allocate higher activation levels for ICs or can use less conductive material in order to reduce costs. Finally, the presented investigation
tool combines experimentation with numerical modeling to obtain a device accurate thermal model, which can be used to efficiently conduct wider-ranging parametric studies [98, 99] that would otherwise require vast experimental efforts and resources.
The previous Chapters have presented technologies where temperature has an adverse effect on microelectronic devices. The thermal characterization techniques were thus applied in conjunction with strategies for thermal management to reduce self-heating and mitigate the degrading effects of elevated operation temperatures. For the technology presented in this Chapter, the thermal effects are not particularly disagreeable but are a major factor in the driving mechanisms behind interesting electrical characteristics, such as those observed in memristive devices. For that purpose, the analysis and modeling of these thermally-driven electrical behaviors cannot occur without a full understanding of the underlying thermal behavior and response of memristive devices.

This Chapter presents a thermal investigation into the response of novel Niobium Oxide based memristors, using the advanced characterization techniques presented earlier. This investigation enables and assists in the development of emerging memristor technologies, which would allow the modeling and integration of the novel devices into functional systems and applications.

8.1. Memristor: the Fourth Fundamental Passive Circuit Element

In addition to the resistor, capacitor, and inductor, a fourth fundamental passive circuit element was postulated by Leon Chua in 1971 by reasoning that there should exist six equations relating each pair of the four fundamental electrical parameters: current $i$, voltage $v$, charge $q$, and flux $\phi$ (Figure 8.1). The fourth equation described memristance behavior which relates the applied flux variation ($\phi$) to a change in electric charge distribution ($q$) through $M$, to be later described. This equation reduces to Ohm’s law for a constant
coefficient $M = d\phi/dq$, as current and voltage are the time-derivatives of flux and charge. However, the applied flux varies the charge distribution and thus also changes the material’s resistance.

![Figure 8.1](image)

**Figure 8.1.** The four fundamental passive circuit devices relating the four electrical parameters ($i, v, \phi$, and $q$).

This activation control over the device’s resistance gives the device a non-linear behavior, which is very desirable characteristic in non-linear circuit design, and would otherwise require active circuit elements to reproduce. Fortunately when compared to active circuits, those based on this passive element would then be exceedingly more compact and energy efficient systems. Moreover, the control over the variable electrical resistance provides a memory effect where the device can hold on to its resistance, much like a data storage device, even when the device is inactive. This promotes applications in memory and computing, and thus granting the descriptive term of the passive element as a memory resistance ($M$) or a "memristor".

Without any actual device implementation of the memristor, later developments observed memristive characteristics in thin oxide layers that experience exceptional behaviors, such as electro-luminescence, electron emission, and most notably variable resistance switching. [69]. Memristive characteristics is observed in transition metal oxides as Negative differential re-
Negative Differential Resistance (NDR) behavior in two types, a current-controlled (S-type) and a voltage-controlled (N-type) behavior. The S and N-type terminology refers to the shape of the I-V characteristics curves as shown in Figure 8.2. The S-type NDR exhibits a voltage drop when subjected to a current sweep, while the N-type exhibits a current drop under a voltage sweep. The displayed response depends mainly on the oxide material, with some oxides, such as Tantalum (Ta) and Titanium (Ti) oxides, can exhibit both types of NDR behavior depending on the fabrication process.

![Figure 8.2. IV characteristics showing N-type and S-type Negative Differential Resistance (NDR) behavior.](image)

This change in resistance allows different resistive states at different activation levels, thus allowing for bistable switching, as well as inducing a resistance change which could be sustained in the device. This sparked an interest in developing multiple transition-metal oxide based devices for applications in non-volatile memory and neuromorphic computing technologies.

### 8.2. Transport phenomena in Thin Oxides

The major transport mechanisms identified for the observed electrical response in memristors are 1) the electric field induced Frenkel-Poole conduction and 2) the Joule heating induced Mott insulator-metal transition. Frenkel-Pool conduction (PFC) is driven by trapped electrons which are loosened from their localized states to the conduction band under the ap-
plication of a high electric field. Higher thermal energy enhances this mechanism as excited electrons have more energy to escape to the conduction band and contribute to electrical conduction. The thermal effect renders this mechanism non-linear in nature and the effect is most dominant at moderate temperatures of around 500 K [238].

Temperature induced insulator to metal transition, also known as Mott transition, is another mechanism attributed to the NDR behavior. The transition occurs at comparatively higher temperatures than PFC, that can be reached under thermal runaway, as the electrical and thermal conductivity of the oxide layer increases with temperature, leading to localization of electrical current and heat generation to a narrow filamentary path (Figure 8.3) which could irreversibly transition to metallic phase if the temperature exceeds that of the Mott transition (≈1000 K).

![Figure 8.3. Metallic filament forms at temperature exceeding the metal-insulator-transition (MIT) temperature (≈1000 K)](196)

Other transport phenomena are also described for insulation material, namely Hopping conduction, Space-charge-limited current (SCL), and Tunneling. However, these phenomena are not dominant for the scale and electrical activation levels studied in this paper.

### 8.3. Applications of memristor devices

The thin transition-metal-oxide devices portray several electrical behaviors that could be employed in several applications. The bistable switching between high and low resistive states can be used in the development of passive microdevices for switching and diode applications.
These devices are based on passive memristive devices and thus require minimal power consumption. Non-linear behavior of the memristive components are also important for non-linear circuit applications and can replicate active components’ oscillatory behavior. This study is mainly concerned in NbO$_x$ current-controlled NDR devices which have been mostly implemented in memory and neuromorphic computing applications.

8.3.1. Non-volatile memory

A memristor consists of a medium that exhibits variable resistance depending on the electrical activation at the two electrodes as shown in Figure 8.4. The resistance value persists even after the device is inactive, allowing for the storage of information as a resistance value, instead of binary states. A demonstration in Figure 8.4 presents the basics of operation where the depletion ions, for instance oxygen, is dispersed through the material upon the application of a voltage potential, thus varying the resistance of the oxide channel. This electro-chemical phenomena became the basis for emerging non-volatile technologies such as resistive random access memory (ReRAM) where a crossbar array of memristors could replace present technology of volatile active memory components.

Moreover, this resistance change persists in the memristor device. In some cases the resistance value is sustained by the current flow, thus removing or reversing the electric field would dissolve this channel. For other materials, the device remains in its resistance state once it is switched without the need for continuous electrical current. This behavior is known as ‘memory switching’ and offers diverse application potential in non-volatile data storage systems.

Another negative differential resistance behavior was observed as an abrupt transition, and also termed ‘threshold switching’. The effect has been contributed to one of the two conduction mechanisms, namely Poole-Frankel and Mott transition, which occur at elevated temperatures and accelerated by Joule heating run-off. ‘Threshold switching’ crosspoint switches where the high or low resistance state can be maintained as long as the suitable activation current is passing through the device.
8.3.2. Neuromorphic computing (neuristor)

Technological progress has been for the past decades based on a steady increase in computing power and a decrease in chip production cost. This trend, known as Moore’s law, has been slowing down as chips reach their size limit and the cost to sustain the trend becomes more inflated. Even with the most advancements, the most powerful supercomputers have failed to replicate biological intelligence of beings more advanced than a spider and a mouse. [118] Emulating the computing speed and complexity of higher brain functionality is exponentially more complex and power consuming, if the same computing architecture is used. To overcome those limitations, researchers are trying to shift the computing architecture to more efficient and faster methods, with options like quantum and neuromorphic computing in development.

Neuromorphic chips are a crude attempt to design electronic chips that model the functioning of the human brain and mimic human intelligence. The method is based on a densely connected network of components that allows the combined storage and processing of data, replaces traditional architecture, and offers the physical implementation to support neural network software algorithms. This requires processing and storage to be combined at the component level, which is is different from the traditional Von Neumann based approach where the CPU and memory are separated. This interconnectivity at the smallest scales is inspired by actual biological neural networks with components aiming to emulate the activation mechanisms in the neural synapses as shown in Figure 8.6.
Memristors have been in consideration for neuromorphic applications for their favorable characteristics. Much like neurons create connections, memristors can change resistance to reinforce a neural node or weaken it. Moreover, CC-NDR memristors can be used for controlled current emission, much like the sodium and potassium ion channels in the synapse. Similar activation response behavior have also been compared by Ju et al. as presented in Figure 8.5 The ‘threshold switching’ characteristic present in memristors can also mimic all-or-nothing action potential that occurs at the axons of neurons.

Figure 8.5. Activation synaptic behavior in memristors as compared to an actual neuron [118].

Memristors size also allows for large scale integration in crossgrid configurations shown in Figure 8.6. This array has the potential to improve computations by performing highly parallelized execution of the weighted sum computation, or the matrix-vector product, thus solving a bottleneck in CMOS based neuron circuits [86]. With such computing capabilities combining memory, computation and communication, memristor based chips would be able to concurrently process information, learn and adapt from its experiences. These similarities between the artificial and the biological systems reveal the feasibility of neural-based computing systems that would allow us to understand and mimic the highly-efficient biological computing capabilities. [86, 196]
8.4. Temperature-induced CC-NDR switching characteristics

The mechanisms behind memristor operation has been associated with, and are largely determined by the device temperature. Thus, the thermal aspect is crucial in understanding and modeling the electrical and switching characteristics of the device. The temperature can affect the switching characteristics of the device [6] and Wang et al. even experimentally demonstrated that thermal properties of the device can be used to tune its electrical characteristics [260].

The dominant mechanism behind the distinctive I-V characteristics has been long debated, in particular the snapback switching and the nonlinear resistance change. Kumar [140] examined NbO₂ and found that S-type NDR is caused by current controlled (Frenkel-Poole) conduction and occurs at temperatures of around 400 K, while SB switching is caused by insulator-metal transition (IMT) and occurs at temperatures of around 1000 K. The study presented thermal and chemical spectro-microscopy results as well as simulations. The microscopy results even show the formation of a metallic phase beyond the high temperature IMT switching. Alexandrov [6] studied the similar effect using global heating to determine the temperature effect on the current characteristics in TiO₂. The author concluded that a temperature 600 to 650 K is sufficient for TiO₂ transition to anatase phase. These re-
results seem to show conclusive evidence about the dominant switching mechanism, however other researchers have presented differing viewpoints. Gibson [230] showed using a thermal compact model that thermally induced NDR occurs at temperatures of $380 \text{ K} \approx 400 \text{K}$, fairly lower than metal-insulator-transition temperatures ($\approx 1000 \text{K}$). This was also observed by Wang et al. where the maximum temperature observed from quasi-static electrical modeling does not exceed $300 \text{°C}$, which is well below the MIT temperature of $800 \text{°C}$. [259] Slesazeck [230] also estimated the switching temperature to be around $550 \text{ K}$. The researchers argued that moderate temperature rise of $150 \text{ K}$, well below the MIT temperature, is sufficient for switching. They stating that the Mott transition is not dominant, but can exist in conjunction with the switching if the temperature exceeds the MIT temperature ($\approx 1000 \text{K}$), and that the low-temperature thermal-feedback mechanism is responsible for the switching.

As there has not been any final consent on the dominant mechanism governing the switching, with low filament temperature ($<500 \text{ K}$) or high temperature ($>1000 \text{ K}$) pointing to the most presently debated mechanisms of Poole-Frankel and Mott transition respectively, conducting an advance experimental investigation into the thermal characteristics of NDR devices would provide an insight on the controlling mechanisms and allow accurate prediction and modeling of the device behavior. Presented in this chapter, and for the first time, in-situ thermoreflectance imaging and quasi-static current voltage characteristics is used to investigate the thermal mechanism behind snap-back switching in NbO$_2$ and to validate the previously core-shell model presented by the co-authors. The TR high-resolution thermal maps also provide an estimate of the filament temperature to finally resolve the long-standing controversy of the origins of the NDR characteristics. In addition, TR measurements will be used to validate the model and investigate the effects of stoichiometry, device dimensions, polarity, and temperature on the observed NDR device switching behavior. This would provide a physical basis for understanding and predicting (i.e., modeling) the behavior of the novel device technology.
Figure 8.7. Transmission electron microscopy images showing the metallic phase formed after device switching [140].
8.5. Sample Description and Methodology

The thermal study is performed on Pt-NbO\textsubscript{x}-Pt crosspoint devices. The devices are fabricated on thermally oxidized Si (100) with a patterned electrode terminals (Figure 8.8) of Pt deposited using e-beam evaporation. The sub-stoichiometric NbO\textsubscript{x} is deposited using RF-sputtering of an Nb\textsubscript{2}O\textsubscript{5} target. The crosspoint devices (metal-oxide-metal (MOM) junctions) are patterned with of different overlap sizes of 2, 5, 10 and 20 µm. Each device consists of a central bottom electrode (BE) that is common for four identical peripheral junctions connected each to a top electrode (TE). Thin deposited electrodes are utilized to minimize their heat absorption and to obtain measurements closest to the actual filament temperature. The different widths will allow the analysis of size effect on the device electrical, thermal, and switching characteristics.

Initial WL scan and test runs revealed that the platinum (Pt) surface material of the electrode has a very low thermo-reflectance response ($C_{TR} < 0.5 \times 10^{-4}K^{-1}$) which lead to a low signal-to-noise ratio SNR.

Prior to utilizing thermoreflectance imaging, collaborators had used the photoresist (PR) cover on the Pt electrode surface to identify the location of the filament formation, indicated by the a burn-through the PR due to the generated hotspot from the localized electrical conduction. In our initial measurements, the PR covered Pt revealed a high TR response, which was beneficial to observe the temperature rise with a good SNR. However, the PR covered electrode was limited to moderate temperature rises since the PR would evaporate and leave surface defects. For that purpose, An initial investigation to observe the temperature rise pre-electroforming, where the observed temperature rise is uniform within the cross-point region, and no hot-spots are present.

Mapping the temperature response post-electroforming would require another approach since the PR would be denatured at the filamentary region and would not be viable for a temperature measurement. To improve the TR response (i.e., better SNR) while having a stable surface for measurement, a thin layer of Gold (Au) is added over the Pt electrode. Gold is known to exhibit a desirable TR response and thus would serve as a "transducer"
Figure 8.8. a) Device structure as seen under 20× magnification showing the four patterned top and bottom electrodes (TE and BE) and the resulting four cross-point devices. b) Intensity map of crosspoint device under 100× as seen with the camera under monochromatic illumination. c) Schematic of the crosspoint devices showing the Pt electrodes and the NbO\(_x\) thin-film.
layer for the TR measurement \[131,258\] providing a higher TR signal.

The sample test structures are thermally mapped using the T°Imager system presented in Chapter 2. The samples are placed on the thermal chuck under an illumination that ensures maximum TR response as determined initially by a WL scan. The devices are mapped with a 50× magnification which provides sufficient resolution for the 5, 10 and 20 µ size devices (pixel size \(\approx 0.15 \, \mu m\)).

The devices are activated in-situ (i.e., during thermal mapping) with a current control (CC) activation using the Keithley 2010. Ex-situ electrical measurements (I-V sweeps) are performed independently to characterize the switching characteristics of the devices. The bottom electrode is grounded while forward (positive) or reverse (negative) bias are applied to the top electrode.

Other device characterizations are performed by the collaborators to complement the thermal investigations such as Rutherford backscattering spectrometry (RBS) to measure the NbO\(_x\) stoichiometry \(x\), with \(x\) directly related to the film electrical resistivity, and X-ray diffraction (GIAXRD) to verify the amorphous crystal structure of the grown NbO\(_x\).

Initially in Section 8.6, the activated temperature profiles are measured with activation levels prior to electroforming. After electroforming, the temperature is later measured with respect to the snapback switching. The method will approximate the electroforming temperature and switching temperature of the memristor and infer the dominant switching mechanism. In Section 8.7, a model is presented to predict the different NDR switching mechanisms based on a thermal-runaway and current bifurcation process. The thermal response behind the observed electrical behavior is presented to verify the reduced electrical model. In Section 8.8, the polarity dependent NDR characteristic is modeled by investigating the Schottky barrier asymmetry in shell resistance. Thermal investigation verifies and correlates the different NDR characteristics with the thermal signature of electrical conduction in the crosspoint device.
8.6. Electric Field- and Current-Induced Electroforming Modes in NbO$_x$ [176]

To establish the memristive behavior in NbO$_x$ devices, one step electroforming is required, during which the crosspoint device is subjected to high current or potential field conditions. This generates a conductive electronic path (i.e. soft dielectric breakdown) which results in the characteristic NDR characteristics.

For NbO$_x$ two modes of EF observed, one being field-induced and another thermally-induced under high current conditions. The field induced EF relies on high-field electrolytic process that occurs on the oxide/electrode interface by oxidation/reduction reactions and creates a conduction path (Figure 8.9). The second observed mode of EF occurs is current induced and occurs due to current bifurcation and thermal runaway. This occurs since the electrical conductivity of the oxide film increases with temperature, resulting in a feedback mechanism that localizes the current in the regions of higher temperature, and leading to more localized Joule heating. Due to the thermal boundary conditions, the current localization leads to filament generation in a central region of the cross-device, while the field-induced EF results in randomly generated filament formation depending on local electric fields.

The EF mode that occurs for an NbO$_x$ device depends mainly on film conductivity where low-conductivity films (high $x\approx 2.6$) exhibit field induced EF with oxygen-vacancy filament. High conductivity films with significant current densities (low $x\approx 1.92$) exhibit sufficient thermal heating within the device which results in transient current-induced filament. correlated by filament formation location, I-V measurements and thermoreflectance imaging. More results and discussion are presented in [176].

To validate the thermally-induced electroforming in high conductivity NbO$_x$ films, thermoreflectance imaging is applied to the a 10 µm test device (high-$\sigma$, $x=1.99$) while activating with increasing current levels. The I-V characteristics and the thermal mapping results (Figure 8.10) reveal that initially, the temperature distribution is broad and rises monotonically with current levels. A snapback response is observed at around 6 mA indicative of an electroforming event while, concurrently, the temperature distribution narrows abruptly and exhibits a large jump in the peak temperature. This peak rise in temperature of around 200
Figure 8.9. Electroforming and filament distributions: (a) electroforming characteristics of a 5 µm Pt/Nb/NbO2.6/Pt device, (b) electroforming characteristics of a 5 µm Pt/Nb/NbO1.92/Pt device, (c) schematic representation of filament distribution in a 5 µm × 5 µm cross-point device with low-conductivity NbO_x (x = 2.60) films, (d) schematic representation of filament distribution in a 5 µm × 5 µm cross-point device with high-conductivity NbO_x (x = 1.92) films, (e) schematic of an oxygen vacancy filament formed by the field-induced generation, drift, and diffusion of oxygen vacancies, and (f) schematic of a transient current filament due to current bifurcation. Note that each dot in (c, d) represents a single device and the filament distributions represent 75 separate measurements for each film [176].
K is accompanied with a reduction of about 50 K in the region surrounding the filament.

The observed thermal response associated with electroforming and the current bifurcation is also verified using a finite element model of the bifurcation process. A 2D axisymmetric model where the electrical conductivity was governed by Poole-Frenkel conduction supports the experimental observations (Figure 8.11). The model self-consistently solves the heat transfer and current continuity equations. More results and discussion are presented in [176].

The results presented above however do not explain why some devices exhibit S-type or SB-type NDR behavior, and the dependence on device parameters. For that purpose, a lumped electrical model is presented in the next Section to explain the above NDR behaviors.

Initial thermoreflectance results show different heating profiles for the three activation schemes. Before electroforming of the conductive filament, global heating is seen at the electrodes with heating extending to the leads. After the filament forms, localized heating is observed at a central region of the crosspoint. Moreover, observations show that Joule heating in the filament causes a parabolic rise in temperature up to a steady value of around 450 K where switching occurs. At higher activation levels, the thermal maps show that the hot spot changes location and that the peak temperature rise remains constant. Thus any increase in current intensity causes the filament region to expand accommodating the increased Joule heating keeping the temperature uniform. The post-switching temperature correlates well with the values measured by Gibson and Slesazeck [87,230,259]. Another explanation for the dynamic phenomena is that the heating could have evaporated or deformed some material from the electrode or the oxide, thus destroying the filament and causing a new filament to form at another location [18].

8.7. Current Localization and Redistribution as the Basis of Discontinuous Current Controlled Negative Differential Resistance in NbO$_x$ [177]

Beyond the initial electroforming of the crosspoint devices, two distinctive modes of negative differential resistance are observed during a current-controlled activation. The first is characterized by an S shape I-V behavior is denoted as an S-type NDR. The second
Figure 8.10. In situ thermoreflectance measurements: a) in situ (points) and ex situ (line) current-voltage characteristics of a 10 µm × 10 µ cross-point device with a high conductivity NbO$_x$ (x=1.92) film, and the temperature of the filamentary and surrounding regions of the film (shown by circles in panel b), b) Two-dimensional temperature maps of the top electrode surface for different device currents. The inset shows a back-scattered electron image of the filamentary region in the oxide film (i.e., after removing top electrode).
Figure 8.11. Finite element simulation: a) current-voltage characteristics during bidirectional current sweep, b) maximum temperature in oxide film as a function of device current, c) current density distribution before and after bifurcation, d) temperature distribution before and after bifurcation, e) before (point A) and f) after current bifurcation.
Figure 8.12. Core-Shell model and current bifurcation: a) schematic of the CS model of a negative differential resistance device, b) electroforming characteristics of a 5µm × 5µm device with NbO$_x$ (x-2.22 film), c) electroforming characteristics of a 10µm × 10µm device with NbO$_x$ (x-2.22 film), d) in situ thermoreflectance maps of 5µm × 5µm device during electroforming e) in situ thermoreflectance maps of 10µm × 10µm device during electroforming.
Figure 8.13. Structure of the studied NDR crosspoint device showing snapback switching and saturation of filamentary conduction after current bifurcation.
observed characteristic denoted as a snapback switching (SB-NDR) shows a sudden voltage drop in the I-V sweep beyond a certain threshold (Figure 8.14).

Figure 8.14. I-V characteristics showing a) S-type NDR behavior and b) snapback (SB) NDR behavior. c) model of conductive paths in transition metal oxide (NbO$_x$) layer showing core and shell with representative electrical resistance. d) representative electrical circuit with constant current source, and the respective resistance of shell and variable resistance (memristor) of core [177].

Extensive I-V measurements are conducted by the authors and has lead to observations that the switching type and threshold values are dependent on crosspoint area, area and temperature. The threshold current and voltage ($I_{th}$ and $V_{th}$) are extracted from the I-V characteristics as well as the switching mode. Results and the major conclusions are presented in the article [177].

While the origins of S-type behavior in memristive device have been agreed upon, the origins of the SB-type switching remains elusive and controversial. Initially thought to
originate from IMT, but the phase transition fails to account for similar SB behavior occurs in other transition-metal oxides such as TaO$_x$, NiO$_x$, and in SiO$_x$. It has been correlated through modeling that the switching behavior is caused by a current redistribution into regions of high and low current density, which served to develop a electrical model to represent the different observed NDR behaviors in NbO$_x$ [145].

From initial observation, the switching behavior of NbO$_x$ based memristors can be described by a mechanism of current localization from two regions with different electrical conductance (Figure 8.14). A thermal runaway behavior due to positive feedback between increase in thermal conductance and electrical conductance leads to a filamentary conduction and a sudden drop in voltage.

The developed electrical resistance model is based on representing two regions, one region of high current localization within the filament (core), and another within the surrounding film (shell) which acts as a parallel resistance to the core. The threshold switching would thus depend on the relative magnitudes of the core NDR and shell resistance, denoted by $R_{NDR}$ and $R_S$.

The distinct NDR behaviors are then explained by comparing the relative magnitudes of the core and shell resistances.

1. If $R_S > R_{NDR}$, the current is localized at the filament and S-type characteristics are observed.

2. If $R_S < R_{NDR}$, SB characteristics are observed. Given that $R_S$ depends on conductivity (stoichiometry), area and thickness, then for some critical value of those parameter the device might exhibit a transition from S-type to SB-type NDR

To verify the model presented by the co-authors in [145], some sort of physical investigation is required to observe the current bifurcation and analyze the spatial distribution. For this purpose, the thermal signature within the cross-point device is experimentally investigated to observe the current density distribution during the different experienced NDR behaviors. Three devices, when observed under high-resolution thermoreflectance imaging,
helped correlate the observed electrical characteristics with the current localization and redistribution. The devices of three different observed features are presented (Figure 8.15): one with S-NDR and a permanent filament, one with a SB-NDR and permanent filament, and one with SB-NDR without a filament.

1. Devices with permanent filaments show localized heating

2. Shell temperature for S-NDR increases monotonically with filament temperature

3. Shell temperature for SB-NDR decreases while filamentary temperature increases rapidly

Show that snap-back is associated with current localization to a pre-existing filament, or with current bifurcation and the redistribution of current between low and high current-density domains [177]. Moreover, the measured highest filament temperature beyond switching does not exceed a few hundred degrees (200 to 400 K rise) and does not reach the Mott temperature needed for metal-insulator transition, verifying that a sustained switching can be achieved without any phase transition.

8.8. Schottky barrier induced asymmetry in the negative differential resistance response of Nb/NbOₓ/Pt cross-point devices [178]

The NDR response in Nb/NbOₓ/Pt is observed to be dependent on the polarity of the applied potential. This in some cases has resulted in distinct NDR types at the different polarities, in addition to the bipolar S-type and bipolar SB-type. The different NDR behaviors are observed using TR imaging to be associated with strong current localization thus justifying the validity of the developed core-shell model.

To account for the polarity dependence, the contact resistance contributed by the metal/oxide Schottky barrier is added to the model and verified for the observed NDR behavior at different stoichiometric values and device areas. This contact resistance results in an asymmetric $R_S$ which can be designated as $R_n$ and $R_p$ in negative and positive polarity respectively. This asymmetry can lead to three scenarios according to the reasoning presented in the core-shell model:
Figure 8.15. In-situ temperature measurements of S-type and snap-back NDR. a) Current-voltage (I-V) characteristics and average temperature rise of the filament and surrounding area (the area used for averaging is indicated by boxes in b) as a function of applied current, b) 2D maps of the surface temperature rise in a 5 µm device operating at 1 mA and 6 mA, c) Temperature (current) localization of S-type NDR in the post-formed device at different current levels as shown in a, d) I-V characteristics and average temperature rise of the permanent filament and surrounding area (indicated by boxes in e) as a function of applied current, e) 2D map of the surface temperature rise in a 10 µm device at pre-threshold (4 mA) and post-threshold (10 mA) currents, f) Temperature (current) localization in a post-formed device at different currents as shown in d (the blue arrow indicates snap-back transition), g) I-V characteristics and average temperature rise of the current filament and surrounding area (indicated by boxes in h) as a function of applied current, h) 2D map of the surface temperature in a 10 µm device at pre-threshold (4 mA) and post-threshold (12 mA) currents, i) Temperature (current) localization due to current bifurcation in the device without permanent filament as shown in g (the blue arrow indicates snap-back transition) [177].
1. $R_p > R_n > R_{NDR}$ results in bipolar S-type

2. $R_p > R_{NDR} > R_n$ results in SB in negative bias and S-type in positive bias

3. $R_{NDR} > R_p > R_n$ results in bipolar SB-type

These three scenarios are observed when different stoichiometric values and device areas are tested (Figure 8.8)

Figure 8.16. (a-c) Asymmetric NDRs observed in Nb/NbO$_x$ 1.99/Pt MOM under positive and negative polarity. (d) Matrix representation of dependencies of NDR on stoichiometry and device area.

"In situ thermoreflectance imaging is performed on 10-µm NbO1.99 devices to better understand the current distribution during post-forming current sweeps. The results are
summarized in Figure 8.8 and include in situ I-V characteristics and the measured temperature distributions for devices subjected to positive and negative bias. Under positive bias the temperature distribution is highly localized in the filamentary region over the entire range of operating currents and the temperature in this region and that of the surrounding film increase monotonically with increasing current. This clearly demonstrates that the resistance of the filamentary path is lower than that of the surrounding device and that the S-type NDR characteristic is dominated by the temperature-dependent conductivity of the filamentary region.

In contrast, under negative bias, the temperature distribution is near uniform at currents below the threshold for snapback and only becomes localized for currents near the threshold value [indicated by the point “F” in Figure 8.8(b)]. At this point the temperature increases abruptly while that of the surrounding area decreases, consistent with current redistribution and localization due to the positive feedback created by local Joule heating, as previously reported [28]. This demonstrates that the resistance of the filamentary region is initially comparable to, or greater than that of the surrounding device, but becomes relatively less resistive as the current increases to the threshold value. Such a change implies that the conductivity of the filamentary region increases more rapidly with temperature than that of the surrounding device. For filaments near the center of the device this may also be facilitated by the nonuniform temperature distribution created by the device geometry [see Figure 8.8(c)]. It is important to note that the extent of the filamentary region is smaller than that of the temperature distribution observed at the top electrode surface due to the high thermal conductivity of the metal electrode [25,28].”

8.9. Thermal Conductivity of Amorphous Sub-stoichiometric NbO_x for Memristors Applications using Transient-Domain Thermoreflectance

Despite of the general understanding of thermal environment during electroforming or switching, the description of the temperature field belying memristors operation remains incomplete due to lack of their thermal properties. Therefore, it is very important to know to
Figure 8.17. a), b) In situ I -V characteristics of a 10×10 µm² Au (75 nm)/Nb(5 nm)/NbO1.99 (approximately 45 nm)/Pt (25 nm) device showing S-type NDR under positive bias and snapback under negative bias, respectively. Insets show the in situ thermoreflectance (ΔR/R) maps of the 10×10 µm² device during corresponding bias polarity. c) Temperature profile through the filamentary region for different points in the current-voltage curve during positive bias. d) Temperature profile of the same device for different points in the current-voltage curve during negative bias. Note: the temperature map and the I-V characteristics shown in Figure 3 are also obtained for 5×5 µm² devices with lower threshold currents, and a snapback response observed under negative polarity and S-type NDR under positive polarity. But in this case, the filament is observed to form around the edges due to the dominance of edge effect in smaller area cross-point devices with a thick (75-nm Au in this case) capping layer [25]. The short-dashed curves in c) and d) represent the experimental data and the smooth lines show Gaussian fits of the corresponding experimental data. A, B, and C denote sub-threshold, threshold, and post-threshold points respectively in the I-V characteristic obtained under positive bias, while D, E, and F denote the same for negative bias polarity [178].
have thermal conductivity values as a function of stoichiometry. However, the computational prediction of their switching dynamic depends on the temperature distribution relies on approximations, as the thermal properties of the NbO$_x$ as a function of stoichiometry ($x$) have not been investigated.

8.9.1. Experimental Methods

NbO$_x$ dielectric layers of variable composition were subsequently deposited using either RF sputtering of an Nb$_2$O$_5$ target in an Ar ambient or DC sputtering of an Nb target in a variable O$_2$/Ar ambient to deposit sub-stoichiometric NbO$_x$ films. Details of the deposition conditions are given in Table 1.

A gold absorptive/transducer layer with a thickness set to maximize the measurement sensitivity is deposited using electron beam evaporation. The conductivity measurements are conducted using the time-domain thermoreflectance (TDTR) method that provides the thermal diffusivity of a thin film stack in the through plane direction. The Transometer TDTR is used to conduct the conductivity measurements. Details of the TDTR setup and methodology are presented in Chapters 2 and 3. The NbO$_x$ samples are placed on a hotplate with temperature control (Temptronic TP04310) that can vary the sample temperatures between 20 and 180°C for the temperature dependent analysis.

8.9.2. TDTR Results and Discussion

<table>
<thead>
<tr>
<th>Layer</th>
<th>Ar/O$_2$ (sccm/sccm)</th>
<th>Power (W)</th>
<th>Pressure (mT)</th>
<th>x in NbO$_x$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb$_2$O$_5$</td>
<td>20</td>
<td>180</td>
<td>4</td>
<td>2.6±0.05</td>
</tr>
<tr>
<td>NbO$_2$.11</td>
<td>19/2</td>
<td>150</td>
<td>2</td>
<td>2.10±0.03</td>
</tr>
<tr>
<td>NbO$_2$.03</td>
<td>18.5/1.5</td>
<td>150</td>
<td>1.5</td>
<td>1.98±0.04</td>
</tr>
<tr>
<td>NbO$_1$.88</td>
<td>19/1</td>
<td>150</td>
<td>1.5</td>
<td>1.88±0.03</td>
</tr>
</tbody>
</table>

* Layer thickness is function of deposition time
Figure 8.18. a) Schematic representing the Pump-and-Probe setup used to conduct transient-domain thermoreflectance conductivity measurements. b) TDTR normalized response showing the heating and temperature decay profile as measured at the gold surface of the blanket film samples.

Figure 8.19. a) The TDTR response is presented as a function of temperature for the thickest (100 nm) sample. b) Thermal conductivity as a function of temperature for different stoichiometric NbOx films.
At higher temperatures, the total thermal resistance of the sample film decreases and the slope decreases revealing an increase in intrinsic conductivity of the NbO$_x$ sample with temperature, which is observed when $\kappa_i$ is plotted against $T$, varying approximately from 1.5 to 1.8 W/m·K. The measured conductivities in this study closely match the improved model by Agne et al [3], as it considers diffusons diffusivity. In regards to interface thermal effects, there is no significant increase with temperature where $R_{int}$ remains at a constant value around 30 m2K/GW.

The thermal conductivity of the sub-stoichiometric NbO$_x$ for the three measured values of $x$=1.88, 2.03, and 2.11 are reported for the same temperature range between 20 and 180°C (Figure 8.9.2). The intrinsic thermal conductivity for the different samples is seen to increase from around 1.2 – 1.5 W/m·K at 20°C to around 1.8 – 2.5 at a base temperature of 180°C. This increase with temperature is expected and is the basis of the thermal runaway mechanism that occurs in NbO$_x$ crosspoint devices which results in a filament region of higher thermal and electrical conductivity than its surrounding. The thermal conductivity range is within expected range for amorphous NbO$_x$ and shows a lower conductivity than crystalline NbO$_x$ (3-4 W/m·K [54]) due to the amorphous structure of the lattice. The effect of oxygen deficiency is not significant where the room-temperature thermal conductivity and the variation with temperature was similar for the three samples. The effect of oxygen deficiency on thermal conductivity is observed in crystal structures of reducible oxides (such as PrCrO$_2$ [155], LaSrCoO$_3$ [267]) where reduction of Pr$^{4+}$ to Pr$^{3+}$ causes lattice distortion. The direct effect of oxygen is slight since oxygen does not contribute to the mass of the compound.
Chapter 9

Characterization of Temperature Rise in AC Electro-thermal Flow Using Thermoreflectance Method [129]

Thermal concerns mostly related to degradation and premature failure in electronic devices have been the driving motivation for thermal characterization of devices operating under extreme power-density applications. However, in the micro-fluidic application presented in this Chapter, the reliability of the device operation is temperature sensitive instead of the devices itself. An example of a micro-fluidic operation, AC electrothermal flow (ACET) induced by Joule heating is utilized to transport biologically relevant liquids in microchannels using simple electrode designs. The operation is compromised by increasing temperature since the ACET Joule induced temperature rise may result in degradation of the chemical or biological fluidic specimens, and hence, rendering ACET impractical for biomicro-fluidic sensors and other possible applications. For that purpose, strict temperature control and characterization is necessary in micro-sensing applications to ensure a proper functioning, or else unchecked temperature levels may invalidate any sensing measurements.

For the above thermal concerns, the thermoreflectance method is applied in this study to investigate the temperature rise at the electrode/electrolyte interface during ACET flow. The TR imaging setup generally used to thermally map microelectronics systems, so to meet the challenging requirements of imaging within an active microfluidic channel, the method is slightly modified and an innovative approach which has not been previously reported is presented. The experimental findings reveal that Joule heating could result in an excessive temperature rise, exceeding 50°C at higher voltage levels (20 V_{pp}). The measured data are compared with the results of the enhanced ACET theoretical model, which predicts the temperature rise accurately, even at high levels of applied voltages. Overall, our study provides a temperature measurement technique, which is used for the first time for electrode/electrolyte
systems. The reported results are critical in designing biomicro-fluidic systems with significant energy dissipation especially for conductive fluids.

9.1. AC Electrothermal Flow for Electrokinetic Actuation of Liquid Media

Microfluidic platforms employed for point-of-care (POC) diagnosis in micro-total analysis systems (µ-TAS) are a great lab-on-chip driver due to their low cost, low power consumption, short analysis time, and their use of minute amounts of sample volumes. Most of the µ-TAS use physiological samples including urine, bile, cerebrospinal fluid, blood, saliva and other biological buffers such as Lysogeny broth, Mueller Hinton broth, phosphate buffer saline, and Dulbecco’s modified Eagle’s medium, where the conductivities of solutions vary from 0.4 to 1.8 S/m [83,153]. Among various techniques, alternative current (AC) electrokinetics represents a promising approach towards the development of a fully integrated µ-TAS where electrodes embedded in micro-fluidic channels are commonly used for electrokinetic actuation of liquid media [24,70,139,265]. AC electrothermal flow (ACET) is a dominant transport phenomenon for operating in high conductive media, and it is applied by the biomicro-fluidics community in the past two decades due to its predictable nature and ease of implementation, while requiring simple electrode structures [71,188,229].

ACET flow has been used to enhance the performance of µ-TAS, for instance, when the sensitivity is usually limited by the diffusion of target molecules to the ligand surface [223]. It was utilized to enrich the binding efficiency of on-chip immunoassay biosensors by Li et al., who presented an immunoassay experiment using PBS, which is a conventional solution in most biosensor and high ionic strength biological applications [116]. The ACET mechanism for pumping of cell suspended Dulbecco’s Modified Eagle Medium (DMEM, ≈1.7 S/m) has been investigated by Lang et al. who designed a configurable ACET circulatory pump by switching the voltage over the electrodes [142]. In another study, Lysogeny broth buffer (≈1 S/m) was used as a carrier medium, and the pumping was established by the ACET flow [266]. An ACET micromixer was presented by Cao et al., whose results showed that a mixing efficiency of 98% can be achieved in a microchannel [34]. ACET flow was also
used for active control of ion transport which can find broad applications in improving on-chip electrodialysis by inducing thinner ion-depletion layers [189], lowering the microchannel resistance for biosensing at nanoscale [144], and controlling the location of pre-concentrated plug of charged nanoparticles [181, 182]. Besides the aforementioned applications, ACET has been also employed for many diverse fields such as the homogenization of solutions of reagents in biochemical reactions [221], clinical diagnosis [150], and drug delivery [162].

9.2. Temperature Control for ACET Applications

It is of great importance to monitor and control temperature within the micro-fluidic device while utilizing ACET flow as a favorable tool. However, maintaining a relatively low temperature rise is a major challenge for high ionic strength processing, since higher conductivity fluids substantially increase the effects of Joule heating at the electrodes and within the fluid medium. At high conductivities, temperatures beyond a few degrees (T\textdegree C) are likely to occur under an applied AC potential of even a few volts. For instance, using PBS of electrical conductivity $\sigma=1$ S/m, the temperature rise could exceed 100\textdegree C for the 20 $V_{pp}$ applied voltage depending on the geometry of the system [90, 213]. Therefore, the ACET-induced self-heating must be characterized prior to any electric field excitation to avoid unwanted effects and ensure that ACET is applicable and biologically safe for the required applications [39, 66, 80, 150].

A variety of methods have been employed to measure the temperature of active micro-electronic devices. Infrared (IR) thermography is the most widely recognized and is mainly used for macroscale applications. The IR thermal metrology technique measures the amount of thermal radiation emitted from a surface and thus, it requires the surface material to be sufficiently emissive [102, 154]. This renders the method less than advantageous for electronic devices as they contain mostly reflective metallic features that would require special high-emissivity coatings. Moreover, the resolution of IR is diffraction restrained by the infrared wavelength to around 3-10 $\mu$m which further limits the application for micro-devices of similar scales. Other methods like Scanning Thermal Microscopy (SThM) [244] and $\mu$-
Thermocouples [142] can improve resolution but are contact methods that require direct access to the measurement region, which can interfere with the thermal system and affect the accuracy of the measurements [130]. Furthermore, direct access is unfeasible for most devices since the electrodes are embedded within the fluidic microchannel, and hence, an optical method such as Thermoreflectance (TR), or Micro Raman Spectroscopy would be required. Micro Raman Spectroscopy is a high resolution optical method that measures temperature changes based on inelastic or Raman Scattering of visible light [137, 233]. The method is highly complex in nature and is a single-point measurement approach that would require a translation stage for scanning, resulting in large acquisition times in addition to the expensive system setups. Raman scattering can also result from mechanical and electrical phenomena which would need to be decoupled from the temperature effect to avoid additional accuracy errors.

Recently, Rhodamine B (RhB) has been used as a temperature sensitive dye, exhibiting strong temperature dependent fluorescence in the range of 0-100 °C [137]. Although it provides a high spatial resolution, it is not preferable for biologically relevant studies due to its invasive nature. Besides that, thermosensitive dyes can be absorbed on a polymer-based channel surface because of its hydrophobic nature. Dye adsorption causes fluctuations in the baseline fluorescence intensity, resulting in inaccurate temperature data. In contrast, the TR method presented in Chapter 2, is an optical thermal metrology method that is non-destructive, non-invasive, and non-contact with a submicron spatial resolutions and without requiring special sample preparations, all of which make the TR method suitable for this experimental thermal investigation.

To the best of our knowledge, the extent of temperature rise due to Joule heating was not experimentally reported in the literature. In this study, the TR imaging technique described above is used to measure the temperature rise distribution on the electrodes in contact with a high conductivity medium in ACET activated flow. The findings will help in the design of new generation of high accuracy, low-cost, low power biosensors. This paper is organized as follows: In Section 2, we first present the samples fabrication methods and the experimental
setup. In Section 3, the results of self-heating for different activation frequencies, potential, and power dissipation are provided and discussed. Finally, conclusions are presented.

9.3. Materials and Methodology

![Process flow diagram and image of electrodes](image)

Figure 9.1. a) Process flow of electrode fabrication using photolithography technique and b) image of interdigitated electrodes fabricated on glass substrate. The micro-fluidic channel is indicated with red dashed rectangle.

The interdigitated electrode array used herein is a well-known geometric configuration for AC electrokinetic-based micro-fluidic platforms, as it yields large electric field gradients even at small applied potentials. The micro-fluidic device is fabricated using a standard photolithography technique (Figure 9.1). 2.5 × 2.5 cm² glass slide are used as substrate where the electrodes are fabricated by using a sputter-coater (EMS300TD, Emitech). Sputtering generates the thin layers of chromium (≈ 5nm) and gold (≈ 55nm) required for the electrodes. The substrate is then immersed in PG remover (Sigma-Aldrich) to remove the gold on the positive photoresist. The emerging device pattern consists of 3 pairs of interdigitated electrodes with 200 µm width (w) and 200 µm spacing (l) as shown in Figure 9.1. A micro-fluidic channel is made by sandwiching 400 µm height double-sided tape between two glass slides. The rectangular portion of the tape is cut using a craft cutter to create a micro-channel with 0.4mm×10mm×25mm (height (H)×width (W)×length (L)) dimensions and centered at the electrode region. Inlet and outlet ports are drilled with a diamond drill.
bit on the top glass slides where the electrodes are placed. The wires used for electrical connections are bonded using conductive silver epoxy (MG Chemicals). Figures S-1a) and S-1b) show the microelectrode fabrication procedure and the electrodes fabricated on a glass substrate, respectively. The electrical ports are connected to a function generator (Tektronix AFG3102) to induce the required sinusoidal excitation at a given frequency and peak-to-peak voltage ($V_{pp}$).

The sample structures are thermally mapped using the T°Imager thermoreflectance setup presented in Chapter 2. To measure the temperature rise, the device is placed on a thermally controlled stage and illuminated with a monochromatic light source. As shown in Figure 9.2, the light is directed through a 5× objective lens, reflects from the electrode surface through the transparent glass slide, and is collected on a 16-bit camera that provides $1024 \times 1024$ pixels intensity map of the reflected light from the region of interest (ROI).

![Figure 9.2](image)

Figure 9.2. a) Schematic of the thermal imaging measurement setup used for measuring the temperature rise at the electrodes. b) In calibration, the electrodes are heated using a thermal chuck. c) In activation, the electrodes are self-heated in the presence of applied voltage via Joule heating effect.
The thermoreflectance response is largely dependent on the wavelength of illumination, and thus the optimal wavelength must be determined initially to maximize the measurement signal. A wavelength (λ) scan is conducted using a Variable WL Monochromatic Light Source (Figure 9.3) while resistively heating the center electrode with a current density of 1000 µA/µm (200 mA) to induce an observable temperature change, and compare the response for the different λ values. After determining the optimal wavelength for illumination at 520 nm (530 nm LED used), the measurement procedure is performed in two runs, one for device calibration and another for activation.

![Figure 9.3. Wavelength scan showing the thermoreflectance response (ΔR/R) at different illumination wavelengths. Only the middle electrode is electrically activated using a constant current source and shows a positive reflectance change response for λ < 490 nm, a negative response for λ > 490 nm, and an optimal response at λ=520 nm. A 530 nm LED is used for higher intensity.](image)

To measure the thermoreflectance signal during device activation, the reflectance from the electrodes must be monitored. Since the micro-fluidic flow may distort the optical path above the electrodes, the measurement is performed from the back side of the electrodes through the glass substrate. A set of 50 cold frames are first obtained during the OFF state.
from the ROI. Another set of 50 hot frames are collected next after activating the device with a sinusoidal signal of prescribed amplitude ($V_{pp}$) and frequency ($f$). Since an async activation is performed, a quasi-steady state temperature rise is measured. The unit change of reflectance ($\Delta R/R$) caused by the Joule-heating temperature rise is then acquired with the camera by averaging the two sets of unit intensity change maps ($\Delta I/I$) at each pixel in the ROI. The activation run is repeated and the reflectance change maps are acquired at several voltage levels and frequencies to observe the Joule heating dependency on the different activation parameters.

Calibration is next performed to measure the thermoreflectance coefficient of the electrode material. Only the upper glass slide containing the electrodes is used for calibration and placed in direct thermal contact with the heating stage. This ensures that the electrodes are at the set chuck temperature. Moreover, this also ensures that the optical path is identical to that of the activation runs since the light reflects from the backside of the electrodes through the transparent glass (Figure 9.2b and 9.2c).

To infer the Joule heating rate within the device under AC activation, the resistance value in the frequency domain is extracted by impedance spectroscopy using the Agilent 4193A high precision impedance analyzer. An equivalent circuit analysis is conducted to extract the important impedance and resistance values of the micro-fluidic system. The values are used in the simulation model of the fluidic and thermal system. Numerical simulations of the ACET flow in our microchannel are conducted using COMSOL Multiphysics which incorporates the electrical, thermal and Stokes flow physics. The model was solved for the fluid domain representing the tested micro-fluidic device. More details on the equations, model domain, and boundary conditions are presented in [129]. The model results will be used to verify the experimental measurements.
9.4. Results

9.4.1. Frequency-Dependent Self-Heating

An initial impedance study is conducted to determine the frequency at which the heat generation rate is highest. Thermal measurements are conducted between 10 kHz and 100 MHz with 20 V\textsubscript{pp}. The results (Figure 9.4) reveal three impedance regions within the micro-fluidic system.

In the low frequency range, EDL capacitance overshadows the impedance spectra and this behavior is diminished after a critical frequency (f\approx 1 MHz). The frequency range of interest for ACET flow is located in an intermediary region ranging from 1 MHz kHz to 10 MHz, where the signal obtained is dominated by the electrolyte resistance. The impedance magnitude shows a resistive plateau and a frequency independent impedance behavior as can be noticed from Figure 9.4a. The stray capacitance takes into account the alternative current paths found in parallel with the detection or reference volume. At higher frequencies (10-100M MHz), the stray capacitance from the connectors and electronic board shunts the channel impedance.

TR measurements are conducted to determine the corresponding temperature rise at different discrete frequencies under a fixed applied potential of 20 V\textsubscript{pp}. At this activation level, the measurements are performed at discrete frequencies starting from 100 MHz to a minimum frequency of 1 kHz while avoiding lower frequencies at which electrolysis occurs. In the high frequency regime, the applied voltage is annihilated by the stray capacitance. One notable observation can be that the ACET flow generated by the temperature rise does not exist after a critical frequency (50 MHz) of the system even at high applied potentials.

In the intermediate region, the measured temperature difference remains almost constant, which indicates frequency independent electrolyte resistance. This characteristic behavior also validates that the applied voltage is transmitted to the solution without being stored in the EDL and converted into heat.
Figure 9.4. a) Impedance magnitude and b) phase angle of 1.4 S/m PBS solution in 400 µm micro-fluidic channel from 100 Hz to 100 MHz. c) Experimental temperature rise at the electrode/electrolyte interface with electrode excitation at 20 V_{pp} and frequencies varying from 10 kHz to 100 MHz. Maximum temperature rise at these conditions is obtained at 51.5 °C ± 5.5 oC (standard deviations were constant for all data points).
In the low frequency regime, most of the applied voltage is stored in EDL capacitance, and therefore, the voltage transmitted to the solution becomes less than the applied voltage. As a consequence of low transmitted voltage, the Joule heating rate and thus the temperature rise decreases drastically. At 1 kHz, the electrodes peel off because of Faradaic reactions, which occurs at low frequency spectrum and high applied voltage. Therefore, any data below 10 kHz in Figure 9.4 are lacking.

9.4.2 Voltage-Dependent Self-Heating

The voltage dependent temperature measurements are performed at 5 MHz where the resistance of the system is accurately known ($R_{sol} = 940\Omega$). Figure 9.5a shows the temperature rise profile between the side electrodes at different activation voltage ranging from 2 to 20 $V_{pp}$, with 2 $V_{pp}$ increments. The data are shown with a rainbow color coding where red and dark blue represent the highest and lowest temperature rises, respectively. As expected for Joule heating, the temperature rise increases with increasing activation voltage. Note that the temperature rise profiles obtained from the simulations (solid lines) at different applied voltages were collapsed on experimental results (symbols), showing that the temperature rise due to Joule heating is accurately predicted by the numerical model. The deviations between numerical and experimental values can be attributed to the ideal considerations adopted in the numerical modeling, where the electrode edge effects are not considered, resulting in smooth temperature profiles.

Figure 9.5b summarizes the average temperature rise on the side and middle electrodes as a function of applied voltage at 5 MHz. The standard deviation is for the set of 50 different measurements for an applied voltage and the highest standard deviation was obtained at 5.5 °C when the electrodes were activated at the highest potential of 20 $V_{pp}$. It is important to state that good agreement between the numerical model and experiments are due to the chosen applied frequency where EDL and stray capacitance effects are negligible, and hence, nearly the entire applied electrical field contributes to Joule heating. Figure 9.5c shows the thermal map obtained on the electrodes for 20 $V_{pp}$ at 5 MHz. The temperature rise between
Figure 9.5. Comparison of experimental (symbols) and numerical (solid lines) temperature fields on the electrode regions a) for different applied voltages at 5MHz. b) Experimentally and numerically obtained average temperature rise on the side and middle electrodes as a function of the applied peak to peak voltage. c) Measured temperature field on chromium-gold surface at 20 $V_{pp}$ (Top view). d) Flow field below the electrodes (represented with yellow) bars is shown with streamlines and the velocity magnitude contours given in terms of $\mu$m/s. The simulation is performed at 5 MHz and 20 $V_{pp}$. 
the electrodes (black region) could not be measured because no TR signal reflects back from
the transparent glass region nor from the fluid channel, only the reflective gold electrode
surface is measured.

As our model is validated with the electrode measurements, the temperature of the
remaining domain can be inferred from the validated numerical model. Moreover, for a given
case, the average temperature rise at the center electrode is seen to be higher than that of
the side electrodes (Figure 9.5a and 9.5b). This distinction can be explained by examining
the ACET flow behavior predicted by the numerical simulations (Figure 9.5d). Two global
vortices are created beneath the electrodes, which convect the flow from the outer edges of
the side electrodes to the center of the middle electrode, and then circulate the flow. As the
simulations show, convection effects are stronger near the side electrodes than those under
the middle electrode, which explains the additional cooling at the side electrodes compared to
the middle electrode. Another reason is that the electric field magnitude becomes maximum
between the electrodes and decreases away from inner edges of side electrodes, which can
also explain the temperature drop towards the outer edges.

Since Joule heating is proportional to the square of the electric field strength (Joule heat
\( \approx E^2(E = V/r) \)), the volumetric heat generation decreases with square of the distance and
it results in lower temperature rise at the other edges. The maximum velocity magnitude
is found consistent with the experimentally measured velocities using micro particle image
velocimetry (\( \mu \)-PIV) technique for similar electrode design [152].

9.4.3. Power-Loss Dependent Self-Heating

Applying an external electric field \( E \) to manipulate bulk fluid will result in some elec-
trical power dissipation in the form of Joule heating. Using the results from the impedance
measurements, the relationship between dissipated power and temperature rise was estab-
lished for different applied voltage. The power level is calculated as
\[
P = \frac{(V_{rms})^2}{R_{sol}}
\]
where
\[
V_{rms} = \frac{V_{pp}}{(2\sqrt{2})}
\]
The resistance of the solution (\( R_{sol}(T_0) \approx 940\Omega \)) has been measured at
5MHz-20mV at which the Joule heating effect is negligible. However, the resistance of the
solution is temperature dependent and would decrease at higher applied voltages. Since the
temperature-dependent electrical conductivity is considered for the numerical simulations,
the power dissipation is calculated and corrected for the temperature variation in the mea-
sured $R_{\text{sol}}$, using the measured temperature rise where $R_{\text{sol}}(T) = R_{\text{sol}}(T_0)/(1 + c_\sigma(T - T_0))$.
The results of the temperature measurements are plotted against the power dissipation and
presented in Figure 9.6.

![Figure 9.6. Temperature rise at the electrode/electrolyte interface as a function of the
dissipated power when the side electrodes are excited at 5 MHz and with different voltages.](image)

A linear relationship is observed between the variations of temperature and dissipated
power, with regression factors of 0.988 for both middle and side electrodes (Figure 9.6. This
relation, also known as the “Arrhenius Principle”, is expected for our results. As the power
generated increases at the electrodes, the temperature differences between the heat source
and the heat sink also increase to maintain the higher power dissipation to the substrate.
The linear relation reveals that higher order effects thermal effects can be neglected and that
the thermal resistance of the system does not vary much with activation level (468 K/W and
509 K/W for side and center electrodes respectively).
As defined in the equation presented earlier, the Joule heating rate is linearly proportional to the conductivity of the medium. As a result, reducing the conductivity of the solution should in principle reduce Joule heating, even under high levels of applied potential. To demonstrate this expected behavior, the temperature measurements were repeated for the same micro-fluidic device but while using deionized (DI) water ($\sigma = 5 \times 10^{-6} S/m$) as the fluidic medium. The obtained temperature rise data presented in Figure 9.7 show no significant Joule heating, with a temperature rise that does not exceed a couple of degrees. This also proves that the heating observed in PBS is solely due to Joule heating. The important insight from Figure 9.7 is that for low conductivity solutions, ACET operation is thermally-safe and even the highest electrical field is unable to induce a measurable temperature rise in the fluid.

![Figure 9.7](image.png)

Figure 9.7. Temperature rise as a function of applied voltages at 5 MHz for DI water showing negligible self-heating due to low Joule heating in electrically non-conductive fluidic medium.
9.5. Conclusions

This study presents a thermal investigation into ACET Joule heating induced temperature rise. The study also validates the enhanced numerical model by characterizing the thermal behavior of the electrode/electrolyte interface under ACET flow operation. The temperature rise of the activated electrodes is measured using a high-resolution, non-invasive thermoreflectance imaging method. Numerical simulations are performed for different applied voltages (2-20 $V_{pp}$) and varying frequencies (1 kHz – 100 MHz) and show substantial heating at the electrodes, with the maximum temperature rise reaching $\approx 52^\circ C$ at $20 V_{pp}$. The experimental results are consistent and thus validate the numerical solutions, both showing similar temperature rises and distributions along the middle and side electrodes. The good agreement between experimental and numerical results can allow one to assess temperature variations for ACET based devices prior to experimentation.

This study provides a useful guideline for experimental temperature determination in ACET-based micro-fluidic applications, which allows one to assess Joule heating effects. It also validates the appropriate numerical model to predict the temperature rise prior to ACET flow experiments with precious biological samples. The outcomes of this study ultimately can provide insights to improve future biosensor designs, as thermal considerations are crucial for tailoring the sensor geometry to keep the Joule heating effects within acceptable ranges for biological samples, or other temperature sensitive applications. Accordingly, our work makes substantial contributions to the field of electrokinetic biomicro/nanofluidics, by presenting a unique microscopic interpretation of temperature rise. In the future, the thermoreflectance technique can also be adapted to a variety of micro/nanofluidic applications to explore fundamental observations such as nanoscale interfacial transport at extended evaporating menisci.
Chapter 10
Conclusions

As emerging technologies push the boundaries of the achievable in activation power levels, operation frequencies, and energy efficiency, thermal concerns become more and more crucial for accurate and reliable device performance. Characterizing the thermal response and temperature rise of active devices is becoming a determining factor across the spectrum of applications and requirements, as temperature affects operation, lifetime, and reliability. In other advanced device technologies still under development, thermal response is key in understanding and predicting operating mechanisms behind devices operation to allow their modeling and integration into working systems.

This work has presented the advanced methods used for experimental and numerical thermal analysis of emerging microelectronic devices and structures. Thermoreflectance-based surface temperature mapping offers the benefits of non-invasive, non-destructive, high resolution measurements and suitable for reflective metallic electronic components, without requiring any special sample preparation. The presented ultra-fast numerical technique provides a complete 3D thermal model to accurately represent the physical test device, in regards to structure, properties, and thermal response. The combined approach finally helps bridge the gap across each of the methods’ shortcomings, resulting in an extremely powerful tool for full thermal characterization, where the experimental measurements provide accurate validation, and the simulation model provides a full temperature map that reaches embedded and inaccessible features, and provides estimates for uncertain thermal and physical parameters specific to the device-under-test at the moment of testing.

Developing an accurate digital representation of microelectronic devices serves many design benefits. The model can be used for studying the effect of different model parameters and performing parametric studies to optimize design. This saves time and cost by requiring
fewer samples for experimental validation and device testing. The method also can serve as an in-situ property measurement technique instead of other thin-film measurement setups that require expensive sample preparations. All the mentioned benefits would eventually lead to a reduction in prototyping, testing, and evaluating steps in the design cycle, with shorter time to push new, advanced, and reliable technologies to market.


Jackson, W., Jeon, Y., Henze, D., Li, Z., and Stanley Williams, R. An accurate locally 
active memristor model for S-type negative differential resistance in NbOx. 

[88] Goodson, K., and Flik, M. Effect of microscale thermal conduction on the packing limit of 
silicon-on-insulator electronic devices. IEEE Transactions on Components, Hybrids, and 


Higashiwaki, M., Jena, D., and Luo, T. Anisotropic thermal conductivity in single crystal 

[92] Han, Y., Lau, B. L., Zhang, X., Leong, Y. C., and Choo, K. F. Thermal Management of 
Hotspots with a Microjet-Based Hybrid Heat Sink for GaN-on-Si Devices. IEEE Transactions on 
Components, Packaging and Manufacturing Technology 4 (9) (sep 2014), 1441.

[93] Hatta, I. Thermal diffusivity measurements of thin films and multilayered composites. 

[94] He, L., Kahng, A., Tam, K., and Xiong, J. Design of integrated-circuit interconnects with 
accurate modeling of chemical-mechanical planarization, 2005.

[95] Heller, E., Choi, S., Dorsey, D., Vetury, R., and Graham, S. Electrical and structural 
dependence of operating temperature of AlGaN/GaN HEMTs. Microelectronics Reliability 53, 6 

[96] Heller, E., Choi, S., Dorsey, D., Vetury, R., and Graham, S. Electrical and structural 
dependence of operating temperature of AlGaN/GaN HEMTs. Microelectronics Reliability 53, 6 (jun 
2013), 872–877.

[97] Helou, A., Raad, P., and Venugopal, A. Assessment of CMP Fill Pattern Effect on the 
Thermal Performance of Interconnects in Integrated Circuits BEOL. In 2019 IEEE 69th Electronic 
Components and Technology Conference (ECTC) (may 2019), IEEE, pp. 405–409.

[98] Helou, A. E., Raad, P. E., Kande, D., and Venugopal, A. Thermal Modeling and 
Experimental Validation of Heat Sink Design for Passive Cooling of BEOL IC Structures. In 2018 
24rd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC) (sep 
2018), IEEE, pp. 1–3.

[99] Helou, A. E., Raad, P. E., Kande, D., and Venugopal, A. Thermal Modeling and 
Experimental Validation of Heat Sink Design for Passive Cooling of BEOL IC Structures. In 2018 
24rd International Workshop on Thermal Investigations of ICs and Systems (THERMINIC) (sep 
2018), pp. 1–3.
[100] Helou, A. E., Raad, P. E., and Komarov, P. Temperature dependence of the


lattice thermal conductivity of group IV and group III-V semiconductors. Physical Review B - 

[174] Moultif, N., Latry, O., Ndiaye, M., Neveu, T., Joubert, E., Moreau, C., and Goupy, 
J. F. S-band pulsed-RF operating life test on AlGaN/GaN HEMT devices for radar application. 

Thermal Boundary Conductance across Bonded Heterogeneous GaN-SiC Interfaces. ACS Applied 
Materials and Interfaces 11 (aug 2019), 33428–33434.

[176] Nandi, S. K., Nath, S. K., El-Helou, A. E., Li, S., Ratcliff, T., Uenuma, M., Raad, 
P. E., and Elliman, R. G. Electric Field- And Current-Induced Electroforming Modes in NbOx. 
ACS Applied Materials and Interfaces (2020).

[177] Nandi, S. K., Nath, S. K., El-Helou, A. E., Li, S., Liu, X., Raad, P. E., and Elliman, 
R. G. Current Localization and Redistribution as the Basis of Discontinuous Current Controlled 

[178] Nath, S. K., Nandi, S. K., El-Helou, A. E., Li, S., Ratcliff, T., Raad, P. E., 
and Elliman, R. G. Schottky-barrier-induced asymmetry in the negative-differential-resistance 

[179] Naumov, A., Kolomys, O., Romanyuk, A., Tsykaniuk, B., Strelchuk, V., Trius, M., 
Avksentyev, A., and Belyaev, A. Self-heating effects in AlGaN/GaN HEMT heterostructures: 
Electrical and optical characterization. Semiconductor Physics, Quantum Electronics & 

[180] Naumov, A. V. Self-heating effects in AlGaN/GaN HEMT heterostructures: Electrical and optical 
characterization. Semiconductor Physics Quantum Electronics and Optoelectronics 18, 4 (dec 2015), 
396–402.

and Boltasseva, A. Long-range and rapid transport of individual nano-objects by a hybrid 

and Boltasseva, A. Photothermal heating enabled by plasmonic nanostructures for electrokinetic 

[183] Nguyen, H. V., Salm, C., Wenzel, R., Mouthaan, A. J., and Kuper, F. G. Simulation and 
experimental characterization of reservoir and via layout effects on electromigration lifetime. 

Effect of self-heating on electrical characteristics of AlGaN/ GaN HEMT on Si (111) substrate. AIP 
Advances 7, 8 (2017).


