ABSTRACT

Title of dissertation:	PHYSICAL ASPECTS OF VLSI DESIGN WITH A FOCUS ON THREE-DIMENSIONAL INTEGRATED CIRCUIT APPLICATIONS
	Zeynep Dilli Doctor of Philosophy, 2007
Dissertation directed by:	Professor Neil Goldsman Dept. of Electrical and Computer Engineering

This work is on three-dimensional integration (3DI), and physical problems and aspects of VLSI design. Miniaturization and highly complex integrated systems in microelectronics have led to the 3DI development as a promising technological approach. 3DI offers numerous advantages: Size, power consumption, hybrid integration etc., with more thermal problems and physical complexity as trade-offs. We open this work by presenting the design and testing of an example 3DI system, to our knowledge the first self-powering system in a three-dimensional SOI technology. The system uses ambient optical energy harvested by a photodiode array and stored in an integrated capacitor.

An on-chip metal interconnect network, beyond its designed role, behaves as a parasitic load vulnerable to electromagnetic coupling. We have developed a spatially-dependent, transient Green's Function based method of calculating the response of an interconnect network to noise. This efficient method can model network delays and noise sensitivity, which are involved problems in both planar

and especially in 3DICs.

Three-dimensional systems are more susceptible to thermal problems, which also affect VLSI with high power densities, of complex systems and under extreme temperatures. We analytically and experimentally investigate thermal effects in ICs. We study the effects of non-uniform, non-isotropic thermal conductivity of the typically complex IC material system, with a simulator we developed including this complexity. Through our simulations, verified by experiments, we propose a method of cooling or directionally heating IC regions.

3DICs are suited for developing wireless sensor networks, commonly referred to as "smart dust." The ideal smart dust node includes RF communication circuits with on-chip passive components. We present an experimental study of on-chip inductors and transformers as integrated passives. We also demonstrate the performance improvement in 3DI with its lower capacitive loads.

3DI technology is just one example of the intense development in today's electronics, which maintains the need for educational methods to assist student recruitment into technology, to prepare students for a demanding technological landscape, and to raise societal awareness of technology. We conclude this work by presenting three electrical engineering curricula we designed and implemented, targeting these needs among others.

PHYSICAL ASPECTS OF VLSI DESIGN WITH A FOCUS ON THREE-DIMENSIONAL INTEGRATED CIRCUIT APPLICATIONS

by

Zeynep Dilli

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Professor Neil Goldsman, Chair/Advisor
Professor Martin Peckerar
Associate Professor Reza Ghodssi
Associate Professor Timothy Horiuchi
Associate Professor Linda Schmidt

© Copyright by
Zeynep Dilli
2007

Dedication

To all my teachers

Acknowledgements

First and foremost I am indebted to Dr. Neil Goldsman for his invaluable support and guidance. He has been a great advisor and teacher.

I am grateful to Dr. Martin Peckerar and Dr. Linda Schmidt for the privilege to work with them on different sections of this dissertation, and for joining my committee. I am grateful to Dr. Reza Ghodssi and Dr. Timothy Horiuchi for agreeing to serve in my committee, and for their time and effort.

I have been lucky to collaborate with brilliant and supportive people. I must mention Dr. Akın Aktürk especially. I also owe thanks to Dr. Janet Schmidt, Mr. Todd Firestone, Ms. Datta Sheth, Ms. Bo Yang, Mr. Yves Ngu and Mr. Bai Yun for their assistance and collaboration, for Mr. Jay Renner and Mr. Shyam Mehrotra for technical support, and to Dr. Volkan Cevher, Mr. Susitha Jayaratne, Dr. Gary Pennington, Mr. Siddharth Potbhare and Dr. John Rodgers for very profitable discussions. Through the years I have had many inspiring teachers, and while I cannot name them all, I thank them all.

I would like to thank Dr. George Metze at LPS and the Dept. of Electrical and Computer Engineering for supporting the majority of this work.

My thanks and gratitude go to my mother Yıldız for teaching me how to read resistors (and for much more), to my father Budak for teaching me how to analyze filters at a glance (and for much more), and to my sister Gökçe for being the ray of sunshine she is (and for much more). I give thanks to all my extended family for

their support and love. *Size gerektiği kadar teşekkür edemem hiç.*

Many precious people have supported me and offered their confidence, counsel, experience and friendship during this work. Dr. Breno Imbiriba, Dr. Hilmi Volkan Demir, Mrs. Paula and Mr. Chort Montrie, Dr. Kristy Henscheid, Dr. Jeff Huo, Mrs. Maggie Brazeau, Mr. John Novak, Mrs. Loreen and Mr. Martijn de Kort, Mrs. Frederica and Mr. Herb Baer, all those who attend the Conservatory and members of Three Left Feet, all the newsgroup people, more others than I can name and once again my family—I hope I have been and will be able to give back some of all that I received from you.

Table of Contents

List of Figures x

1 Introduction 1
- 1.1 Three-Dimensional Integration and Self-Powering: Motivation and Effects of VLSI Physics . . . 2
- 1.2 Issues in the Physics of VLSI Design: Interconnect Modeling . . . 6
- 1.3 Issues in the Physics of VLSI Design: Thermal Effects and Modeling . 9
- 1.4 Issues in the Physics of VLSI Design: Reactive Components . . . 12
- 1.5 Electrical Engineering Education: Motivation . . . 14

2 Three-Dimensional Integration: Physical Design of A Self-Contained Electronic System 16
- 2.1 Introduction and Overview . . . 16
- 2.2 Example Implementation: A Self-Powering 3-D System on SOI CMOS 17
  - 2.2.1 Design Introduction . . . 17
  - 2.2.2 Process Information . . . 19
  - 2.2.3 Photodiodes: Design Issues . . . 22
    - 2.2.3.1 Photocurrent Calculation . . . 22
    - 2.2.3.2 Photodiode Design and Layout . . . 28
  - 2.2.4 Layout and Chip Microphotographs . . . 34
    - 2.2.4.1 Tier 1: The Local Oscillator . . . 34
    - 2.2.4.2 Tier 2: The Capacitor . . . 35
    - 2.2.4.3 Tier 3: Photodiodes, Measurement Pads . . . 36
    - 2.2.4.4 Chip Microphotographs . . . 37
  - 2.2.5 Circuit Analysis and Simulations . . . 40
    - 2.2.5.1 Circuit Operation: Qualitative Description . . . 40
    - 2.2.5.2 Circuit Operation: Quantitative Analysis . . . 45
  - 2.2.6 Measurement Results . . . 49
    - 2.2.6.1 Rail Voltage Measurements . . . 49
    - 2.2.6.2 Oscillator Output Measurements . . . 49
  - 2.2.7 Concluding Remarks on the First Self-Powering SOI 3DIC . . 50
- 2.3 Second-Generation Design . . . 51
  - 2.3.1 Introduction . . . 51
  - 2.3.2 Overview of of the Second-Generation Chip . . . 52
  - 2.3.3 New Diode Arrays . . . 53
    - 2.3.3.1 Diode Layouts with No-Silicide . . . 54
    - 2.3.3.2 Diode Array with Two Serial Diodes per Branch . . 55
  - 2.3.4 Integrated Externally-Powered Amplifier . . . 57
  - 2.3.5 Self-Powered Amplifier . . . 59
- 2.4 3-D Integration: Current Research, Challenges and Directions 62
  - 2.4.1 Sequential Fabrication Techniques for 3DI . . . 63
  - 2.4.2 Parallel Fabrication Techniques for 3DI . . . 64

2.4.2.1 Chip Stacks with Peripheral Connections 64
2.4.2.2 Chip Stacks with Through-Wafer and Through-Die Connections 66
2.4.2.3 Alternate Vertical Connection Methods 68
2.4.3 An Analysis of State-of-the-Art in the Advantages and Problems of 3-D Integration 69
2.4.3.1 Gains in System Size 69
2.4.3.2 Enabling Higher System Complexity 70
2.4.3.3 Gains in System Speed and Power Consumption . . . 71
2.4.3.4 Noise and Crosstalk Problems and Proposed Solutions 72
2.4.3.5 Thermal Problems and Proposed Solutions 73
2.4.3.6 Layout Issues 74
2.4.3.7 New Process Requirements 74
2.5 Recent Progress in Self-Powering Methods 75
2.5.1 Photoelectric Methods 76
2.5.2 Piezoelectric, Vibrational and Thermoelectric Methods 76
2.5.3 Rectifying Antennas 77
2.5.4 Small Batteries 80
2.6 Summary 80

3 Physical Aspects of VLSI Systems: A Transient Spatially-Dependent Green's Function Approach to Modeling 3-D On-Chip Interconnect Networks 82
3.1 Introduction, Motivation and Background 82
3.1.1 Methodology: Linear Time-Invariant Systems and Green's Function 83
3.2 Numerical Modeling 85
3.2.1 Theory 85
3.2.1.1 An Example Implementation using SPECTRE 87
3.2.2 Computational Cost 90
3.3 Implementation 91
3.3.1 Interconnect Network Construction 91
3.3.2 Impulse-Response Solver 94
3.3.2.1 KCL Network, Equations and Discretization 94
3.3.3 Convolution Routine 96
3.4 Simulation Results 97
3.4.1 Comparison with SPECTRE 97
3.4.1.1 Convergence Test 97
3.4.2 Example 3D Network Simulations 100
3.5 Summary 106

4 Physical Aspects of VLSI Systems: On-Chip Heat Generation and Dissipation 107
4.1 Introduction and Motivation 107
4.2 Background 109
4.2.1 The Heat Equation 109

4.2.2 A Simplified Solver and the Effects of Non-Isotropic Thermal Conductivity . . . 111
4.2.3 A Thermal Solver with Individual IC Layers Represented . . . 117
4.3 Experimental Design . . . 119
4.3.1 Heaters . . . 120
4.3.2 The Temperature Sensor Array . . . 121
4.3.2.1 Temperature Dependence of the Diode Current . . . 121
4.3.2.2 Diode Array for Distributed Temperature Sensing . . 124
4.3.3 Ring Oscillators as Temperature Sensors and Heat Sources . . 125
4.3.4 The Metal Layers . . . 129
4.4 Measurements of Temperature Distribution in Integrated Circuits . . 131
4.4.1 Experimental Setup . . . 131
4.4.2 Parasitic Effects . . . 132
4.4.3 Measurement Results . . . 135
4.5 Simulations: Controlled Heating and Cooling on ICs . . . 137
4.5.1 Simulator Validation . . . 137
4.5.2 Controlled Heating and Cooling with Thermal Via Frames . . 140
4.5.3 Effects in SOI Technologies . . . 142
4.6 Summary . . . 147

5 Physical Aspects of VLSI Systems: Reactive Structure Design and Parasitic Capacitive Load Effects in Integrated Circuits 149
5.1 Introduction, Motivation and Background . . . 149
5.2 Basics of On-Chip Inductors and Transformers . . . 150
5.2.1 The Concept of Inductance . . . 150
5.2.2 Physical Design . . . 151
5.2.3 On-chip Inductor Behavior and Design Guidelines . . . 152
5.3 Experimental Design and Measurement Results . . . 158
5.3.1 Experimental Setup and Methodology . . . 158
5.3.1.1 De-Embedding . . . 159
5.3.2 Inductor Measurements . . . 160
5.3.3 Transformer Measurements . . . 165
5.4 Tunable LC Tanks: Photoelectric Effect . . . 168
5.4.1 Experimental Results . . . 169
5.4.2 Interpretation and Modeling . . . 171
5.5 Concluding Remarks on On-Chip Inductors . . . 172
5.6 A Study of Performance Improvement by Capacitive Load Reduction with 3D Integration . . . 175
5.6.1 System Design . . . 175
5.6.1.1 Measuring Bonding Pad Effects Using Ring Oscillators 175
5.6.1.2 Internal and External Ring Oscillators . . . 177
5.6.2 Measurement Results . . . 180
5.7 Summary . . . 183

6 Looking to the Future: Improving Electrical Engineering Education 184
6.1 Introduction, Motivation and Background 184
6.2 An Experiential Introduction to Electrical and Computer Engineering: A Summer Course for High-School Students 185
6.2.1 Introduction and Program Goals 185
6.2.2 The Syllabus and the Textbook 188
6.2.2.1 The Textbook 189
6.2.3 Implementation 191
6.2.4 Program Outcomes 192
6.2.5 Conclusion 196
6.3 Aiding the Mentors: A Short Course for High-School Maths Teachers 196
6.3.1 Introduction and Motivation 196
6.3.2 Implementation 197
6.3.3 Conclusion 200
6.4 Multi-Track Capstone Design: A Senior-Year Course 202
6.4.1 Introduction, Motivation and Course Goals 202
6.4.2 Implementation 204
6.4.3 Outcomes 205
6.5 Summary 207

7 Summary and Future Research 209
7.1 Three-Dimensional Integration 209
7.2 Interconnect Network Modeling 210
7.3 Thermal Modeling 212
7.4 Reactive Structures and Load Effects in ICs 214
7.5 Electrical and Computer Engineering Education 215

A Temperature Dependence of Diffusion Constants and Diffusion Lengths 217

B Temperature Dependency of MOSFET Current in Saturation 219

C Example Lab Template Excerpts from "An Experiental Introduction to Electrical and Computer Engineering" 222

Bibliography 224

List of Figures

2.1 Three tiers in a conceptual 3-D system design: The sensor (top), storage (middle) and functional electronics (bottom) levels. The components in our specific design are also shown here. 19

2.2 The particular circuit being fabricated in our 3-D design. The photodiode symbolizes a diode array comprised of many diodes in parallel. 20

2.3 The full integration structure for the 3DL1 process. In our design, the photodiodes will be built in the active regions of the top tier, Tier 3. 21

2.4 The schematic representation of a pn-junction diode fabricated in this SOI process and of the photosensitive area to be taken into account. Since the diode tier is integrated into the 3DIC upside-down with respect to its fabrication orientation, the metal connections to the anode and cathode of the diode are shown as going "downwards". In this case, the effective photosensitive area is $A = W_d W_j$, and the light passes through the active silicon depth d for a single pass. In this picture, the following are not shown for clarity: The oxide layers surrounding the active silicon layer which houses the diode, and the polysilicon layer covering the junction, which is between the metal layers and the active layer in the process. 24

2.5 The multiple reflections in our multi-layer thin film system. The layer thicknesses are not shown to scale. The transmission layers are numbered 1 through 6 and interfaces named A through F. Below the SiO_2 layer 6 a metal layer is shown, with 100% reflection assumed at that interface. Every time an incident ray hits an interface, a green arrow points at it splitting into transmitted and reflected components. With the incident power level set at 1.0, the fractional numbers show the percentage reflected or transmitted at each interface. The numbers shown in red are thus power levels entering and re-entering the active silicon layer. When the power level in a particular ray trace falls below 0.1 (i.e. 10%), that number is given in gray and that particular ray is not traced any further. Note that through multiple reflections layer number 6, eventually all of the power entering layer 6 from layer 5, 33.4% of the incident power, will be reflected back into layer 5; all the back and forth of that process is not shown in the figure. 27

2.6 Implants for the pin-type diode. 30

2.7 Implants plus poly for the pin-type diode. 30

2.8 The full layout for the pin-type diode. 30

2.9 Implants for the CBN/CBP diode. 33

2.10 Implants plus poly for the CBN/CBP diode. 33

2.11 The full layout for the CBN/CBP diode. 33

2.12 Local oscillator (Tier 1—bottom tier) layout. To the left and right are the 3-D vias coming from the higher tiers, for GND and VDD rails respectively. 34

2.13 A zoom of the local oscillator, output buffers and the 3D via carrying the output signal up. 34

2.14 The capacitor tier (Tier 2—middle tier) layout. The capacitor top plate is a poly square, 67.9 μm on one side. The bottom plate is produced using the CAPN n-type implant. The expected capacitance of this structure is 30 pF and the Cadence-extracted capacitance is 29 pF. 35

2.15 The diode tier (Tier 3—top tier) layout. The bondpads have overglass cuts to allow access to the Metal 1 layer of the tier with probes or bondwires. Most of the tier is covered by the array of 2062 annular CBN/CBP diodes, and there are also 52 lateral "pin" diodes near the top of the layout. All features are identified in the chip microphotograph, Figure 2.16. The layout is a square 250 μm to a side. 36

2.16 The fabricated 3DIC microphotograph. The top tier is visible. Top right: VDD pad. Top left: GND pad. Bottom middle: The oscillator output pad. The pin-diode array is next to the VDD pad. The rest of the tier is covered by the annular pn-diode array. The intertier (3-D) via arrays to the lower tiers from the VDD and GND pads are below the pads; the intertier via from the lower tier to the output pad is to its right side. Lower tiers are somewhat visible, though out of focus, at the bonding pad edges, where there are no top tier features defined. 37

2.17 A close-up of the intertier via landing next to the GND pad and the corner of the surrounding annular pn-diode array. The four dark squares at the center of the each diode are the stacked contacts and vias from the n-active region to metal-1 and metal-2. The eight squares on the sides of each diode are contacts from the p-active region to metal-1. The polysilicon square covering the junction for silicide protection is also visible in each diode. 38

2.18 A close-up of the bonded die. Surrounding our design are test structures from the process. 39

2.19 System operation with $i_p = 15$ nA. The voltage scale is between -50 mV and 300 mV, and the blue line rising faster in the beginning is the rail voltage, which reaches a steady state of 242 mV in this simulation. The lighter solid line is the ring oscillator output feeding the buffer and the thick solid line is the buffer output to a 15 fF load (bonding pad only). 41

2.20 Zoom in on the full oscillation region of the simulation in Figure 2.19. The oscillation frequency in this case is ∼1.82 MHz. 42

2.21 Zoom in to Figure 2.19 around the region where oscillation is beginning. 42

2.22 Simulated steady-state rail voltages (top) and oscillation frequencies (bottom) vs. photocurrent. 43

2.23 System operation with $i_p = 40$ nA and storage capacitor 30 pF. . . . 43

2.24 The PMOS source current of the fourth inverter (first buffer stage) during full oscillation. The ratio of the cross-hatched area to the highlighted area gives Φ_4, which is interpreted as the effective duty cycle. 47

2.25 **a)** Simplified equivalent circuit to demonstrate the system operation. **b)** The rail voltage of the circuit builds up in time to reach a steady state as the current drawn by the inverters from the storage capacitor, rising with increasing rail voltage, reaches an equilibrium with the supplied photocurrent. The solid line with the triangle markers display the solution to the differential equation Eqn. 2.19, which governs this process. The dashed line reproduces the Spectre simulation of this process from Fig. 2.19. 48

2.26 Amplifier output with 3DIC under red laser pointer light. The maximum signal amplitude is 130 mV, and dominant frequency is 1.8 MHz. The bottom signal level is at ∼39 mV. 50

2.27 An at-a-glance floorplan of the 3DM2 chip. 53

2.28 The no-silicide version of the CBN/CBP annular diode. The dimensions are the same as the diode in Figure 2.11. 55

2.29 The layout for one branch of the two-serial-diodes-per-branch photodiode array. The overlay shows how the diodes are connected by the metal-2 bridge in the center. 56

2.30 The layout for a quartet of branches for the two-serial-diodes-per-branch photodiode array. The layout was constructed by creating a mosaic out of these quartets. 56

2.31 The externally-powered amplifier used in the third quarter of the 3DM2 chip. 57

2.32 Top: The self-powered oscillator and buffer driving a 1 pF off-chip load. Bottom, first graph: The output of the self-powered oscillator and buffer driving the externally-powered built-in amplifier. Bottom, second graph: The output of the externally-powered built-in amplifier, with the above input, driving a 1pF off-chip load. 58

2.33 Layout of the externally-powered amplifier used in the third quarter of the 3DM2 chip. 59

2.34 Schematic of the self-powered amplifier implemented in the fourth quarter of the 3DM2 chip. The amplifier is a simple common-source amplifier, biased by a current mirror which is fed by the photodiode arrays. 60

2.35 Top: The full layout for the self-powered amplifier, and the externally-powered level shifter and output amplifier to drive off-chip load. Full resistor layouts are once more cropped out for visual clarity. Left: a schematic representation of the components in the above layout. . . . 61

2.36 A 3-D inverter formed by fabricating a PMOS over an NMOS using epitaxial lateral overgrowth (ELO) techniques (adapted from [2]). The n layer which forms the channel of the PMOS is grown using ELO. The poly gate is common to both transistors. 63

2.37 **a)** A schematic representation of a stack of integrated circuits connected to each other through interconnects extended to the edges and routed on the sidewall. **b)** A cross-section of a stack of two chips connected vertically through a T-connection formed on the sidewall, with outside connections enabled by a solder bump, adapted from [64]. **c)** The schematic representation of the sidewall connections for a stack of VLSI associative string processor chips, from [62]. 65

2.38 A schematic representation of wafer-level 3DI. The substrates of the upper tiers are thinned, but not removed entirely, and the process in question is a bulk process. The bottom tier has the active side facing up, while the higher tiers have the active devices facing down. (Adapted from [66].) 67

2.39 Using the same maximum interconnect linear length, a central unit in a three-dimensional configuration can access many more neighboring units, depending on the relative sizes of the central unit vs. layer-to-layer distance in a stack. 70

2.40 Photos of the rectifying antenna system: Left, the full board with an LED load; middle: close-up of the back side of the board with the Schottky diode. Right: Schematic of the circuit. 79

2.41 **a)** The rectifying antenna of Fig. 2.40 charging a 2.2 mF capacitor over twenty seconds, with $\sim$ 5mW of RF power provided at $\sim$467 MHz. **b)** The voltage drop across an LED load being suppled by the rectifying antenna as the RF power is turned on and off. 79

3.1 Unit cell examples for a two-metal process. 84

3.2 Left: A continuous time-and-space dependent function. Right: A discretized time-and-space dependent function, which can be written as the sum of its samples using the coefficients α_{ij}. 86

3.3 Time-dependent responses over all space to two individual impulse-type inputs applied at different x-points at different times t_j. 86

3.4 A linear network with six nodes defined. 88

3.5 Top: Simulated impulse responses at node x_3 to impulses at x_2 and x_5. Bottom: Simulated output at node x_3 to the input of Eqn. 3.7. . 89

3.6 Response at node x_3 calculated using Eqn. 3.8 (black triangles) together with the simulated output (red circles), exact agreement. . . . 89

3.7 Two-metal network; a simplified equivalent RC network shown. Possible unit cells partitioned with dashed outlines. Note that where two metal lines on different layers overlap with a via, the connection between the mesh points on the overlap is resistive, whereas at locations without a via the same connection is capacitive. 92

3.8 The element values that need to be defined per mesh point by the mesh creator code. 93

3.9 A 5x5x3 mesh full response simulation. Top: Our solver combines impulse responses at three points using the input data. Bottom: SPECTRE performs full mesh solution at each timestep for the same input pattern. The solutions match. 98

3.10 The time evolution of the Green's Function of a 21x21x5 system on the topmost leayer (layer 5). The unit impulse was injected at t=0 into point (1,1,1): The lower left corner of the lowest level, layer 1. The uneven spread is due to the lower resistivity of the network in the N-S direction. 102

3.11 3-metal interconnect network. Vias: **X**. Inputs: $\sqcup$. Outputs: $\bullet$. 103

3.12 Responses of the network in Fig. 3.11 to a combined input signal injected at points (5,15,1) and (25,25,3). Distance as well as the effect of inter-level coupling and presence of paths to ground affect the responses. 103

3.13 Responses of the network in Fig. 3.11 over time to an impulse injected at point (5,15,1). Left: Response at Level 1 (Metal 1 network) over 300 femtoseconds. Right: Response at Level 2 (Metal 2 network) over 300 femtoseconds. 104

3.14 Responses of the network in Fig. 3.11 over time to an impulse injected at point (25,25,3). Left: Response at Level 2 (Metal 2 network) over 24 femtoseconds. Right: Response at Level 3 (Metal 3 network) over 24 femtoseconds. 105

4.1 An example mesh for a 2-D thermal network. The temperature is measured at nodes, numbered with the red (italic) numbers. The thermal conductivity is the property of the material on the branches, numbered with the blue (sans serif) numbers, between the nodes. . . 112

4.2 Left: Resulting temperature distribution with a uniform κ distribution. Right: Resulting temperature distribution with the same uniform κ distribution, with the southern boundary three times as thermally conductive to ambient than the other boundaries. 114

4.3 A non-uniform thermal conductivity distribution. The light-colored areas are five times more thermally conductive than the dark areas. . 115

4.4 Resulting temperature distribution with the nonuniform κ distribution given in Fig. 4.3. 115

4.5 A directionally non-uniform thermal conductivity distribution. The left-side and right-side shows the κ distribution of the E-W branches and N-S branches respectively. The gray (orange) areas are three times, the light-colored areas are five times more thermally conductive than the dark areas. 116

4.6 Resulting temperature distribution with the nonuniform κ distribution given in Fig. 4.5. 116

4.7 *Left*: The A-type heater/temperature array sensor chip, microphotograph. *Right*: The B-type heater/temperature array sensor chip, layout. 119

4.8 An example polysilicon microheater resistor. The poly layout shape is designed to surround temperature-sensor diodes, which are shown faintly within the microheater layout, with one example layout to the right. 120

4.9 Schematic representation of five columns and five rows of the diode array. For reference purposes, for instance, the diode on the top left corner of this schematic would have the label **d08p01n**. 125

4.10 IV curves of two diodes on the Type A and Type B chips. The dashed and dot-dashed lines with circle and square markers are measurements taken at room temperature. The solid lines with triangle markers are measurements taken with the heater over these diodes turned on. . . 126

4.11 SPECTRE simulation results for 31-stage minimum-size ring oscillator oscillation frequency vs. rail voltage for T=300 K, 307 K and 314 K, using the AMI 1.5 μm technology. 129

4.12 A Type-B measurement chip mounted on the experimental PCB. . . 132

4.13 The circuit for diode IV curve measurement. 133

4.14 **a)** The formation of parasitic BJTs in the diode array. **b)** The simplified equivalent circuit of a small (3x2) diode array, with the parasitic BJT and n-well/substrate diodes shown. The red (thick) wire indicates the shorting of the n-side of the diodes of the row under measurement, Row 2 in this example, to the p-substrate. 134

4.15 Measurement results from the Type-A chip with the uniform metal network. 135

4.16 Measurement results from the Type-B chip with the asymmetric metal network. 135

4.17 Measured ring oscillator frequency vs. rail voltage from chip Type A, when one (R24) or two (R24+R34) of the heaters next to the oscillator are operating. 136

4.18 Simulation results for the Type-A chip (uniform metal network), displaying the temperature distribution at the substrate level, matching Fig. 4.15. 139

4.19 Simulation results for the Type-B chip (asymmetric metal network), displaying the temperature distribution at the substrate level, matching Fig. 4.16. 139

4.20 Simulation results for the Type-A chip with a vertical thermal via frame consisting of metal poles from the topmost layer all the way down to the substrate placed around the operating heater. 141

4.21 Simulation results for the Type-A chip, with the same vertical thermal via frame placed around the heater, operating at 0.46 W to obtain a comparable temperature range as the original measurement. 141

4.22 Simulation results for the Type-A chip substrate temperature, with an unbroken vertical thermal via frame from the topmost layer all the way down to the substrate placed around the operating heater. . . . 142

4.23 Simulation results for the Type-A chip polysilicon layer temperature with the same frame as Fig. 4.22. Note that the axis of this figure are zoomed in from the full-chip scale. 142

4.24 The temperature distribution when a heater outside of and to the east of the via frame is operating. The via frame, consisting of poles, is marked with a dashed white outline. 143

4.25 Simulation results for the substrate temperature in an SOI version of the Type-A chip (uniform metal network), with the same heater operating as the results in Fig. 4.18. 145

4.26 Simulation results for the substrate temperature in an SOI version of the Type-B chip (asymmetric metal network), with the same heater operating as the results in Fig. 4.19. 145

4.27 The lighter-colored region is the area on which there are the extra metal interconnects in Chip Type B, also visible in the chip photograph. It is clear from the contour lines that the thermal gradients change at the boundaries of this region. 145

4.28 The SOI version of the Type-A chip simulated with a vertical via frame consisting of poles connecting the substrate to the ambient. The peak substrate temperature for the area under the substrate falls to 572.59 K and the areas outside the via frame experience no temperature rise. 146

5.1 Top and side views of typical square spiral on-chip inductors. 151

5.2 Lumped-element models for a single microstrip and coupled microstrips on a lossy substrate, which can represent the metal-oxide-semiconductor system of on-chip inductors. 154

5.3 The measured inductance of a typical on-chip spiral inductor vs. frequency, as defined by Eqn. 5.2. 155

5.4 A very basic lumped circuit model which gives the same sort of impedance response with the same three regions of an on-chip inductor response. 156

5.5 Changes in the elements of the simple lumped circuit of Fig. 5.4 reflected by changes in its inductance curve. 156

5.6 An on-chip inductor within a probe-pad structure designed for two-port S-parameter measurements. 158

5.7 Open and thru de-embedding structures for on-chip measurements. . 159

5.8 De-embedding. 160

5.9 Results of the de-embedding process. 161

5.10 Inductance and quality factor plots for inductors built on p-substrate using different metal layers. (M3 (“m3psub”): Metal 3, M2 (“m2psub”): Metal 2 , M1 (“m1psub”): Metal 1.) 162

5.11 Layouts for planar and stacked inductors; schematic representation of the latter. 162

5.12 Inductances and quality factors of planar and stacked inductors. (M3, M1: Planar Metal 3 and Metal 1.) 163

5.13 Schematic representations of the “Coil1,” “coil2” and “coil3”. 164

5.14 Layouts of the coil inductor variants coil2 and coil3. 164

5.15 Inductances and quality factors for planar and coiled inductors. (M3: Planar Metal 3.) 164

5.16 Different transformers: M2:M3 (metal 2 to metal 3), M2:M1M3 (metal 2 to metal 1+metal 3) , interwound, spiral-within-spiral. 165

5.17 Layouts for the M2:M3 and spiral-within-spiral transformers. 166

5.18 Reflection and transmission coefficients, for the M2:M3 (“m2m3xf,” metal 2 to metal 3, blue) and M2:M1M3 (“m2m1m3xf,” metal 2 to metal 1+ metal 3, red) transformers. 167

5.19 “Turn ratio” (left) and voltage gain (right) for the M2:M3 (“m2m3xf,” metal 2 to metal 3, blue) and M2:M1M3 (“m2m1m3xf,” metal 2 to metal 1+ metal 3, red) transformers. 167

5.20 Reflection and transmission coefficients for the interwound (interxf, green) and spiral-within-spiral (swsxf, black) transformers. 168

5.21 “Turn ratio” (left) and voltage gain (right) for the interwound (interxf, green) and spiral-within-spiral (swsxf, black) transformers. . . . 168

5.22 The measured inductances of two metal-3 on-chip spiral inductors, one on p-substrate and the other on an n-well. Solid lines: External light off. Dashed lines: Light on. This is data prior to de-embedding. 170

5.23 Zoom near the f_{sr} point of Figure 5.22. 170

5.24 Zoom near the peak impedance point of Figure 5.22. 170

5.25 A more advanced lumped circuit model for on-chip inductors [131]. . 171

5.26 Spectre s-parameter simulation results for the lumped-circuit inductor model of Fig. 5.25 with the element values in the text. Solid line: ”Light off”. Dashed line: ”Light on,” with the substrate resistance reduced. 173

5.27 Zoom near the f_{sr} points of Figure 5.26. 173

5.28 A five-stage ring oscillator. 176

5.29 The CMOS inverter and its digital model. The input and output capacitances are proportional to the transistors’ gate oxide capacitances: $C_{in} = 1.5C_{ox}$ and $C_{out} = C_{ox}$ approximately. The effective switching resistances are the average resistances of the relevant MOSFETs between the off and the on stages of the full swing: $R_{n,p} = V_{DD}/I_{Dn,p}$. 177

5.30 A five-stage ring oscillator with a capacitive load between two of the stages. 178

5.31 The 31-stage ring oscillator comprised of minimum-size transistors in the 0.6 μm technology chip. The output is taken from between the stages on the lower and upper right hand corners and goes through an additional inverter, which serves as a buffer (not shown). 179

5.32 The layout for the 31-stage ring oscillator and 6-bit counter/divide-by-64 frequency divider. 179

5.33 Three stages of the external ring oscillator. The bonding pad connections are visible on the top and left edges of the figure. 180

5.34 The full chip layout for the C5N chip. 180

5.35 The output waveform of the 11-stage external ring oscillator. 181

5.36 The output waveform of the divide-by-64 counter obtaining its input from the 31-stage internal ring oscillator. 181

6.1 Example exam questions. These were the associated question statements: *(Left circuit)* For this op-amp circuit, what are I_o and V_o? Is the LED on, why or why not? *(Right circuit)* For the BJT circuit β=100 and V_{BE}=0.7. Draw the DC equivalent circuit. Calculate I_B, I_E, I_C, V_B, V_E and V_C. 195

6.2 Example digital-to-analog conversion question. The sequence: **011 101 110 111 110 101 011 001 000 000 000 001 011 101 110 111 110 101 011**. 200

6.3 Example grayscale image converted into a 8-level 32x32 grayscale array. Top left: Original image. Top right: Empty grid and grayscale levels provided. Bottom left: Grayscale values given. Bottom right: Reconstructed image. 201

C.1 Lab report template excerpt example for an experiment studying the voltage drop across a diode under different biases. 222

C.2 Lab report template excerpt example for an experiment studying the behavior of a half-wave rectifier circuit. 223

Chapter 1

Introduction

Current microelectronics research displays two interlinked trends: Miniaturization and increasing system complexity. Innovation and economics push the drive for miniaturization, which is most emphatically visible in the evolution of CMOS technology. The development of increasingly complex applications is encouraged by developments in signal processing, communications and computing. Novel materials and technologies, attempts to use disparate technologies together, and increased flexibility in semiconductor processes are aspects of both of these trends.

The emergent effect on the integrated circuit design process is that certain physics-based issues gain increased significance. For instance, silicon-on-insulator (SOI) technologies and three-dimensional integration (3DI) ease certain power consumption and scaling problems, but they require more careful thermal modeling. Decreasing feature size and processing techniques allow on-chip passive elements for radio-frequency applications and higher operating frequencies, but this implies that the interconnect network, now itself more complex, also needs more detailed modeling. Both these thermal and electromagnetic modeling problems have to be considered along with the electronics for a holistic approach to integrated circuit design and design techniques.

In this dissertation, we present our approaches to the different facets of de-

sign, characterization and physical modeling involved in integrated circuit research. This introductory chapter overviews the research areas we will address. Chapter 2 opens the main body of the dissertation with a focus on the nascent technology of three-dimensional integration (3DI), detailing the design and implementation of a self-powering three-dimensional system and continues with further review of 3-D and self-powering technologies. Chapter 3 follows with a novel approach to modeling the response of on-chip interconnect networks to internal and external noise. Chapter 4 is concerned with thermal modeling, where we present our experimental and theoretical work on on-chip heat generation, dissipation, cooling and controlled heating. Chapter 5 is a look at the design and fabrication of on-chip reactive structures required for RF circuits as well as the extra power and speed cost of parasitics created by intra-chip connections. Finally, Chapter 6 provides a look at our activities in engineering education, with the emphasis that the true driving force behind the innovation in electronics applications is an ever-deepening relationship of interdependency between engineering and society.

1.1 Three-Dimensional Integration and Self-Powering: Motivation and Effects of VLSI Physics

The decreasing feature size and increasing single-unit complexity inclinations in integrated microelectronics are driven by considerations of speed, compactness and robustness, at every level of electronics design from device-level to system-level. On the way, the problem of moving signals between parts of the system

implemented on different semiconductor sections arises. Such signals need to go through chip and package structures which can contribute significant amounts of parasitic capacitive and inductive load. They typically require dedicated buffer structures at the output stages of each chip, which increase power consumption, cause extra delay and unwanted extra heating, and waste on-chip real estate.

It is desirable to have a method of connecting different dies without having to go through bonding pads and off-chip traces while conserving chip space. Chip stacking, or three-dimensional integration (3DI) , emerges as a possible answer [1, 2, 3]. The basic idea of 3DI is to place individual substrates, each housing the circuitry of a subsystem, above each other to be connected by vertical interconnects and form the full system. This new approach to integrated circuit design and fabrication brings a number of advantages and certain new complications to the table, all traceable to different physical—geometrical, electrical or thermal—aspects of VLSI technology.

The third dimension enables tighter integration of systems of greater complexity in a single unit by allowing more potential connections as compared to a planar system with the same number of metal layers [3, 8]. 3DI also facilitates hybrid and mixed-signal integration: It lets the digital and analog parts of a mixed-signal circuit, implemented on separate substrates, to be closely integrated as the tiers of a stacked system, yet remain less vulnerable to the significant substrate noise coupling problem of such planar systems [4, 5]. Also, the creation of multimaterial systems might become easier with the additional degree of freedom, as problems like lattice mismatch or material incompatibility are circumvented [6, 7].

Finally, bringing together planar subsystems with vertical connections instead

of some form of off-chip connections allows for shorter electrical connections between them. In a 3-D integrated system, fewer signal and power lines need to go off-chip. This implies less load on driver circuits and thus lower power consumption, lower parasitics and noise, and higher operation frequencies [3, 8, 9, 10].

We can summarize the advantages of 3-D integration as follows:

- Tighter and denser integration and system size reduction, including lower weight considering packaging;
- Delay reduction, making faster clock speeds and higher signal bandwidths possible, through
  - Shorter interconnects between tiers than between planar-integrated sub-circuits,
  - Less parasitic impedance and load impedance;
- Higher node accessibility and connectivity;
- Natural suitability for parallel processing applications;
- Potential noise reduction in the transmissions between different parts of a system;
- Potential substrate noise reduction, aiding mixed-signal integration;
- The possibility to integrate components with substrates of different materials, enabling hybrid integration and multi-functional single ICs;

- Higher degree of geometric freedom, especially for the design of geometry-dependent structures like inductors, transformers and antennas;
- Power consumption reduction.

On the other hand, the main disadvantages of this technique are

- Potential increase in heat-dissipation problems;
- Inter-layer crosstalk potential, both DC and high-frequency coupling,
- Increased (geometric, computational, routing) design complexity;
- Increased process complexity and cost.

Self-powering and power-harvesting are research topics that naturally link to three-dimensional integration in the development of a new technology of considerable interest. Commonly called "smart dust systems", these are conceived to be a network of many very small circuits deployed over a certain geographical area and in communication with the other members of the system. Several features of 3-D integration, listed above, help with this technological trend. Each unit in a smart dust system is ideally miniaturized and self-contained. A fully-self-contained unit would ideally be self-powering as well.

In another vein, critical considerations of energy conservation and the search for alternative energy sources are reflected in the research into self-powering electronics. Miniaturized or not, many applications intended to operate for long times autonomously, without human supervision, can use self-powering as a design feature. We can also consider remote-powering technologies as requiring less mainte-

nance during operation, although these systems might not be considered strictly self-powering as they harvest ambient energy that has been provided intentionally. Efficiency, cost and ease of integration are some of the issues facing the designer of a self- or remote-powered system.

In Chapter 2 we present the design, analysis and testing of a proof-of-concept implementation of a self-powering low-power circuit implemented in three-dimensional silicon-on-insulator technology. The chapter also includes overviews of the state-of-the-art research in 3-D integration and self-powering, with case studies.

1.2 Issues in the Physics of VLSI Design: Interconnect Modeling

The interconnect network on an integrated circuit is comprised of metal lines, typically organized in layers stacked over the semiconducting substrate with isolating layers of dielectric between them. Typical fabrication processes currently feature between two to eight layers of metallization [11]. The metal layers are often named with a numerical scheme (**Metal 1, Metal 2,...**) counting up from the layer closest to the substrate [12]. Metal layer thicknesses and the thicknesses of the isolating dielectric layers (sometimes termed ILD, "Inter-Layer Dielectric", and typically silicon dioxide for silicon processes) vary between hundreds of nanometers to one or two micrometers. Layout in a metal layer needs to be of a minimum allowed width, depending on the process, and two metal lines on the same layer have a minimum separation to prevent fabrication errors and accidental shorts. Vertical vias are used to connect metal layers to each other. They are formed by etching (typically) square

holes in the oxide layer covering a metal layer and filling it with metal before the next layer of metallization is done. Finally, similar connection structures perforating the oxide layer directly on top of the silicon substrate are called "contact"s connect the semiconductor devices to the metallization. Common materials used for metallization and via/contact formation are aluminum, gold and tungsten.

Compared to the intrinsic resistivity of the available silicon and even polysilicon layers, the metal layers exhibit much lower resistivities. Thus they are not used to make on-chip resistors, low-doped regions or polysilicon layers being used for this purpose instead. On the other hand, since the oxide dielectric constant and its thickness between two layers are well-known, metal layers are convenient for on-chip capacitors for a certain capacitance range. Finally, although their low-but-finite resistivity does cause a non-ideal quality factor, the metal layers are the natural available material on-chip for inductor fabrication. But the main functions of the metal layers is to connect single devices to make circuits, to form an interconnect network to carry signals between and power to on-chip circuits, and to become pads which enable off-chip connections.

The complexity, role and effect of the interconnect networks found on VLSI circuits have continuously been evolving [13]. Both in digital and analog circuit design, larger systems with more functionality placed in a single die tend to complicate the interconnect network; likewise the rising number of available metal layers of newer processes. The interconnect network is a physical system: Its behaves like a load with parasitic impedance beyond its design purpose and it can be vulnerable to coupling between its components and to outside electromagnetic signals. Design-

ers need to be aware of and able to model these aspects of interconnect networks. Furthermore, new design approaches have been proposed which intentionally make use of the physical impedance characteristics of interconnects [14]; such approaches require even more precise modeling of the underlying physics. And the interconnect networks of 3DICs will present increased complexity by providing inter-tier connections, as well as unique problems such as the risk of parasitic coupling to not one but two substrates on either side or to the substrate a vertical via passes through.

Interconnect network modeling is a non-trivial problem from the following angles:

- An accurate physical model of single, as well as coupled, interconnects on semiconductor substrates is required. This is not a new research area, with influential works dating back to the 1970s [15, 16]. However, it is still active, as the system of multiple metal interconnects on a conducting substrate is intricate and not straightforward to model with full electromagnetic simulation methods. There is always a demand for accurate models convenient for use in circuit simulators [17]—[21].
- With rising operation frequencies, in digital and analog circuits both, and increasing circuit sizes, several aspects of the entire interconnect network become harder to model rapidly and adaptably [13, 21], while the need to do so emerges:
  - Modeling variable delays from and to different points in the network accurately.

– Modeling the sensitivity of the interconnect network to external noise or intentional interference such as electromagnetic pulses.

– Modeling the sensitivity of the interconnect network to internal noise, such as might originate from a "dirty" power line or from a surge in one of the signal lines.

In Chapter 3 of this dissertation, we present a spatially-dependent transient Green's Function based approach to investigating the response of an interconnect network to internal and external noise. We take advantage of the fact that the interconnect network can be modeled as a linear time-invariant network.

1.3 Issues in the Physics of VLSI Design: Thermal Effects and Modeling

Operating electrical circuits generate heat predominantly through *Joule Heating* or *ohmic heating* [22], calculated by

$$H(\vec{x},t) = J(\vec{\vec{x}},t) \cdot \vec{\nabla}\phi(\vec{x},t) \tag{1.1}$$

where J is current density and ϕ is the electric potential. H is the generated heat density, given in W/m^3. Within an integrated circuit, heat is generated anywhere there is current flow and a finite resistivity, from interconnects to single switching devices. Heat generated within an integrated circuit leads to a temperature distribution throughout the system, governed by the heat conduction equation, [23]:

$$\vec{\nabla} \cdot (\kappa \vec{\nabla} T(\vec{x},t)) + H(\vec{x},t) = \rho C_p \frac{\partial T(\vec{x},t)}{\partial t} \tag{1.2}$$

where T is the temperature, κ is thermal conductivity, ρ is the material density and C_p is its specific heat. The term $-\kappa\vec{\nabla}T$ is described by the Fourier Law as the heat flux vector.

As device sizes shrink and device density as well as operating frequencies increase, heating, caused by the designed circuit operation as well as undesirable processes such as leakage current, becomes highly problematic [105, 106]. It is critical to model the operating temperature of on-chip devices accurately and self-consistently with the electrical operation, since many electrical parameters such as mobility, thermal velocity and the intrinsic carrier concentration depend on local temperature, which will in turn depend on the power output of the devices in question. In the case the heat generated can not be dissipated as rapidly, digital circuit performance suffers due to issues such as timing problems. Rising temperature causes increasing leakage current, which leads to even higher power consumption and further heating [24, 25]. Dissipated power in microprocessors currently in use can reach 130 W and the corresponding power density up to hundreds of kiloWatts per m^2; peak temperatures on the integrated circuits can rise to 120 °K above the ambient; methods used for the requisite cooling, such as packaging design and software-based thermal management, are costly in terms of price and performance [26, 27, 106]. Analog circuits are vulnerable to thermal problems such as thermal memory effects in amplifiers and the positive-feedback effect in bipolar-based RF circuits which operate at high currents to improve high-frequency performance [28, 29].

Thermal effects on microelectronics need to be investigated from another angle for circuits required to operate at cryogenic temperatures. For instance, space

electronics are subjected to very low ambient temperatures, while some infrared detectors are cooled during operation to suppress noise levels. Device behavior naturally changes in cold-temperature operation as well [30, 31], while self-heating effects may become more prominent.

Thus, it is desirable to be able to include heating and cooling effects in the integrated circuit design flow. Furthermore, new techniques such as SOI and 3-D integration often require even more sensitive thermal modeling, as they may improve electrical circuit performance at the cost of exacerbated thermal problems.

To model on-chip heat generation and dissipation, the heat conduction equation can be discretized and solved, in a coupled system with device equations for self-consistent thermal analysis [32]. Discretization of Eqn. 1.2 in the integrated circuit mesh yields a system of equations equivalent to the Kirchoff's Current Law (KCL) equations obtained from an RC network. To account for the varying thermal characteristics of the different materials in an IC, it is possible to use equivalent values obtained by taking into account the conductivities of the materials involved and their particular geometries. However, for several reasons it is worth considering to model the features built of the semiconductor, oxide and metal layers individually:

- The most significant heat sources are often considered to be the devices within the semiconductor layer. However, structures like resistors and inductors can be manufactured on other layers and generate heat.

- Similarly, it has recently been suggested that especially in signal-dense applications such as microprocessors, interconnects are responsible for a significant

fraction of heat generation [33].

- The oxide layer is a thermal as well as an electrical insulator. Heat flux is relatively isotropic in the semiconducting substrate, which is also responsible for the majority of heat dissipating out of integrated circuits, but depending on the layout, it can be directional in the metal layers woven in the oxide. It has been proposed, especially for SOI and 3-D technologies where direct thermal connection to the substrate is restricted, to use metal interconnects and vertical vias to help cool the chip [55].

In Chapter 4, we first review previous work done in our research group to model thermal behavior in integrated circuits. We then present an experimental method to measure on-chip thermal dissipation, and use our results to verify previous simulation work. We then expand the simulator to model different layers in the integrated circuit system individually and comment on the effect of the metal interconnect network. We conclude with suggestions to use the metal layers for localized heating and cooling.

1.4 Issues in the Physics of VLSI Design: Reactive Components

With rising operation frequencies in analog and mixed-signal circuits, placing passive elements such as inductors and transformers on-chip has become practical [34]. Tuned amplifiers, mixers, oscillators and impedance-matching networks are examples of circuits where inductors are required, while transformers are used for current or voltage boosting as well as impedance conversion. Thus there has been

a very active research effort on the design, implementation and modeling of these devices over the 1990s on to today [35, 36, 37]. The on-chip integration of such passive devices is also a part of the trend towards system-on-a-chip implementations, paralleling the development of 3DICs.

On-chip inductors are usually manufactured using the metal layers of CMOS processes. The basic typical design mimics the macro-size coil inductors by laying out a spiral structure, which can be square, hexagonal, octagonal or as nearly circular as the Manhattan-type layout allows. In order to reduce capacitive coupling with the substrate, the highest-level metal is preferred for single-layer spirals. Some newer RF-oriented processes provide a thick metal layer on the top level specifically for inductor layout, as a higher quality factor is achieved with lower parasitic resistance arising from the inductor metal [38, 55]. It is possible to use the multiple metal layers available in semiconductor processes to design multi-layer inductors and transformers [39]. Such inductors exhibit higher inductance per area, with lower quality factors.

By the limitations of their geometry and environment, on-chip inductors are unavoidably LC tanks, with their "inductance" exhibiting inductive, self-resonant and capacitive regions as we move from low to high frequencies. Therefore it is possible to intend their inclusion in a circuit in the role of a passive-LC tank [40]. The resonance point of these devices are set by the design, but there has been efforts to tune them during operation [41].

In Chapter 5, we describe the basic properties of on-chip inductive structures, including their geometry, electrical behavior and design guidelines. We present mea-

surement results of devices we have designed and had fabricated, including our work on shifting the resonance point during operation by changing the semiconductor substrate properties. We examine this effect further using a lumped-element circuit model for the inductor. In the final part of the Chapter, we focus on capacitive, rather than inductive, effects and demonstrate the detrimental effect of extra loads caused by bonding pads for signal transfer on circuit operation frequencies.

1.5 Electrical Engineering Education: Motivation

Miniaturization and increasing complexity of microelectronics applications require novel integration techniques and a better understanding of the physical aspects of VLSI design presented so far. In the meantime they themselves are spurred on by the expanding demand for technological innovation and products in the economy, both for industry and for daily life. From communications to medical technology, modern society is more dependent than ever on the fruits of electrical and computer engineering. Thus there is both a need for more and better-educated engineers and an improved understanding of technology in society in general.

Combined with the rapid development of every subfield of electronics, these needs imply that engineering-oriented education must improve continuously. There are a great number of problems that educational researchers focus on, of which we here give a sample:

- Recruitment of a greater number of new students into technical fields, specifically electrical and computer engineering, and retention once they have started

their studies,

- Helping students gain an ever-expanding skillset to be able to meet industry requirements after their formal education, including teamwork, time management, and a deeper understanding of the real-life design process [42] along with approaches to newer technologies each year,

- Introducing the general public to the capabilities and benefits of technology in a systematic manner with some depth.

In Chapter 6, we present our educational contributions, which are relevant to the items mentioned above. Our work is comprised of the design and implementation of three educational programs with varying lengths and intended for students at different levels. Details of design and implementation as well as the outcomes from each program are included.

Chapter 2

Three-Dimensional Integration: Physical Design of A Self-Contained Electronic System

2.1 Introduction and Overview

In this chapter, we introduce three-dimensional integration technology with the physical design of an example functional fabricated system.

The trend towards tighter integration in every level of electronics design has led to larger integrated circuits, even with the ever-decreasing feature size. The system complexity integrated on a single chip is also rising with the higher functionality demanded by technological progress. Thus it has been natural to consider whether expanding the idea of planar integration into the third dimension is viable.

Three-dimensional integration is the extension of planar microchip integration by using some method to stack chips fabricated in a planar technology. This is a nascent idea, pursued through several different approaches in current research. Successful implementation requires a good understanding of the implications in the physical aspects of VLSI design, whether these aspects are geometrical, electrical or thermal. In three aspects of integrated circuit design, a three-dimensional (3-D) chip offers immediately observable advantages. Systems of greater complexity can be integrated into a single unit; hybrid integration is facilitated; and tighter

integration allows intra-subsystem connections with less parasitic loads and power consumption.

Numerous 3-D integration techniques have been proposed in literature. Our focus here is on an approach which uses through-vias to create connections between planar subcircuits integrated vertically [44].

The heart of this chapter presents the physical design, analysis and test of an example application, a proof-of-concept implementation of a self-powering low-power circuit in three-dimensional silicon-on-insulator technology. To our knowledge, this is the first time a self-powering application is demonstrated using photodiodes in three-dimensionally integrated SOI technology. We continue the chapter with an overview of the current state-of-the-art in 3-D integration, including a case study related to one of the active research aspects, heterogeneous integration. We conclude with presenting current research in self-powering systems and methods, with a case study of rectifying antennas.

2.2 Example Implementation: A Self-Powering 3-D System on SOI CMOS

2.2.1 Design Introduction

Here we describe the physical design, analysis and testing of a novel self-powering three-dimensional integrated circuit (3DIC) implemented in a silicon-on-insulator (SOI) technology [45, 46]. The advantages of 3-D integration stated above, their economy of space and power, superior frequency response, and potential lower

noise coupling, are advantages for many technologies, including wireless sensors and distributed networks. A central question is how to power the individual nodes of such systems [57, 58, 60].

We have designed and implemented a system testing the feasibility of power harvesting from ambient light energy in an experimental SOI technology. This microsystem incorporates subunits for energy harvesting and storage, integrating these with an application circuit in a self-contained 3-D unit. The functionality reported provides a proof-of-concept for minute, low-power, self-powering applications. To our knowledge, this is the first self-powering 3DIC demonstrated on SOI technology.

Our system concept for this design uses the three tiers available in the process. Each tier is visualized as fulfilling a certain general function in a specific way:

1. The bottom-most tier, Tier 1, houses the functional electronics. In the case of this design, our functional block is a local oscillator, comprised of three inverters connected in a positive feedback loop followed by a two-stage output buffer. This oscillator has potential uses as a clock circuit or an RF source.
2. The middle tier, Tier 2, is set aside for some storage function. In our system, we have an energy storage element in the form of a square, 30 pF parallel-plate charge storage capacitor. In a computing or sensor system, one can visualize data storage elements placed on this tier.
3. The top tier, Tier 3, is considered for sensor placement. In our case, the energy harvesting necessary for the operation of the system is placed on this tier in the form of photodiode arrays. It is possible to imagine data sensors fabricated on the top level of a similar system.

Figure 2.1 provides a schematic visualization of this system concept. Figure 2.2 focuses on the particular application circuit present in the design.

2.2.2 Process Information

The process MIT-LL uses in this run, labeled "3DL1", is a 0.18 μm, fully depleted silicon-on-insulator (FDSOI) process. This is a three-metal, single-poly process [54]. Three pairs of dopants are provided: **CBN** and **CBP** are the p-type and n-type body threshold adjustment implants, both at 5×10^{17} cm^{-3}. **PSD** and **NSD** are the degenerately doped p-type and n-type source/drain implants. Finally, **CAPP** and **CAPN** are p-type and n-type island implants provided to allow for

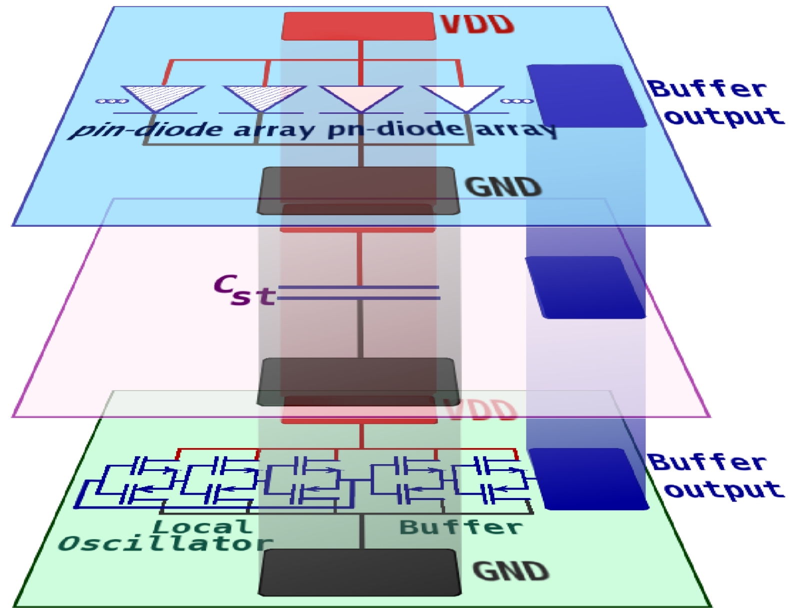


Figure 2.1: Three tiers in a conceptual 3-D system design: The sensor (top), storage (middle) and functional electronics (bottom) levels. The components in our specific design are also shown here.

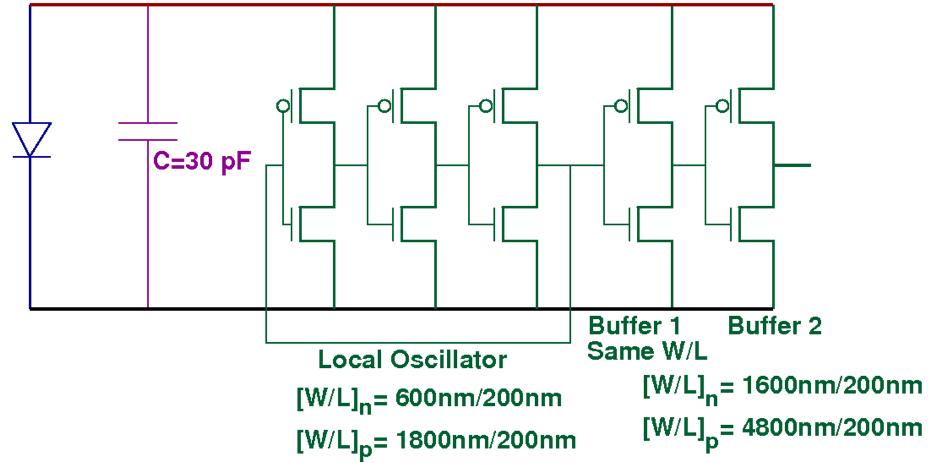


Figure 2.2: The particular circuit being fabricated in our 3-D design. The photodiode symbolizes a diode array comprised of many diodes in parallel.

low-temperature and low-voltage-coefficient capacitors. These implants are doped at 5×10^{18} cm^{-3} and 1×10^{9} cm^{-3} respectively. The undoped silicon is p-type and has a dopant density of about 10^{14} cm^{-3}.

A single tier in the process consists of 50-nm thick silicon islands built on a 400-nm thick layer of buried oxide (BOX), lying over a supportive silicon substrate. The gate oxide is 4.2 nm thick.

The full 3-D chips are assembled after all three tiers are independently fabricated [47]. Tier 1, the bottom tier, retains its silicon substrate and is kept device-side up. Tier 2 is flipped over, aligned and bonded to Tier 1. The silicon substrate is then removed from Tier 2. 3-D (intertier) vias are etched through the Tier 2 oxides and the topmost oxide layer in Tier 1. Tungsten is deposited to create the tier-to-tier connections. After Tier 3 is also flipped over, aligned and bonded to Tier 2 and its silicon substrate removed, the intertier via fabrication process is repeated. Finally,

bond pads are etched down to metal 1 on Tier 3. Figure 2.3, adapted from [47], displays the final 3-D structure.

From the viewpoint of our particular application, it is important to note that since the top tier is inverted, the metal or polysilicon layers do not block outside light from passing through the bottom oxide—net 600 nm thick cap oxide plus BOX,

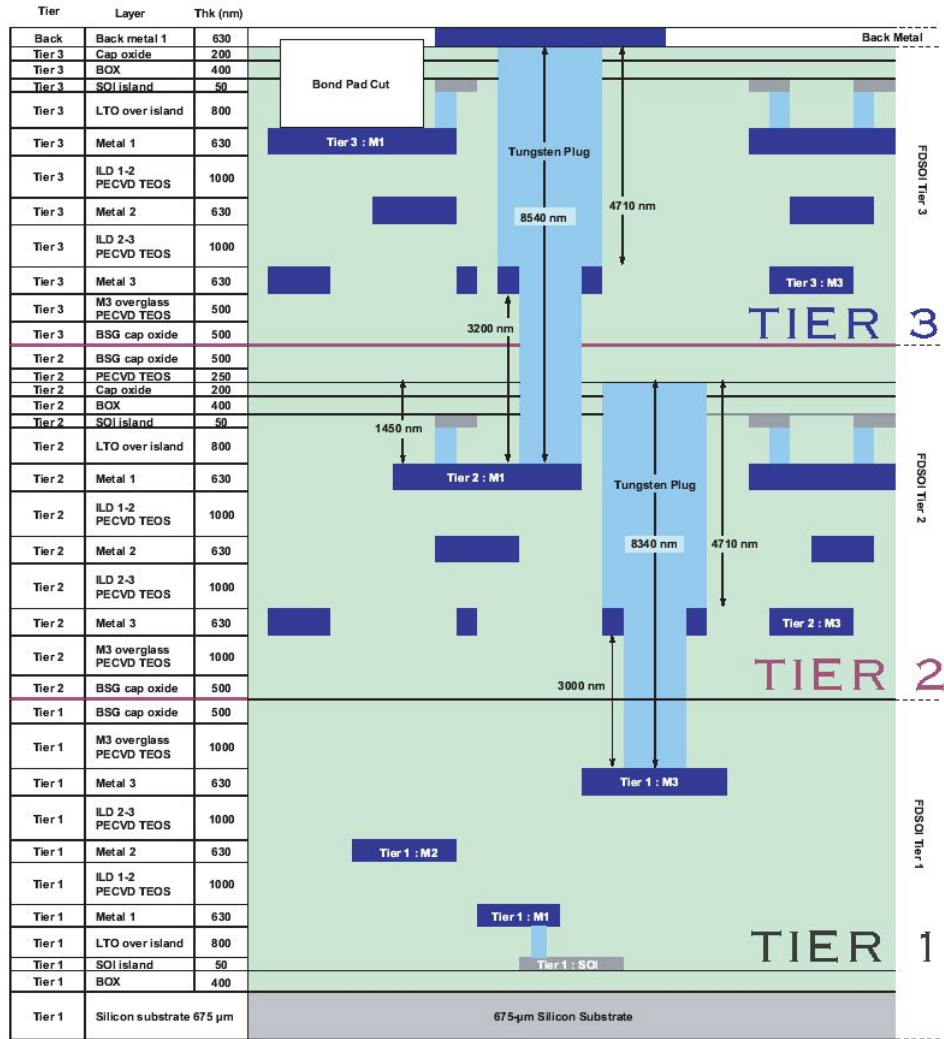


Figure 2.3: The full integration structure for the 3DL1 process. In our design, the photodiodes will be built in the active regions of the top tier, Tier 3.

in this case—to reach the semiconductor islands which house the photodiodes.

2.2.3 Photodiodes: Design Issues

2.2.3.1 Photocurrent Calculation

When photons of sufficient energy fall on the depletion region of a pn-junction, they may be absorbed to create electron-hole pairs. The built-in electric field in the junction then separates these carriers and sweeps them away, the electrons drifting towards the n-side and the holes towards the p-side. Once these carriers hit the bulk regions, if they can diffuse without recombination to the device edges, they contribute to the current outside the device, creating a *photocurrent*.

There are several factors that determine how much photocurrent will be obtained from a photodiode with a given incident optical power falling on the junction [48]. These factors are brought together in the definition of *responsivity*, $\mathcal{R}$:

$$i_p = \mathcal{R} \times P_{inc}, \tag{2.1}$$

where i_p is the photocurrent in amperes, P_{inc} is the optical power incident on the photosensitive region in watts, and $\mathcal{R}$ is thus given in units of A/W.

To calculate the incident power, we need to know the incident intensity and the light-sensitive area: $P_{inc} = I_{ph} \times A$. As a rough guideline, the sunlight intensity on a bright day is about 1000 W/m^2=1$\times$10^{-9} W/μm^2 [49], and a GaInP laser operating at around 670 nm can put out a power of the order of 5 mW, which when focused on 1 μm^2 yields an intensity of 5$\times$10^{-3} W/μm^2 [48].

The responsivity is given by

$$\mathcal{R} = \eta \frac{\lambda}{1.24}, \tag{2.2}$$

where η, *quantum efficiency*, is defined as the ratio of the number of electron-hole pairs that are created to the number of incident photons on the photosensitive area. With that definition in mind, we can show how to arrive at Eqn. 2.2 by relating the photocurrent i_p to the electron flux, ϕ_e and through that, to the photon flux ϕ_p:

$$i_p = q\phi_e = q\eta\phi_p, \tag{2.3}$$

$$\phi_p = \frac{P_{inc}}{h\nu}, \tag{2.4}$$

$$\Rightarrow \quad i_p = \eta \frac{q}{h\nu} P_{inc}. \tag{2.5}$$

$$\frac{q}{h\nu} = \frac{q\lambda}{hc} = \frac{\lambda}{1.24}, \tag{2.6}$$

$$\Rightarrow \quad i_p = \eta \frac{\lambda}{1.24} P_{inc}. \tag{2.7}$$

Hence Eqn. 2.2. For clarification, we reiterate that responsivity is in units of A/W, while η is unitless, and λ is given in units of μm.

The quantum efficiency η is the factor that depends most on the structural design of the photosensitive device. To understand how it is obtained, and also to be able to calculate the incident optical power, we need to know the photosensitive area and depth of the diode structure in question. Due to the resulting structure of this SOI process, every type of n- or p-doped silicon available to the designer takes up the entire active silicon depth in the semiconductor islands of the top tier. Therefore, the photodiode junctions designed naturally have vertical different-dopant-type semiconductor interfaces and lateral current flow. Considering top

illumination on the final integrated 3DIC, the illuminated photosensitive area of a diode will be determined by its junction width (W_j) and depletion region width (W_d), as shown schematically in Figure 2.4.

Keeping this picture in mind, the quantum efficiency η is given as a combination of three factors:

$$\eta = \xi(1 - r)(1 - \exp(-\alpha d)) \tag{2.8}$$

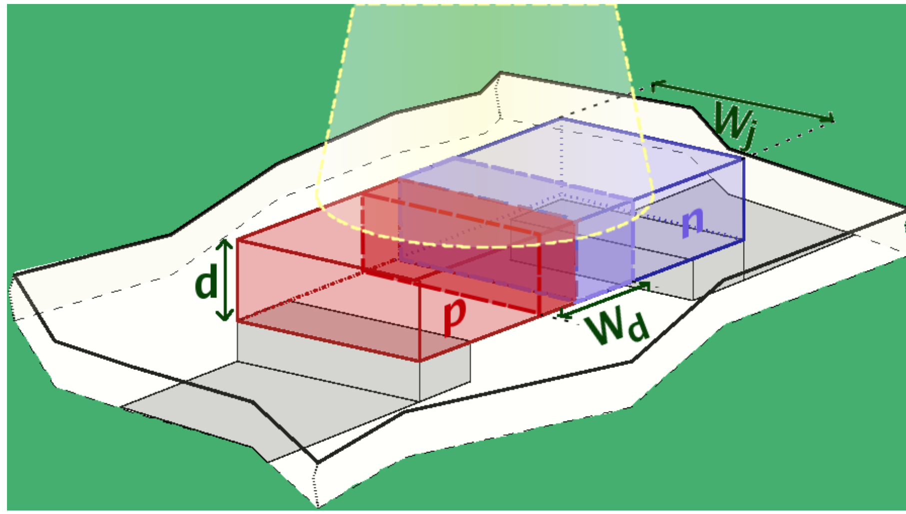


Figure 2.4: The schematic representation of a pn-junction diode fabricated in this SOI process and of the photosensitive area to be taken into account. Since the diode tier is integrated into the 3DIC upside-down with respect to its fabrication orientation, the metal connections to the anode and cathode of the diode are shown as going "downwards". In this case, the effective photosensitive area is $A = W_d W_j$, and the light passes through the active silicon depth d for a single pass. In this picture, the following are not shown for clarity: The oxide layers surrounding the active silicon layer which houses the diode, and the polysilicon layer covering the junction, which is between the metal layers and the active layer in the process.

Let us now examine these factors one by one.

- ξ is the fraction of created photocarriers that reach the outer circuit without recombination. With a value between 0 and 1, it is mostly governed by the material quality and surface recombination. In silicon photodetectors, it is possible to get quantum efficiencies close to 1, due to the high quality of the silicon crystal [48].
- r is the optical power reflectance from the surface. For the first pass of light through our device, there are two interfaces at which some of the incident power is reflected: Air-silicon dioxide and silicon dioxide-silicon. For a simple calculation, we will consider normal incidence and ignore multiple reflections in the oxide layer. The refractive in index of air, n_{air}, is considered 1; of silicon dioxide, n_{SiO_2}, as 1.46 [43], and silicon at 633 nm, n_{Si}, as approximately 4.0 [50]. The reflectance from the surfaces will then be

$$r_1 = (\frac{n_{air} - n_{SiO_2}}{n_{air} + n_{SiO_2}})^2 = 0.035\,, \tag{2.9}$$

$$r_2 = (\frac{n_{SiO_2} - n_{Si}}{n_{SiO_2} + n_{Si}})^2 = 0.216\,. \tag{2.10}$$

Thus, 96.5% of the incident optical power will reach the semiconductor surface, and 78.4% of this will enter the semiconductor—that is, about 75.7% of the incident optical power. Let us assume that in our case ξ is 0.9, and thus

$$(1 - r)\xi \approx 0.68\,. \tag{2.11}$$

It should be noted that the "deeper" interfaces of the stacked tiers work to our advantage after the initial pass of light through the diode layer. Still con-

sidering normal incidence, the first reflection from the oxide layer immediately underneath the diodes allows 16.1% of P_{inc} to reenter the silicon, and reflection from metal layers further underneath allows another 32.4% to rereach the diodes. Figure 2.5 shows some of the multiple reflections in question, demonstrating absorption amounts in the silicon and polysilicon layers.

- The expression $(1-\exp(-\alpha d))$ indicates how much of the photon flux streaming through the photosensitive material will be absorbed. Here d is the total thickness of the material that the photons pass through, and α, the *absorption coefficient*, is a material property. For silicon, α is about 3.5×10^{-4} nm^{-1} at 633 nm (red light). It increases with the photon energy and is around 10^{-3} nm^{-1} at UV frequencies.

  In our design, this factor turns out to be the bottleneck. A challenge of SOI processes is that the thin film of the active layer does not allow for a great deal of photoabsorption; this challenge should be met in the attempt to use SOI to implement a self-powering 3DI circuit while subject to this constraint. Considering top illumination, the relevant silicon depth, d in Figure 2.4, is only 50 nm. This yields $(1-\exp(-\alpha d)) = 0.017$ for red light. In other words, only 1.7% of the incident light power will be absorbed during a single vertical pass through the material.

Following this discussion, we can obtain an estimate about how much photocurrent we can generate in our system per micron-square of photosensitive area, assuming red light (0.633 μm) and an intensity of 1000 W/m^2, and ignoring the

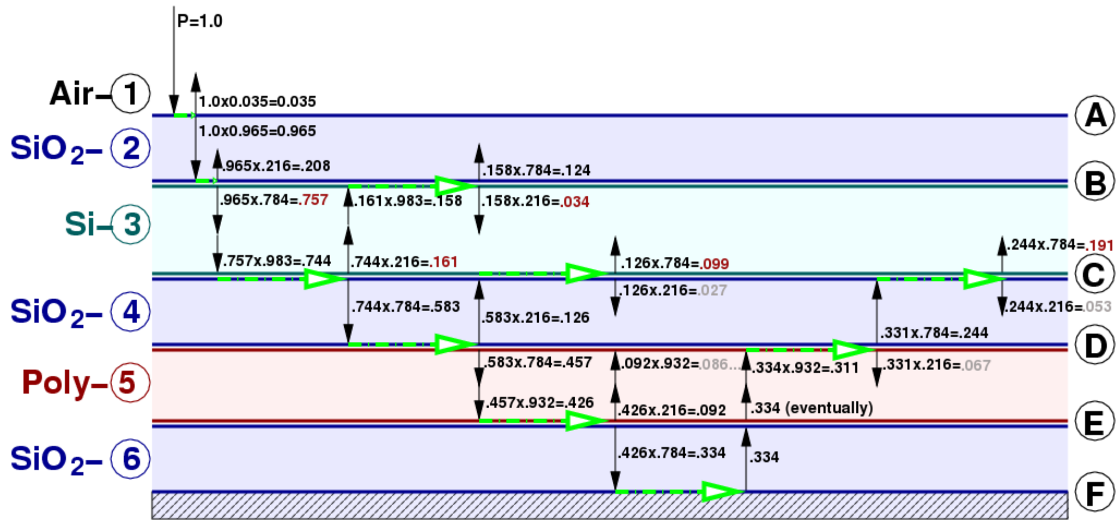


Figure 2.5: The multiple reflections in our multi-layer thin film system. The layer thicknesses are not shown to scale. The transmission layers are numbered 1 through 6 and interfaces named A through F. Below the SiO_2 layer 6 a metal layer is shown, with 100% reflection assumed at that interface. Every time an incident ray hits an interface, a green arrow points at it splitting into transmitted and reflected components. With the incident power level set at 1.0, the fractional numbers show the percentage reflected or transmitted at each interface. The numbers shown in red are thus power levels entering and re-entering the active silicon layer. When the power level in a particular ray trace falls below 0.1 (i.e. 10%), that number is given in gray and that particular ray is not traced any further. Note that through multiple reflections layer number 6, eventually all of the power entering layer 6 from layer 5, 33.4% of the incident power, will be reflected back into layer 5; all the back and forth of that process is not shown in the figure.

multiple reflections for now to get the worst-case yield:

$$i_p = \eta \times \frac{\lambda}{1.24} \times 10^{-9}\ \mathsf{W}/\mu m^2 = \eta \times 0.51 \times 10^{-9} \quad (2.12)$$

$$= (0.68 \times 0.017) \times 0.51 \times 10^{-9} = 0.0116 \times 0.51 \times 10^{-9}\ , \quad (2.13)$$

$$\Rightarrow\ \mathbf{i_p \approx 6\ pA/\mu m^2} \quad (2.14)$$

This current can be increased by using a higher intensity light source like a laser and picking an optimal wavelength—simply increasing the wavelength to exploit the $\lambda/1.24$ component will not work due to decreasing absorption at lower photon energies.

Given these physical constraints setting a limit on the amount of photocurrent generated per area, the design problem now facing us becomes how to maximize the photosensitive area.

2.2.3.2 Photodiode Design and Layout

As was illustrated in Figure 2.4, the useful area for photocurrent generation in our case is the top-view cross-section area of the diodes' depletion regions. This area in turn depends on the layout and depletion region width. For our diodes, we can choose pairs of doping implants from the list of available implants given in Section 2.2.2.

The option that yields the widest depletion region and thus the largest photosensitive area is using the n-type threshold adjust implant (CBP) and undoped silicon. Using the relation

$$W_d = [\frac{2\epsilon_{Si}}{q} V_{bi} \frac{N_A + N_D}{N_A N_D}]^{\frac{1}{2}} \quad (2.15)$$

where $V_{bi} = V_{thermal} \cdot \ln\left(N_A N_D / n_i^2\right)$ is the built-in potential of the junction, this results in a depletion region width of about 1.5 μm. If we design the diode layout such that the junction is 10 μm wide (across), this creates an area of 15 μm^2 and, from Eqn. 2.14, about 90 pA per diode.

However, using the very lightly doped “substrate” material (p-type silicon, doping around 10^{14} cm^{-3}) is not recommended by the fabrication facility [51]. Among the possible problems is the concern that such “intrinsic” regions are vulnerable to possible surface accumulation or inversion if a poly or metal layer over the region becomes charged. Since we need to cover our metallurgical junctions with polysilicon in order to provide silicide protection due to the specific fabrication process steps, this is a real concern for us. Thus we chose not to rely solely on this design type.

Figures 2.6 through 2.8 display this diode: Only the implant regions and contacts, implants plus poly, and the full layout shown, respectively. The junction is between the CBP and “intrinsic” regions, i.e. there are two junctions, one to either side of the central strip, in Figure 2.6. The heavy-n-doping NSD implant is for ohmic contact to the n-side, and PSD for the p-side. Polysilicon covers all the intrinsic region and the junctions and serves the dual roles of silicide protection and implant alignment during fabrication. 26 of these structures are combined to obtain 52 diodes of this type, comprising a pin-diode array of 52 parallel diodes.

The next best option, in terms of wide-depletion-region diodes, is using the two threshold adjust implants, CBN and CBP, as the p-side and n-side respectively. This means the dopant concentrations on both the n- and p-sides will be at 5×10^{17}

Figure 2.6: Implants for the pin-type diode.

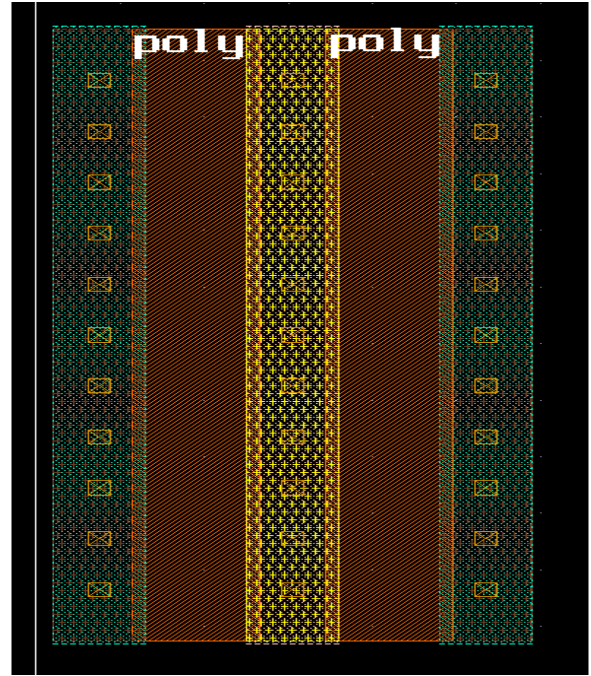


Figure 2.7: Implants plus poly for the pin-type diode.

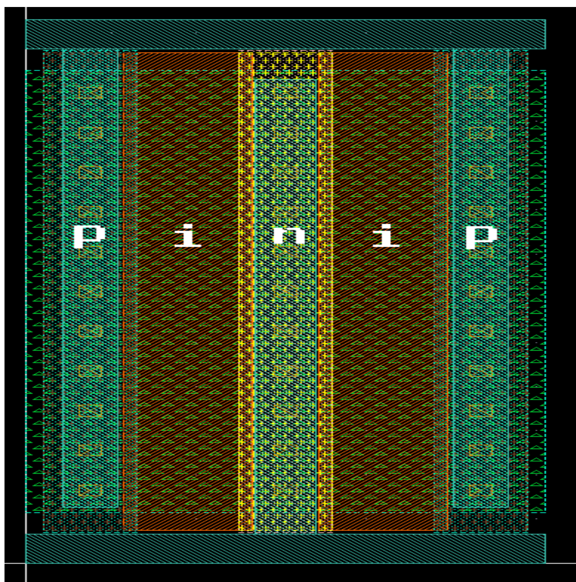


Figure 2.8: The full layout for the pin-type diode.

cm^{-3} and results in a depletion region 0.0684 μm wide. To easily stack many of these diodes into an array, we have chosen an annular design, where the pn-junction is a square 2 microns long on one side; thus the photosensitive area per diode is

approximately 0.55 μm^2 and we can obtain 3.3 pA per diode.

Figures 2.9 through 2.11 display the layout of this diode: Similar to the case for the lateral diode, only the implant regions and contacts shown, implants plus poly shown, and the full layout shown, respectively. The junction is between the CBP and CBN regions, i.e. as a ring around the CBP square in the center, which is two microns per side. The n-side in the center is connected to a plus-shaped metal-2 unit cell to form a 2D metal-2 mesh. The p-side on all four sides is connected to a square-shaped metal-1 unit cell to form a 2D metal-1 network.

The question arises about the optimal size of these diodes in order to achieve the maximum photosensitive area. Our constrains are the design area constraint (250 μm by 250 μm), the depletion region width (fixed by the dopant levels) and the minimum region separations and feature sizes allowed by the process.

Some consideration indicates that a maximum number of diodes, each with the smallest surface area achievable with the allowed separations of the process, should be used to maximize photosensitive area. To demonstrate this conclusion, let W, L, dW and dL be the chip width, length, single diode width, and single diode length respectively. Thus we can fit in $(W/dW) \times (L/dL) = (WL)/(dWdL)$ diodes in our available chip area. Also, let X_d be the depletion region width achieved by the dopants we have chosen. Then the photosensitive area per diode is slightly less than, almost equal to, $A_{psD} = 2X_d(dW + dL)$. Therefore our total photosensitive area will be

$$A_{psT}(dW, dL) = 2X_d WL \frac{dW + dL}{dWdL}. \tag{2.16}$$

Since the $1/dWdL$ term in this function of diode length and width increases faster with decreasing dW and dL than the $(dW + dL)$ term decreases with the same, the maximum area is obtained by the minimum possible dW and dL. In other words, this is a monotonically increasing function both in the decreasing dW and dL directions. Therefore, we chose the smallest dimensions allowed by the process width and spacing rules for our annular diodes. 2062 of these diodes are stacked into an array of parallel pn-diodes in the final layout.

This array of 2062 pn-diodes makes use of all the space on the top tier not occupied by the pin-diode array, the required bonding pads for measurements and the intertier via landings. The next section presents the layout of this tier as well as the other two tiers.

With this total number of diodes and under the rather stringent illumination conditions assumed in Section 2.2.3.1, we expect to obtain about 15 nA of photocurrent. The question is whether this current will be sufficient to run a local oscillator of the design given in Figure 2.2.

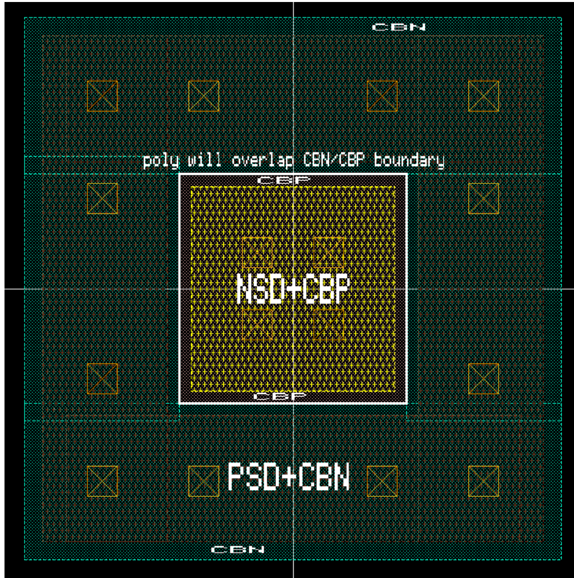


Figure 2.9: Implants for the CBN/CBP diode.

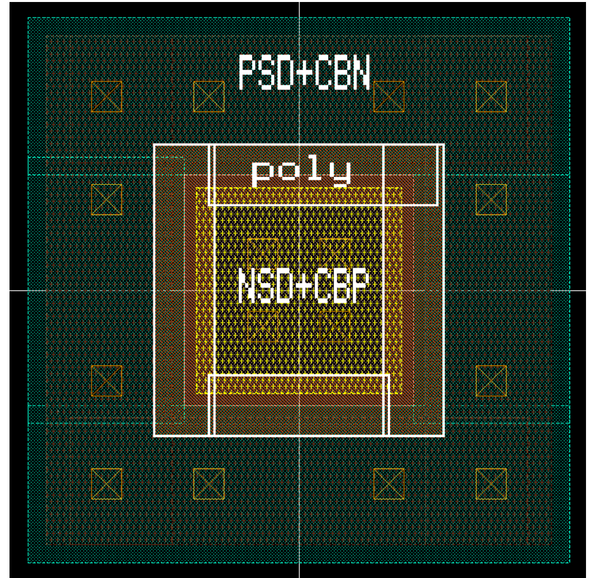


Figure 2.10: Implants plus poly for the CBN/CBP diode.

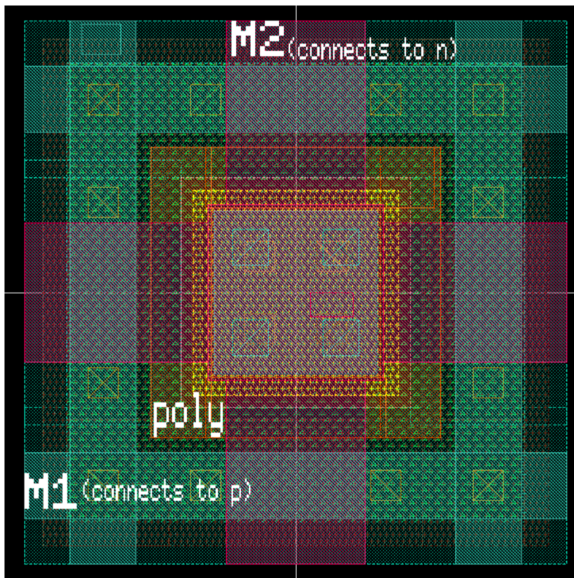


Figure 2.11: The full layout for the CBN/CBP diode.

2.2.4 Layout and Chip Microphotographs

2.2.4.1 Tier 1: The Local Oscillator

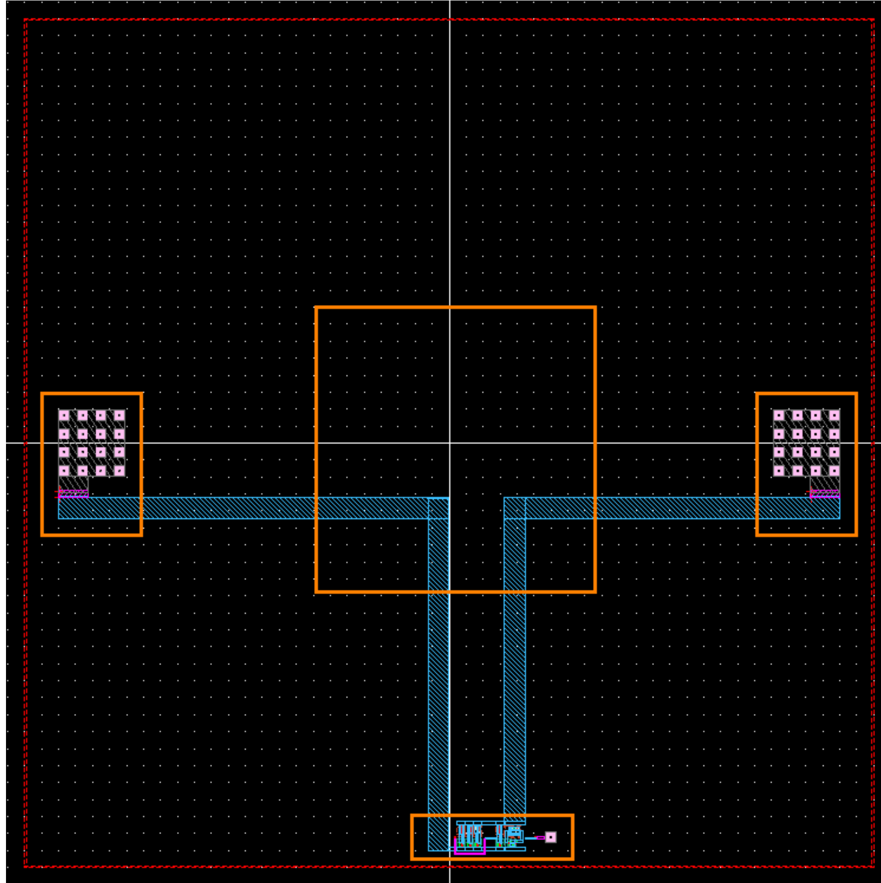


Figure 2.12: Local oscillator (Tier 1—bottom tier) layout. To the left and right are the 3-D vias coming from the higher tiers, for GND and VDD rails respectively.

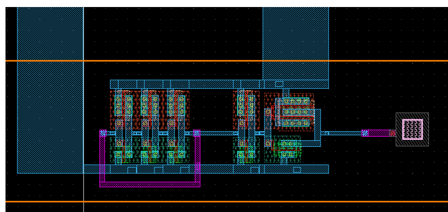


Figure 2.13: A zoom of the local oscillator, output buffers and the 3D via carrying the output signal up.

2.2.4.2 Tier 2: The Capacitor

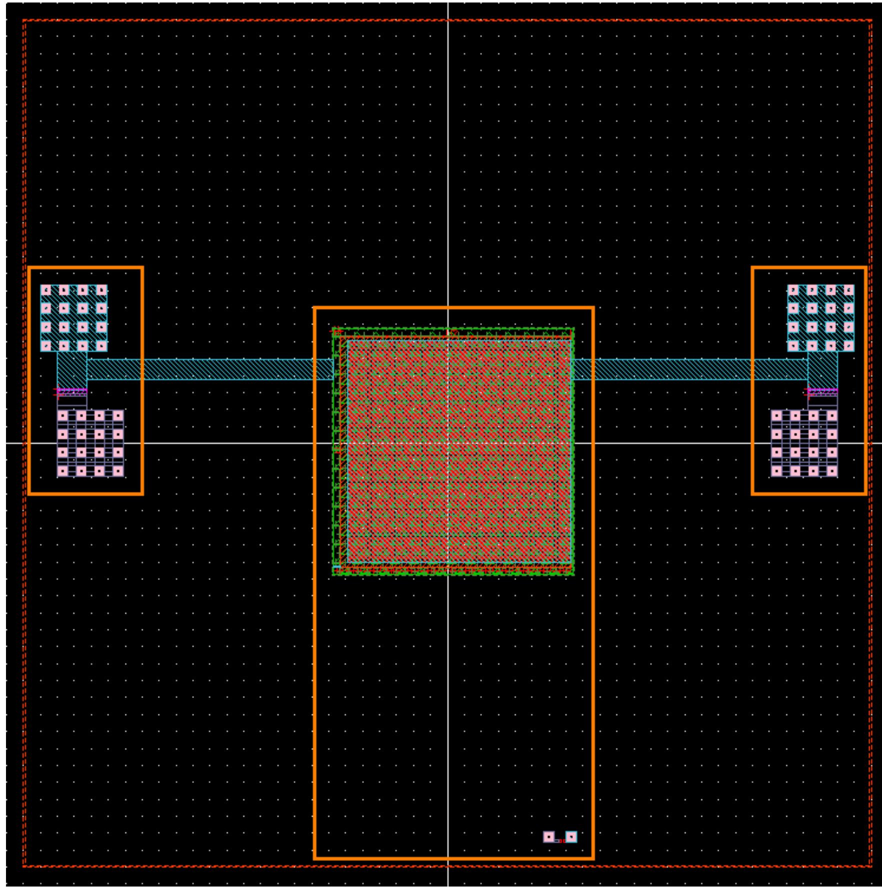


Figure 2.14: The capacitor tier (Tier 2—middle tier) layout. The capacitor top plate is a poly square, 67.9 μm on one side. The bottom plate is produced using the CAPN n-type implant. The expected capacitance of this structure is 30 pF and the Cadence-extracted capacitance is 29 pF.

2.2.4.3 Tier 3: Photodiodes, Measurement Pads

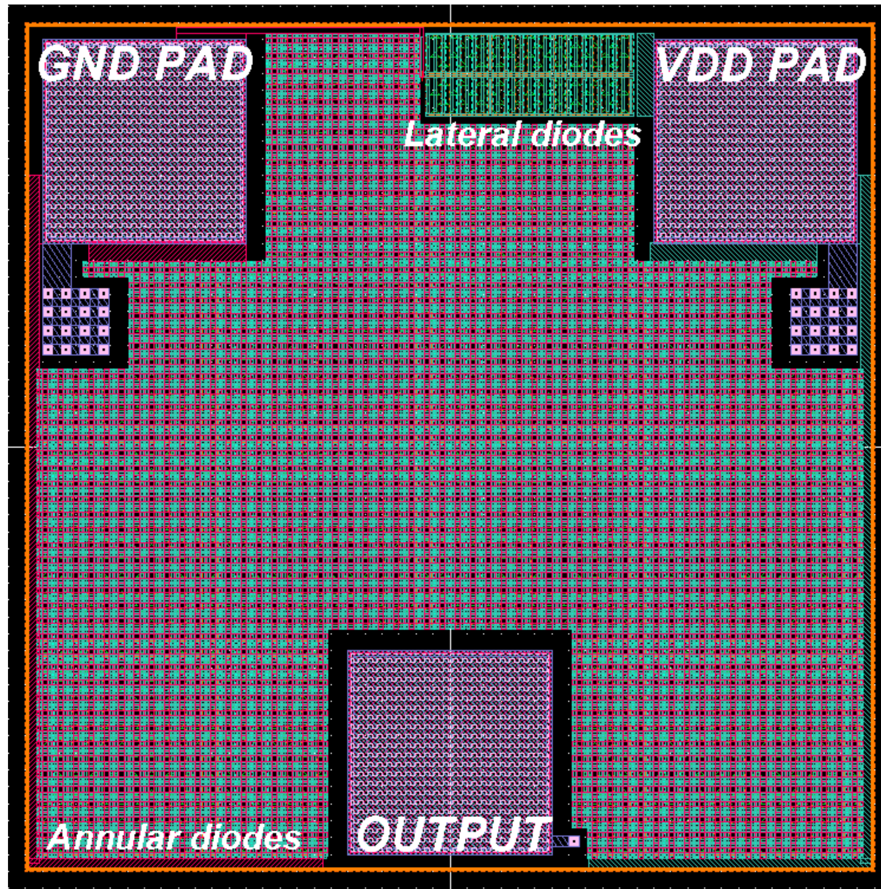


Figure 2.15: The diode tier (Tier 3—top tier) layout. The bondpads have overglass cuts to allow access to the Metal 1 layer of the tier with probes or bondwires. Most of the tier is covered by the array of 2062 annular CBN/CBP diodes, and there are also 52 lateral "pin" diodes near the top of the layout. All features are identified in the chip microphotograph, Figure 2.16. The layout is a square 250 μm to a side.

2.2.4.4 Chip Microphotographs

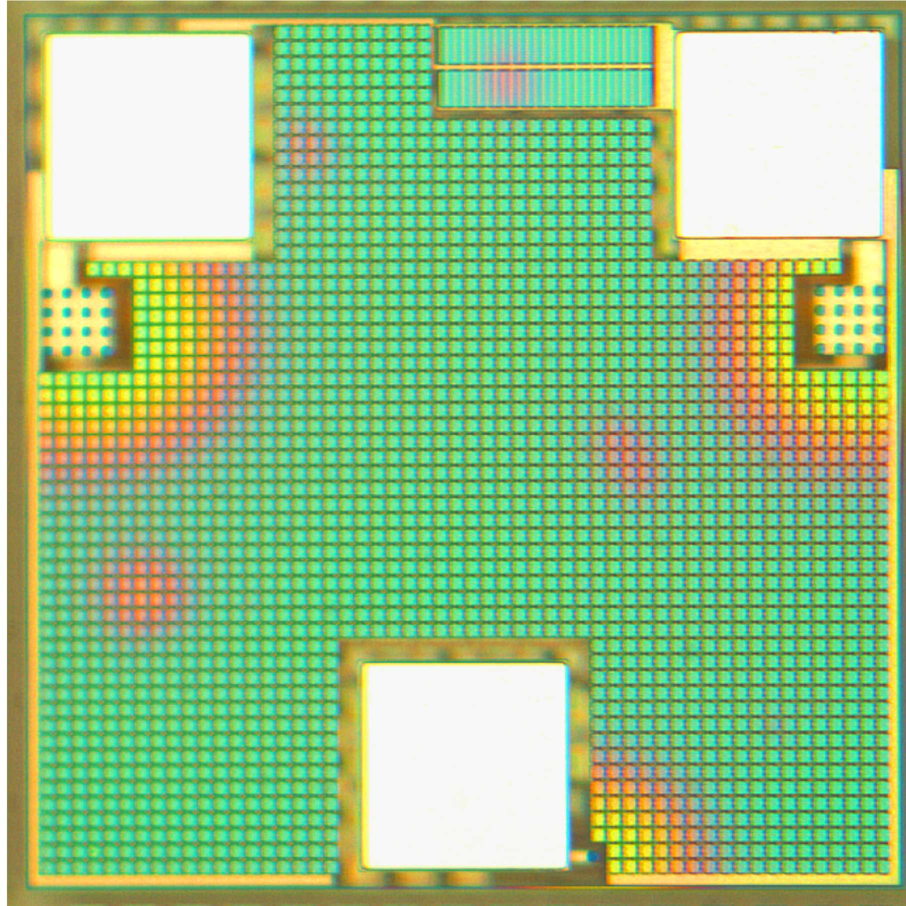


Figure 2.16: The fabricated 3DIC microphotograph. The top tier is visible. Top right: VDD pad. Top left: GND pad. Bottom middle: The oscillator output pad. The pin-diode array is next to the VDD pad. The rest of the tier is covered by the annular pn-diode array. The intertier (3-D) via arrays to the lower tiers from the VDD and GND pads are below the pads; the intertier via from the lower tier to the output pad is to its right side. Lower tiers are somewhat visible, though out of focus, at the bonding pad edges, where there are no top tier features defined.

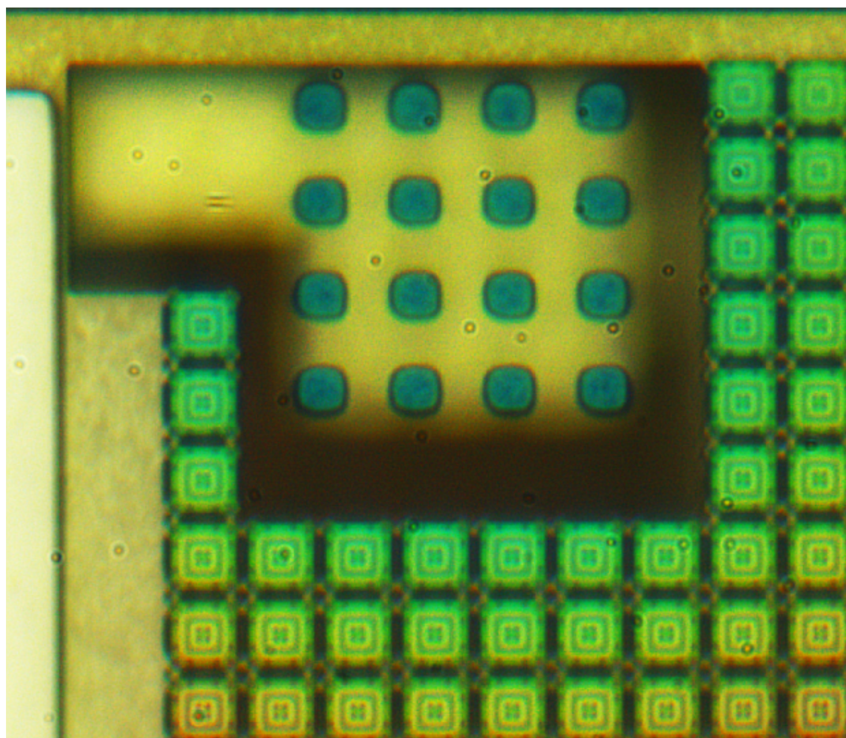


Figure 2.17: A close-up of the intertier via landing next to the GND pad and the corner of the surrounding annular pn-diode array. The four dark squares at the center of the each diode are the stacked contacts and vias from the n-active region to metal-1 and metal-2. The eight squares on the sides of each diode are contacts from the p-active region to metal-1. The polysilicon square covering the junction for silicide protection is also visible in each diode.

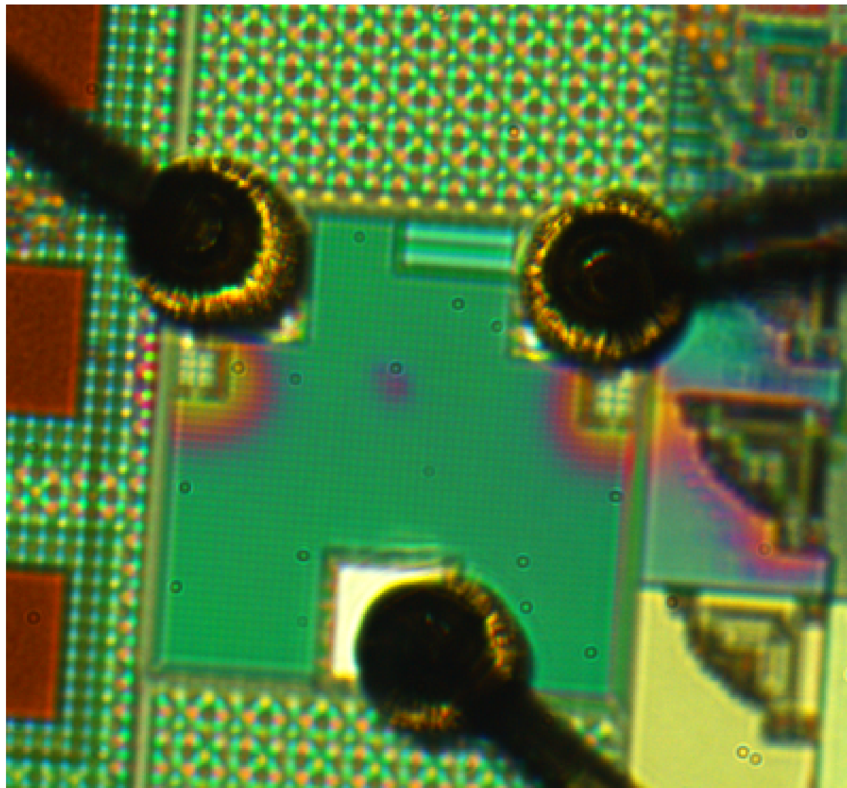


Figure 2.18: A close-up of the bonded die. Surrounding our design are test structures from the process.

2.2.5 Circuit Analysis and Simulations

Our application circuit is a local oscillator of the type which might be used in very low-power wireless sensor notes. It is comprised of three minimum-size inverters with midrange switching points in a ring-oscillator configuration. Following the oscillator is a two-stage buffer (See Figure 2.2).

2.2.5.1 Circuit Operation: Qualitative Description

Assuming a 15 nA photocurrent and using a 30 pF capacitor as the middle tier energy storage device, we simulated the circuit using the MOSFET models provided by the fabrication facility.

The overall circuit operation calculated by the circuit simulator Spectre [52] is given in Figure 2.19. The three traces in the figure are the rail voltage across the capacitor, the signal supplied to a capacitive load by the output buffer and the signal taken between the oscillator stages. The load capacitance is set at 15 fF, the capacitive load of the output bonding pad [45].

We can qualitatively describe the circuit operation as displayed in Figure 2.19 as follows: The photocurrent, charging the storge capacitor C_{st}, builds up the inverters' rail voltage. Inverter output and input voltages follows; at a certain level, the local oscillator has sufficient gain to start oscillation. Afterward all inverters are drawing current. This total current increases with the rising rail voltage as the photocurrent charges the storage capacitor C_{st}, until the current consumption balances the photocurrent and the rail voltage stabilizes. A quantitative analysis of

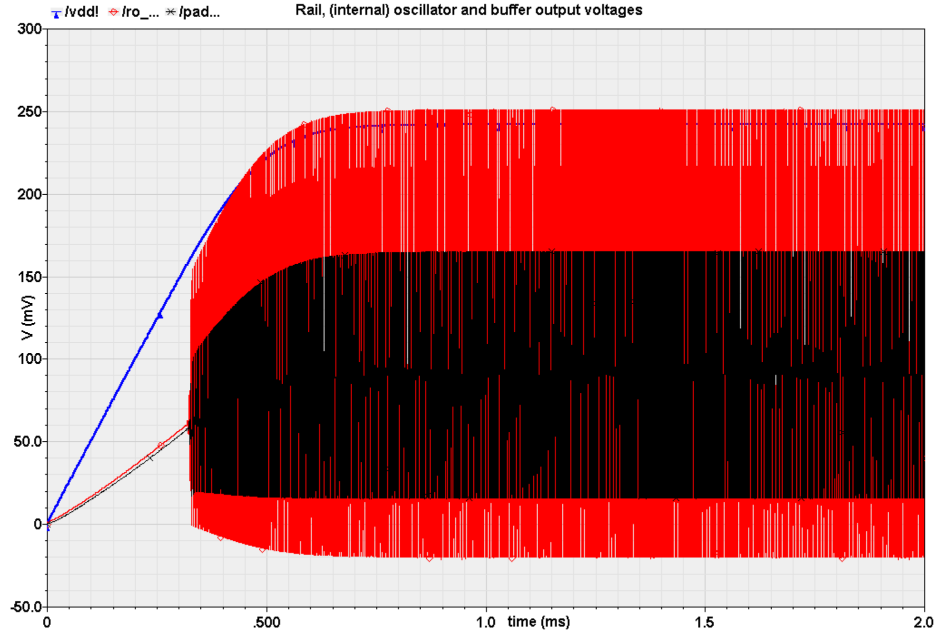


Figure 2.19: System operation with $i_p = 15$ nA. The voltage scale is between -50 mV and 300 mV, and the blue line rising faster in the beginning is the rail voltage, which reaches a steady state of 242 mV in this simulation. The lighter solid line is the ring oscillator output feeding the buffer and the thick solid line is the buffer output to a 15 fF load (bonding pad only).

this operation is provided in the next section.

With a photocurrent of 15 nA, the simulation shows this system driving a 15 fF load at 1.82 MHz, with an equilibrium rail voltage of 242 mV. Figure 2.20 zooms in on the full-oscillation region of system operation. Figure 2.21 focuses on the region where the output signal oscillation is becoming perceptible.

Higher photocurrents yield higher rail voltages and oscillation frequencies. Figure 2.22 displays the rail voltages and oscillation frequencies for a range of photocurrent levels. Figure 2.23 depicts the circuit operation with $i_{ph} = 40$ nA.

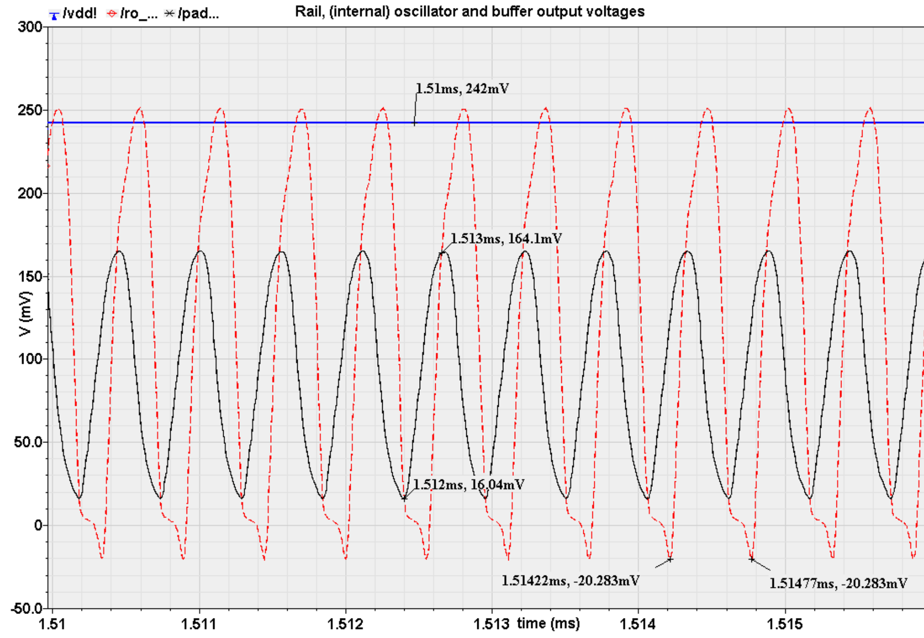


Figure 2.20: Zoom in on the full oscillation region of the simulation in Figure 2.19.

The oscillation frequency in this case is ∼1.82 MHz.

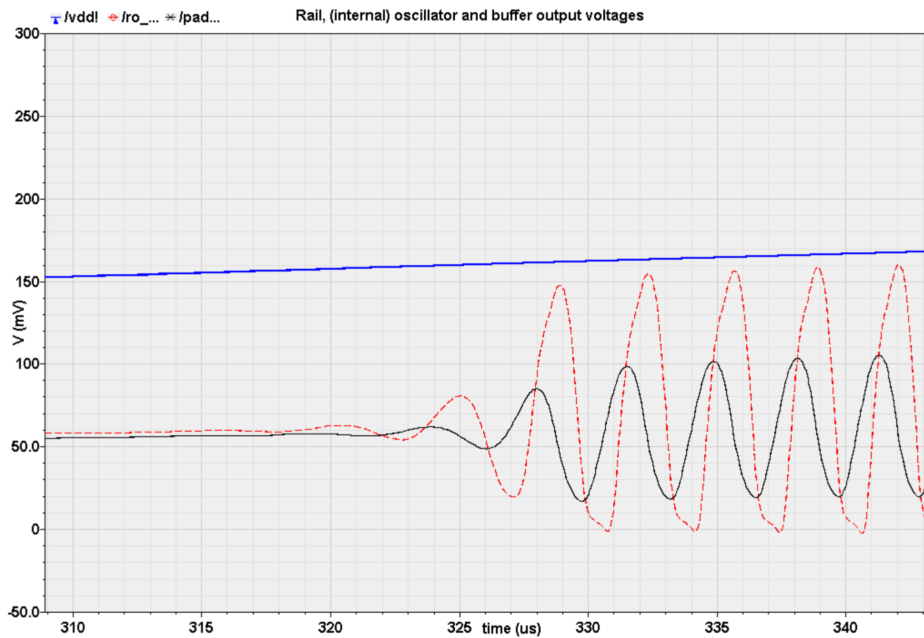


Figure 2.21: Zoom in to Figure 2.19 around the region where oscillation is beginning.

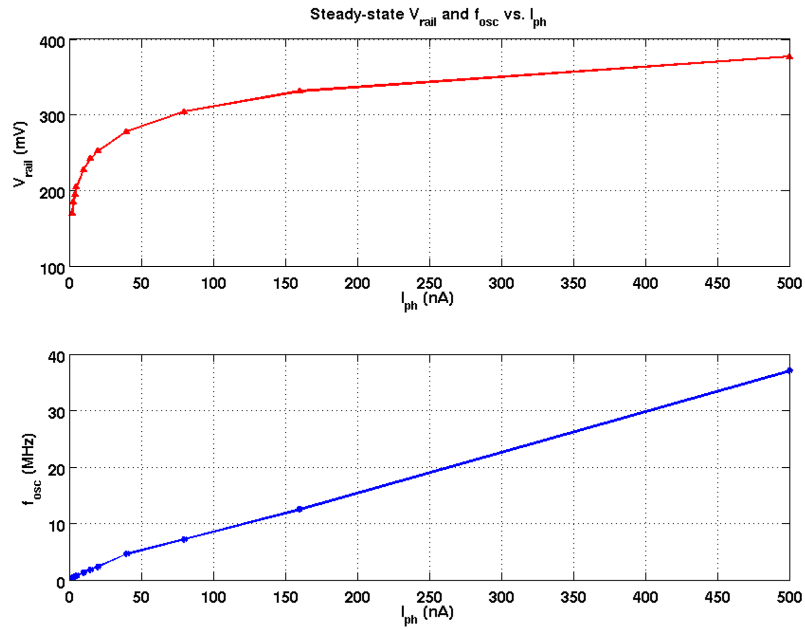


Figure 2.22: Simulated steady-state rail voltages (top) and oscillation frequencies (bottom) vs. photocurrent.

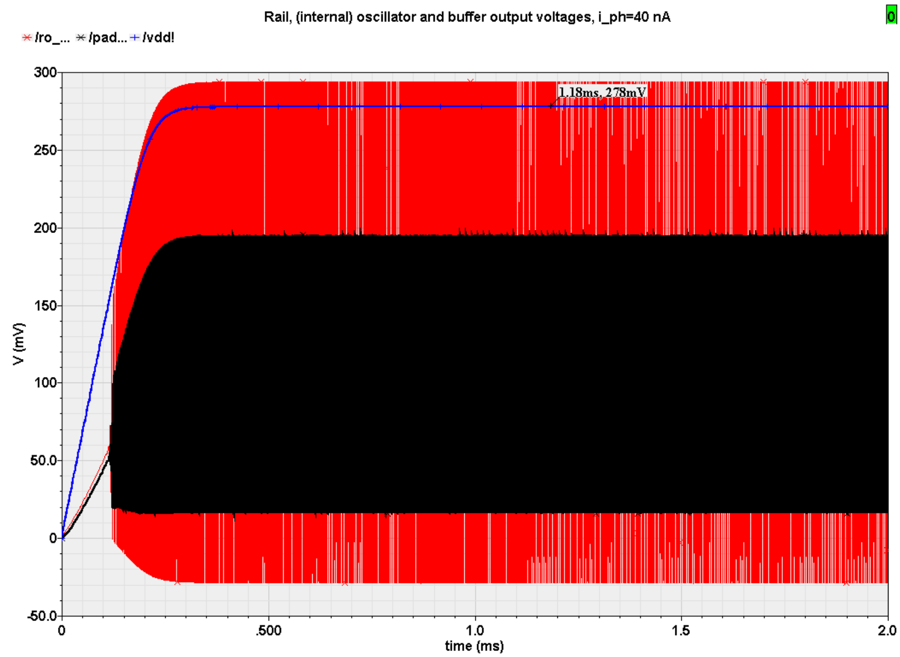


Figure 2.23: System operation with $i_p = 40$ nA and storage capacitor 30 pF.

Note that at higher photocurrent levels, another limiting factor will appear. Since the photocurrent flows from the n-side to the p-side, it charges the storage capacitor to a voltage oriented to provide a forward bias on the photodiodes themselves. This causes the photodiodes to draw a higher forward bias current along with the inverters drawing their own operation current.

To analyze this effect we calculate the forward-bias current of the diodes. For a basic pn-junction diode the current-voltage relationship is given by [43]

$$I_{Dfb} = I_s(e^{\frac{V_{Dfb}}{V_{th}}} - 1) \tag{2.17}$$

where I_{Dfb} and V_{Dfb} are the forward bias diode current and voltage, I_s is called the *saturation current* and V_{th} is the thermal voltage, ∼0.026 V at room temperature. The saturation current depends on the material properties and diode design. Particularly, the saturation current density for a short-base diode (i.e. a diode for which the charge-neutral region length away from the junction past the depletion region boundary is significantly shorter than the diffusion length) is

$$J_s = qn_i^2(\frac{D_p}{N_d W_n'} + \frac{D_n}{N_a W_p'})(e^{\frac{V_{Dfb}}{V_{th}}} - 1) \quad . \tag{2.18}$$

Here, D_p and D_n are the hole and electron diffusion coefficients, N_d and N_a are the n-region donor and p-region acceptor densities (both at 5×10^{17} cm^{-3} for the pn-junction diode), and W_n' and W_p' are the lengths of the quasi-neutral n-type and p-type regions after the relevant part of the depletion region is subtracted from the drawn region length. Taking the hole and electron mobilities at 200 and 400 cm^2/Vsec for these doping levels, the diffusion coefficients are 5.2 and 10.4 cm^2/sec respectively. A calculation of the depletion layer width and consulting the diode

layout for the distance between the drawn junction and ohmic contacts up to the metal layers yields both W'_n and W'_p as 0.491 μm.

This sets the saturation current at around 8.56×10^{-20} A for the CBN/CBP pn-junction diodes. A similar calculation shows that the lateral pin-junction diodes have a saturation current of 0.168 fA.

Our simulations show that a photocurrent of 300 nA results in a steady-state rail voltage of ~350 mV, which using Eqn. 2.17 gives about 60.1 fA per pn-junction diode and 0.12 nA per pin-junction diode. Thus the net forward bias current draw of the two diode arrays will be nearly 6.36 nA. We can conclude that the "true" rail voltage will stabilize at a somewhat lower level than 350 mV and that this effect is not prominent at low illumination levels, where the rail voltage stays low. However, since the diode forward bias current increases exponentially with the rail voltage, at a certain higher illumination level, I_{DFB} becomes the factor balancing the photocurrent. Thus it limits the current available to the inverters and affects where the steady-state rail voltage will be set. We will consider this issue further in Section 2.3, where we describe the design work done for the second-generation version of this 3-D process.

2.2.5.2 Circuit Operation: Quantitative Analysis

Here we provide the analysis behind our description of the circuit operation. At a low photocurrent level where we can ignore the diode forward bias currents, the

evolution of the rail voltage across the storage capacitor is governed by the equation

$$C_{st}\frac{dV_{rail}}{dt} = (I_{ph} - \sum_{i=1}^{5}\Phi_i I_{inv_i}(V_{rail}))\ . \tag{2.19}$$

V_{rail} is the rail voltage across the storage capacitor and inverters. Φ_1 through Φ_5 are the ratios of time-integrated current drawn by each inverter per period to the peak inverter current integrated over one period. $I_{inv_1}(V_{rail})$ through $I_{inv_5}(V_{rail})$ are the currents drawn by the five inverters. The dependence of these currents on the rail voltage allows us to consider a simplified model for the circuit, as presented in Figure 2.25a. In this model, the voltage-controlled current source (VCCS) represents the sum of the currents drawn by the inverters. It is straightforward to demonstrate that, if the VCCS is replaced by an equivalent resistor R_{eq}, where $I_{VCCS} = V_{rail}/R_{eq}$, Eqn. 2.19 yields

$$V_{rail}(t) = I_{ph}R_{eq}(1 - e^{\frac{-t}{R_{eq}C_{st}}}) \tag{2.20}$$

i.e. the voltage exponentially approaching the fixed level of $I_{ph}R_{eq}$. For our purpose, the group of inverters can be considered as equivalent to a voltage-controlled resistor, their current-voltage behavior providing $I_{inv_i}(V_{rail})$ for use in Eqn. 2.19.

For this submicron SOI technology, the threshold voltage magnitudes are given as 0.53 V for the NMOS and 0.64 V for the PMOS [53]. It is therefore safe to assume, given the previous calculations, that our transistors will always operate in the subthreshold region. To obtain the voltage-dependent currents drawn by the three small-size inverters of the local oscillator and the small-size and large-size inverters of the buffer, we performed a DC analysis. A transient simulation reveals the fraction of time each of these inverters would be drawing peak current per period

averages to nearly 33%, which is expected since the oscillator has three stages. More precisely, Φ_1 through Φ_5 are, respectively, 0.2774, 0.4136, 0.2763, 0.3294 and 0.3286 for the inverters in the order of oscillator inverters, first and second buffer stages. Figure 2.24 shows the current through the source of the fourth inverter PMOS; the highlighted area marks a period and the current integrated over a period is cross-hatched. Φ_4 is the ratio of the cross-hatched area to the integrated area.

Using these values, $C_{st} = 30$ pF and $I_{ph} = 15$ nA in Eqn. 2.19, we calculate how the rail voltage builds up over time. Figure 2.25b displays the result of this calculation, comparing it to the $V_{rail}(t)$ obtained by the Spectre simulation of Figure 2.19. The calculation is demonstrated to be accurate at steady-state. The error during the buildup period, a slower rate of increase, is explained by considering that

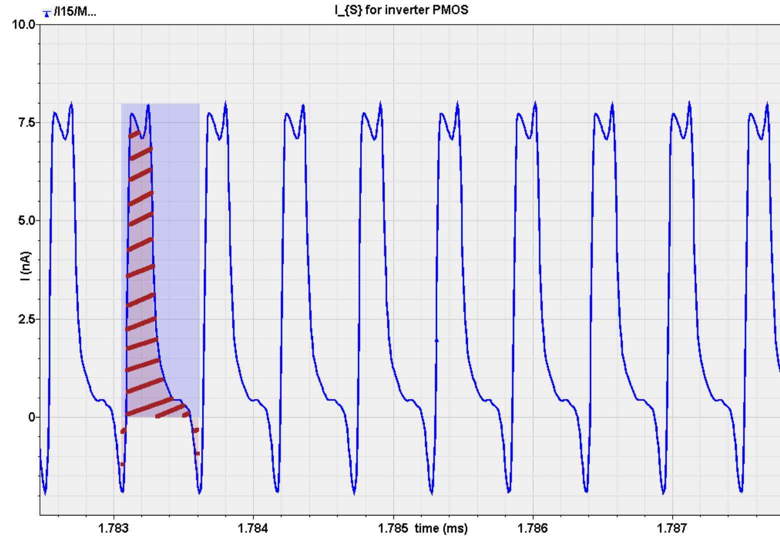


Figure 2.24: The PMOS source current of the fourth inverter (first buffer stage) during full oscillation. The ratio of the cross-hatched area to the highlighted area gives Φ_4, which is interpreted as the effective duty cycle.

both the inverter currents and the coefficients $\Phi_1..\Phi_5$ used in the calculation were for steady state.

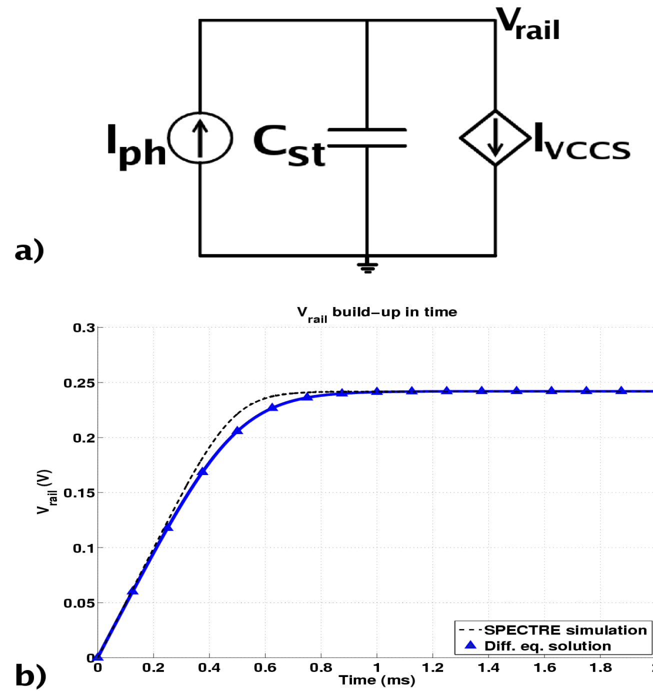


Figure 2.25: **a)** Simplified equivalent circuit to demonstrate the system operation. **b)** The rail voltage of the circuit builds up in time to reach a steady state as the current drawn by the inverters from the storage capacitor, rising with increasing rail voltage, reaches an equilibrium with the supplied photocurrent. The solid line with the triangle markers display the solution to the differential equation Eqn. 2.19, which governs this process. The dashed line reproduces the Spectre simulation of this process from Fig. 2.19.

2.2.6 Measurement Results

Fabricated chips, received as bare dies, were wirebonded to open SOIC-type packages. Our layout allows for measuring the rail voltage across the **VDD/GND** pads and the oscillator/buffer output through the **output** pad.

2.2.6.1 Rail Voltage Measurements

The measured voltage level across the energy storage capacitor, C_{st}, is about 300 mV under a red laser pointer beam (incident intensity$<$330 W/m^2) and 40 to 80 mV under a white LED flashlight. For the laser pointer incident intensity, the calculated current expected from both photodiode arrays would be $\sim$4 nA total, which yields a lower simulated rail voltage. We can explain the disparity by considering multiple passes during which the light is absorbed, resulting in a higher photocurrent than expected. The incident light, after a single pass, is reflected back up from the metal layers of the photodiode tier and lower tiers to pass through the photosensitive layer again. Steady voltages around 100 mV are observed in a sunlit room; exposure to brighter daylight yields higher voltages comparable to or exceeding the laser pointer results.

2.2.6.2 Oscillator Output Measurements

To observe the self-powered circuit operation, we used a test setup consisting of a high-input-impedance, low-input-capacitance (1 pF) instrumentation amplifier and an oscilloscope. We found the oscillator in this 3DIC to oscillate under illumi-

nation. Fig. 2.26 displays a scope screen capture of the oscillator operating at 1.8 MHz, taken at the amplifier output. The 3DIC is under a laser pointer beam. The observed peak amplitude varies between 70-130 mV, with the bottom level at ∼39 mV. Taking the amplifier gain into account, this indicates a 3DIC buffer output amplitude of the order of 4.3 mV, consistent with simulation results for a 1 pF load. The variation in the output amplitude is due to variation in the light source.

2.2.7 Concluding Remarks on the First Self-Powering SOI 3DIC

We have demonstrated the operation of a self-powered three-dimensional SOI integrated circuit. To our best knowledge, this is the first self-powered SOI fully

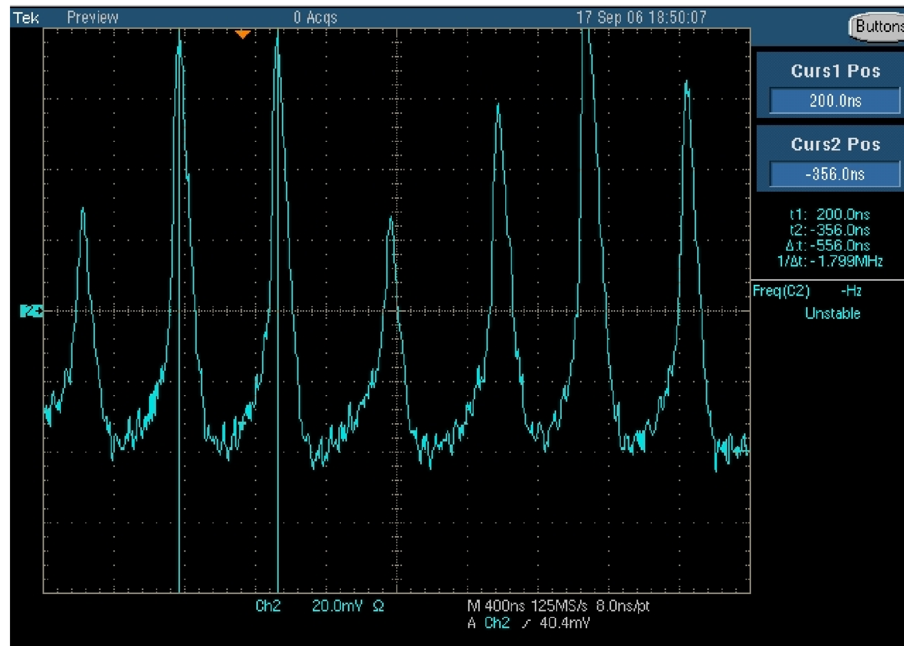


Figure 2.26: Amplifier output with 3DIC under red laser pointer light. The maximum signal amplitude is 130 mV, and dominant frequency is 1.8 MHz. The bottom signal level is at ∼39 mV.

integrated 3-D circuit. Our analysis points out to the self-governing property of this circuit, setting its own point of operation depending on the illumination level. We might make use of this property for a sensor design. The current 3DIC has further stand-alone applications as a built-in system clock, or as a local oscillator for a wireless network 3DIC node.

Deviations in the circuit behavior from design are mainly favorable and explainable by the physical realities of the chip, e.g. internal reflections causing higher absorption. It may be possible to take advantage of this effect by building photodiodes on multiple layers, with careful layout of the metal layers. It may also be possible to use post-processing to create features such as a diffraction grating overlay to increase the absorption depth for more efficient power harvesting. Unfortunately, a true resonator-like structure for the incident light to be trapped for many multiple passes through the active region is not easily designed without blocking the input of incident light to the system initially.

2.3 Second-Generation Design

2.3.1 Introduction

After the completed integrated circuits for the fabrication run 3DL1 utilized for the system described in the previous section were received, Lincoln Labs announced a second open run for a modified version of this process. The new process, termed 3DM2, featured the same three-tier arrangement and stacking method, as well as essentially the same basic SOI CMOS technology. However, there were a number of

differences in the process features, which we exploited in the course of a new design.

2.3.2 Overview of of the Second-Generation Chip

The main differences between the design for the 3DL1 and the 3DM2 run are presented briefly below, after which we give some detail.

- *Design Area* In this new run, the UMCP group was assigned a design area of 1 mm by 1 mm. This allowed us to submit effectively four variant designs placed on quarters of the chip, with each design four times the surface area of the 3DL1 chip.

- *New Diode Arrays* Two main aspects of diode design changed. The first is that the 3DM2 SOI process features a no-silicide layer, enabling the designers to form pn-junctions without the need to use a protective polysilicon layer covering the junction to prevent shorting from silicidation. The second was due to the consideration of photodiodes forward-biasing themselves in the case of high photocurrent output and high rail voltage. To handle this problem, we laid out some new diode arrays which consist of parallel branches, each branch featuring two diodes in series.

- *New Self-Powered Amplifier* The 3DL1 design features the self-powering of a local oscillator. In the 3DM2 design, we included a low-power amplifier to demonstrate the operation of a self-powered analog circuit.

- *Integrated Externally-Powered Output Buffer* For ease of testing without the

need for the circuit to drive high currents off-chip, we have included an amplifier stage with the necessary bonding pads to power it externally.

In summary, the new layout features four quarters, each self-powered by its own independent photodiode array. Figure 2.27 presents the floorplan of this design at a glance.

Q1 Single-photodiode array branches Local osc. + buffer	**Q2** Double-photodiode array branches Local osc. + buffer
Q3 Single-photodiode array branches Local osc. + buffer +external-powered amplifier	**Q4** Single-photodiode array branches Self-powered amplifier

Figure 2.27: An at-a-glance floorplan of the 3DM2 chip.

2.3.3 New Diode Arrays

There are two significant changes in the way the photodiode arrays are designed and constructed. The first one, incorporation of the now-available no-silicide layer, affects all diodes. The second one is the array comprised of two-diodes-in-series branches, implemented in one of the quarters.

2.3.3.1 Diode Layouts with No-Silicide

The 3DL1 process features a compulsory silicidation step after the polysilicon layer definition to improve the conductivity of the exposed active and polysilicon layers. While this step is essential for quality transistors, its effect of shorting junctions at the surface makes the design of intentional diodes impossible unless care is taken to shield the actual junction region from silicidation [47]. In the 3DL1 design, we shielded our diodes' junctions with appropriately-placed layout in the polysilicon layer. In the 3DM2 design, a no-silicide (NOSLC) layer is added to the process steps [55]. The use of this layer is explicitly recommended for diode formation, with the additional requirement of using active layers with no sidewall implants [56].

We revised our diode layouts, both the lateral (pin) and the annular (pn) kinds, incorporating the NOSLC layer. Figure 2.28 shows the layout. This change does not affect the layout dimensions. However, as referring back to Figure 2.5 should demonstrate, it should be beneficial to the incident power/photocurrent power conversion efficiency: The portion of the incident light going through the active layer unobserved during the first pass will not be absorbed by a polysilicon layer on the way down or back up, and thus secondary-pass absorption is likely to yield more extra power. Another advantage is that there is no more danger of the polysilicon layer overlaying the junction getting charged and altering the junction characteristics.

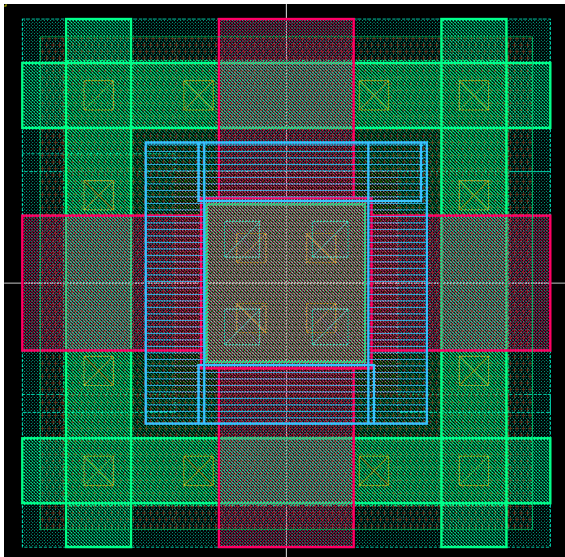


Figure 2.28: The no-silicide version of the CBN/CBP annular diode. The dimensions are the same as the diode in Figure 2.11.

2.3.3.2 Diode Array with Two Serial Diodes per Branch

In Section 2.2.5, we described the case where the photocurrent forward-biasing the diodes themselves by charging the storage capacitor can limit the achievable rail voltage. The rail voltage dropping across two diodes in series instead of a single diode as forward bias can ameliorate this problem. Thus on one quarter of the new chip, we have pn- and pin-diode arrays with two photodiodes in series per branch.

While it is straightforward to layout the pin-junction array comprised of lateral diodes in this fashion, the annular pn-junction diode pair layout requires some care so that the final two-dimensional array covering the top tier is easy to construct. Figure 2.29 displays the layout one branch of this array. Figure 2.30 displays a quartet of branches, which is the basic building block in the two-dimensional full-area layout.

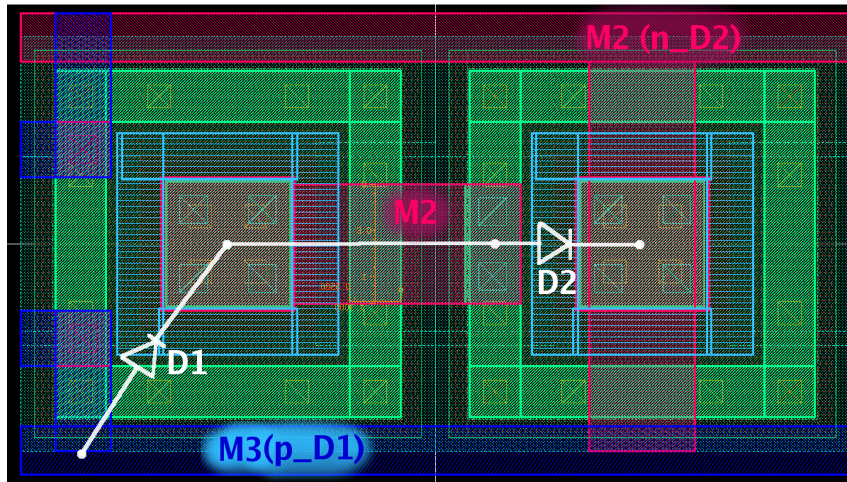


Figure 2.29: The layout for one branch of the two-serial-diodes-per-branch photodiode array. The overlay shows how the diodes are connected by the metal-2 bridge in the center.

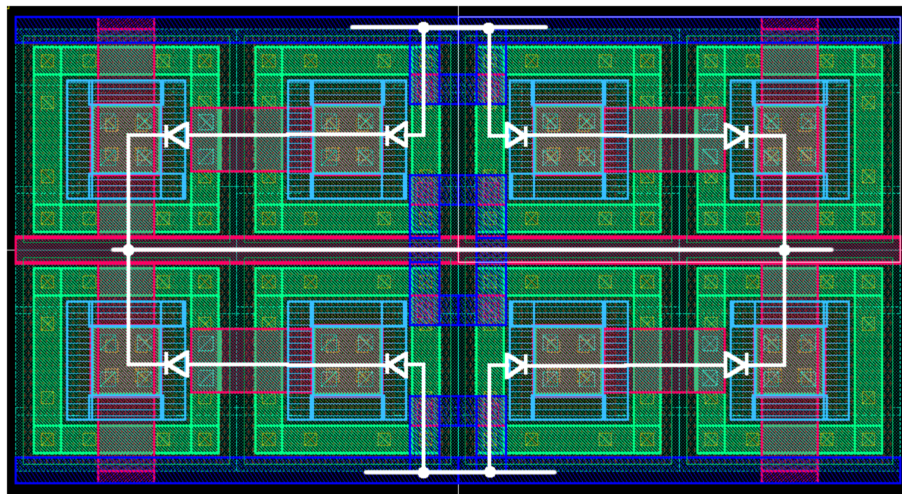


Figure 2.30: The layout for a quartet of branches for the two-serial-diodes-per-branch photodiode array. The layout was constructed by creating a mosaic out of these quartets.

2.3.4 Integrated Externally-Powered Amplifier

To eliminate the need for an external instrumentation amplifier to observe the oscillator operation, in collaboration with Ms. Bo Yang we designed an amplifier with extra bonding pads made available to power this amplifier externally. The amplifier is comprised of a differential stage, followed by two common source stages. The schematic for the circuit is presented in Figure 2.31. To demonstrate the benefit due to this amplifier, Figure 2.32 displays the simulation results of the self-powered oscillator/buffer directly driving a 1 pF off-chip load and the externally-powered amplifier driving the same load with the self-powered oscillator/buffer output as its input. The layout of the amplifier is given in Figure 2.33. For visual clarity, the full layout of the resistor is not shown.

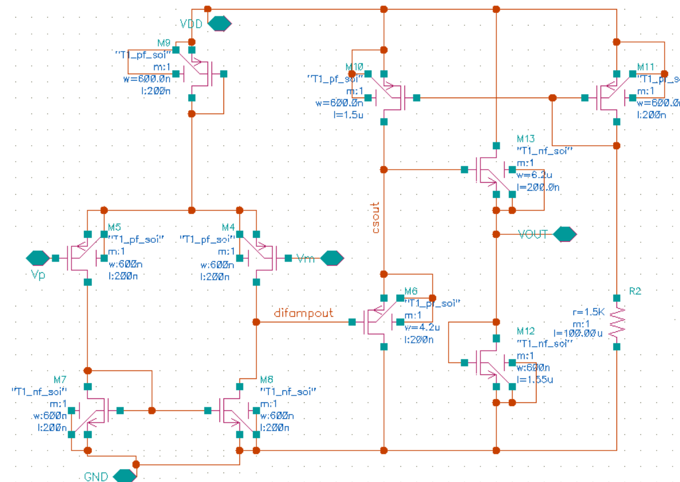


Figure 2.31: The externally-powered amplifier used in the third quarter of the 3DM2 chip.

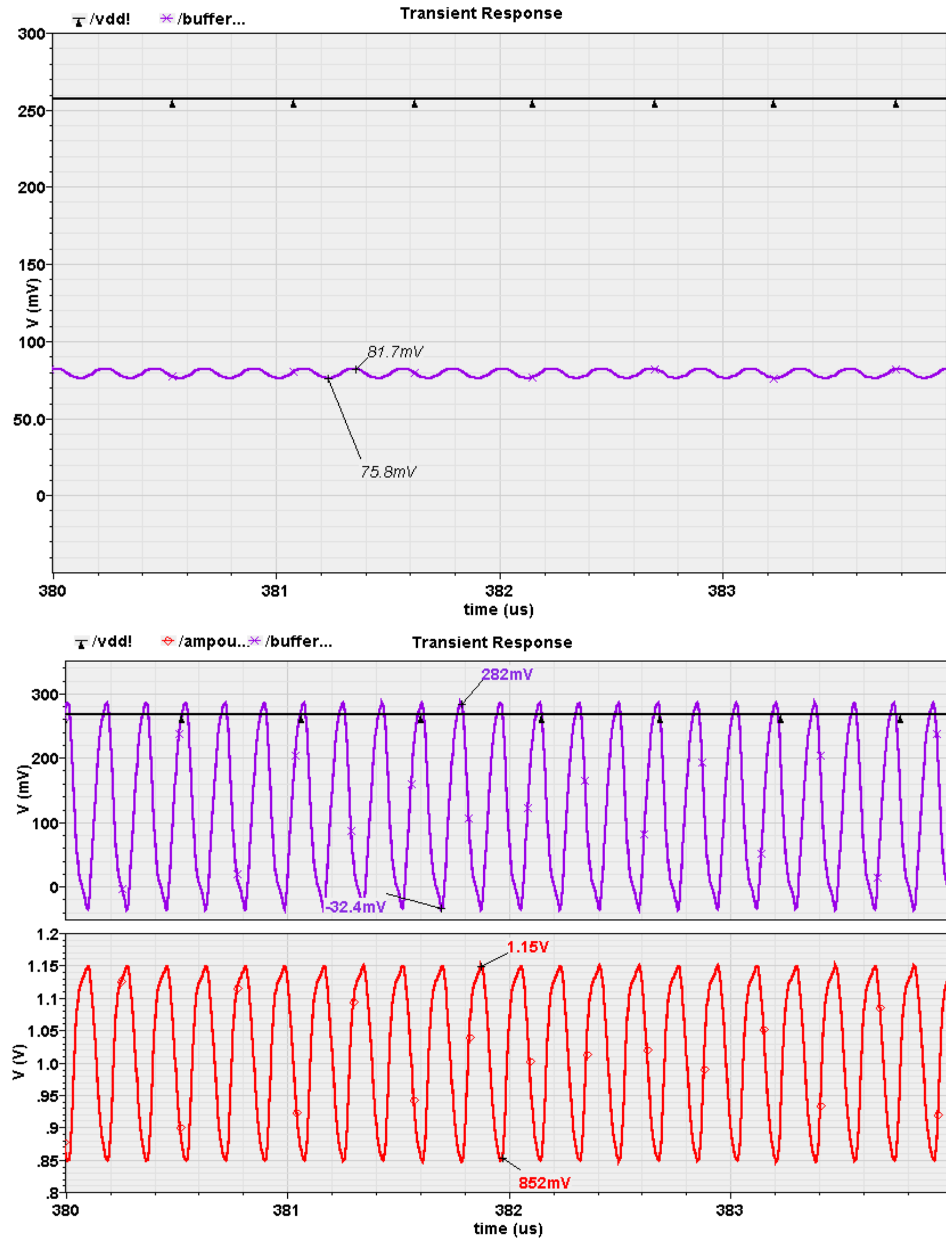


Figure 2.32: Top: The self-powered oscillator and buffer driving a 1 pF off-chip load. Bottom, first graph: The output of the self-powered oscillator and buffer driving the externally-powered built-in amplifier. Bottom, second graph: The output of the externally-powered built-in amplifier, with the above input, driving a 1pF off-chip load.

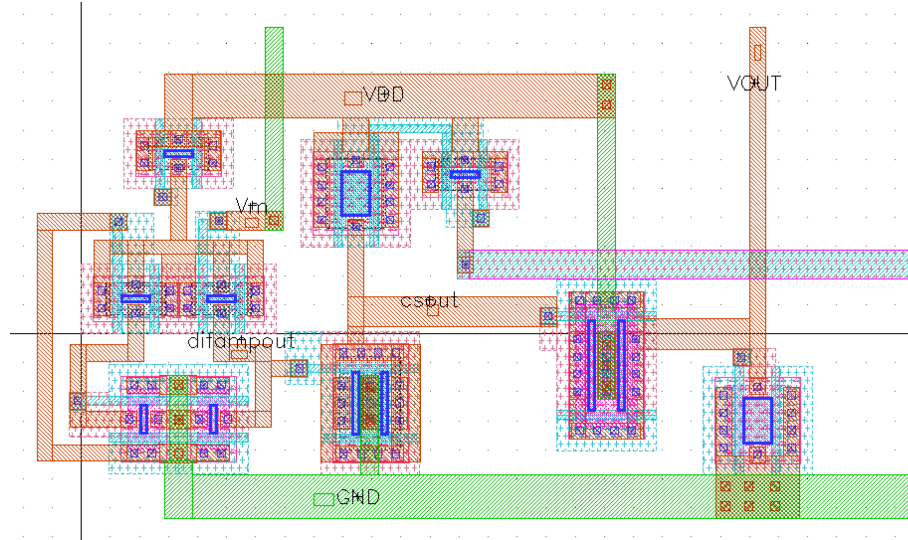


Figure 2.33: Layout of the externally-powered amplifier used in the third quarter of the 3DM2 chip.

2.3.5 Self-Powered Amplifier

To demonstrate self-powering of a second circuit other than the local oscillator/buffer featured in the design until then, in collaboration with Ms. Bo Yang we designed and laid out a small amplifier capable of operating with a bias current provided by the photodiodes. In the layout, this amplifier was followed by an externally-powered level-shifter and output amplifier to drive off-chip loads. Here we present the schematic of the self-powered section of this circuit, followed by the layout of the self-powered amplifier, level shifter and externally-powered output driver, in Figures 2.34 and 2.35 respectively.

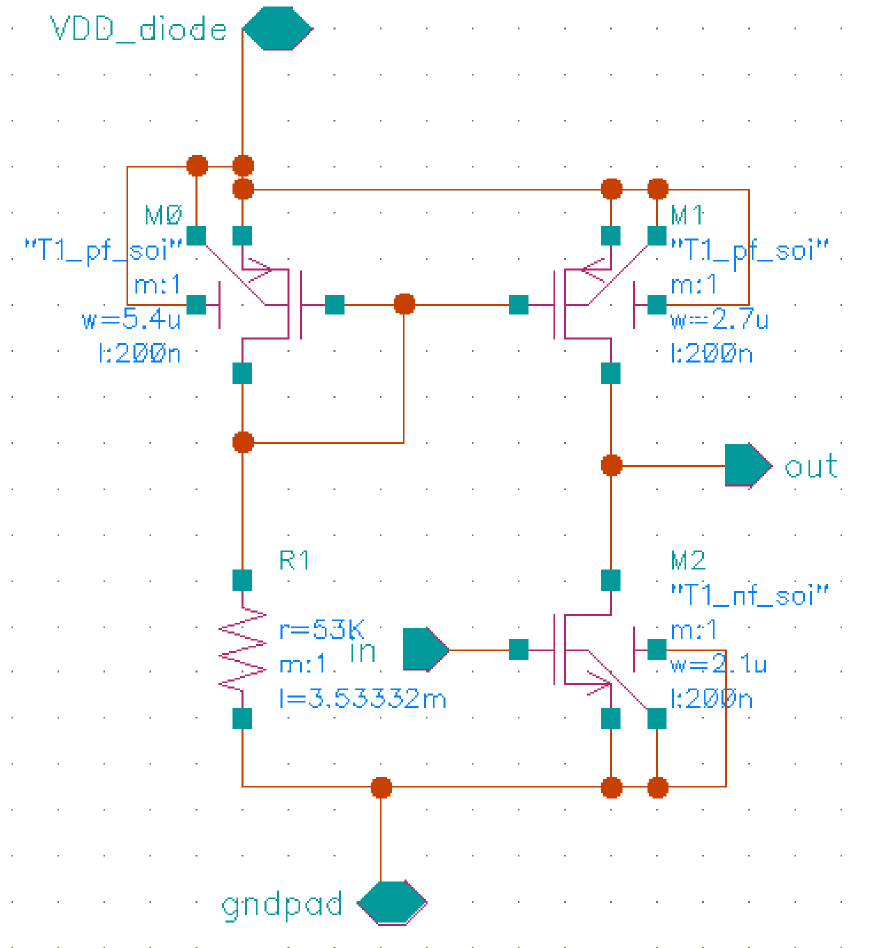
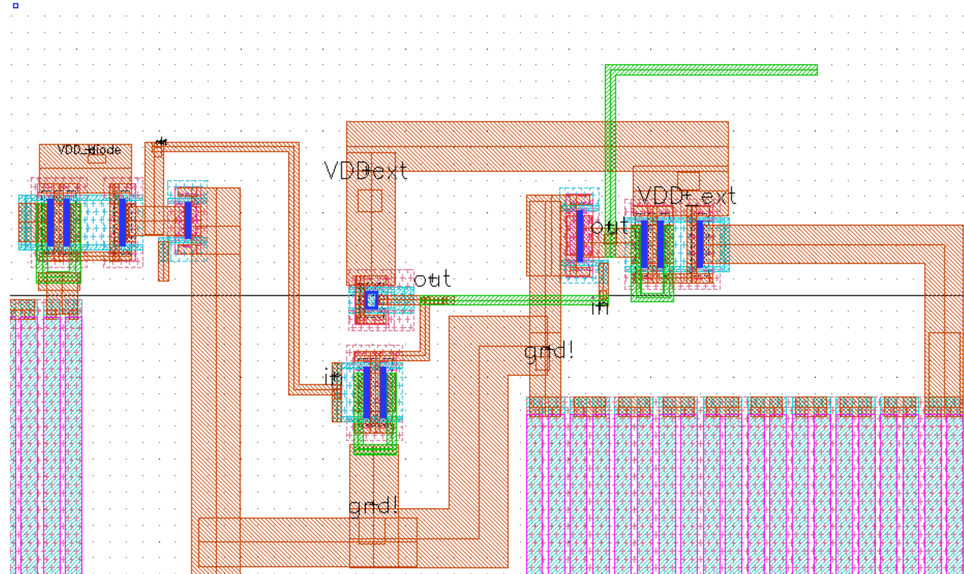


Figure 2.34: Schematic of the self-powered amplifier implemented in the fourth quarter of the 3DM2 chip. The amplifier is a simple common-source amplifier, biased by a current mirror which is fed by the photodiode arrays.



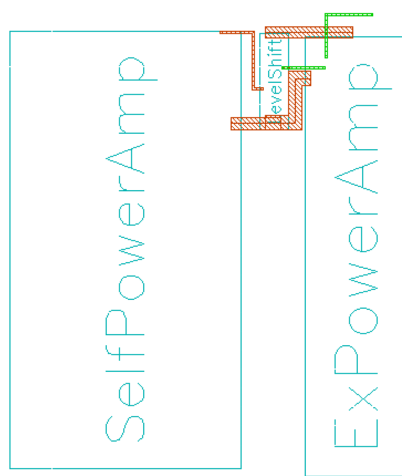


Figure 2.35: Top: The full layout for the self-powered amplifier, and the externally-powered level shifter and output amplifier to drive off-chip load. Full resistor layouts are once more cropped out for visual clarity. Left: a schematic representation of the components in the above layout.

2.4 3-D Integration: Current Research, Challenges and Directions

The basic idea behind three-dimensional integration is to increase integration density by taking advantage of the third dimension, as opposed to packing subsystems increasingly tightly onto a sprawl on a planar substrate. This is achieved through some way of stacking microelectronic circuits and forming the necessary connections vertically.

3-D integration was first considered in early 1980s. A technique for recrystallizing polysilicon was developed, enabling the construction of devices placed above each other within the same dielectric medium over a common substrate [1, 2, 9]. Two different approaches to 3-D integration emerged as the research progressed: *Sequential fabrication*, for which the devices and structures on consecutive stack levels have to be fabricated after the previous level fabrication has been completed, and *parallel fabrication*, where each tier of the final stack is fabricated individually and three-dimensional integration completes the fabrication of the 3DIC [61].

We should note that along with three-dimensional structures from either layers of devices fabricated tier-by-tier on a single die, or from bare (individual) dies or wafers stacked on top of each other with vertical connections, there has also been research on creating three-dimensional structures from packaged die and on special package designs to enable the same [1, 3]. In the scope of this work, we will not cover such integration efforts at the package level.

2.4.1 Sequential Fabrication Techniques for 3DI

By the end of the 1980s, research in creating 3-D integrated circuits had mainly focused on SOI-type technologies. As forming wide areas of good quality single-crystal silicon over amorphous insulators was still a problem during this period, most of the vertically-integrated circuits were of the order of a few transistors. A proposed structure, achieved by selective epitaxial growth and epitaxial lateral overgrowth, was a CMOS inverter with the PMOS stacked on top of the NMOS [2], presented here in Figure 2.36. Further methods for creating monocrystalline silicon for device formation over already-fabricated devices were laser beam recrystallization and metal-induced lateral crystallization.

Such structures require good planarity of the new epitaxial growth over a polysilicon gate, obtained by either chemical-mechanical polishing or restricting the vertical growth through particular techniques. However, thermal restrictions during processing are more critical. Whatever technique is used to deposit or grow new

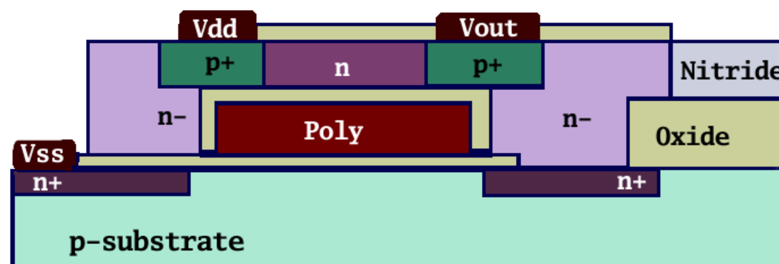


Figure 2.36: A 3-D inverter formed by fabricating a PMOS over an NMOS using epitaxial lateral overgrowth (ELO) techniques (adapted from [2]). The n layer which forms the channel of the PMOS is grown using ELO. The poly gate is common to both transistors.

silicon over the initial devices in the first substrate needs to be a relatively low-temperature so as not to damage the devices already formed. This problem stunted further development in this direction [61].

2.4.2 Parallel Fabrication Techniques for 3DI

Compared to serial techniques, the approach of three-dimensionally integrating die or wafers that have already been completely fabricated individually has been far more successful, with a great variety of integration styles and applications presented. Two broad variants of the approach can be defined: Chip stacks which use peripheral connections and which use through-wafer (or through-die) connections.

2.4.2.1 Chip Stacks with Peripheral Connections

3-D stacks can be formed by designing the interconnect and power networks on the individual-layer chips with the connection points routed to the chip edges. Several techniques can then be used to form the connections between layers and to the outside world, by utilizing the sidewall area of the newly-formed stack.

It is possible to achieve this with techniques such as tape-automated bonding or soldering the conductors on the edge [3]. Another approach is to form interconnects on the side of the stack directly, by patterning an interconnect network on the sidewall through photolithography and sputtering or laser trimming [62, 63, 64]. Figure 2.37 illustrates this approach. An extra substrate dedicated to interconnections might be soldered or otherwise attached to the stack side or top face [65].

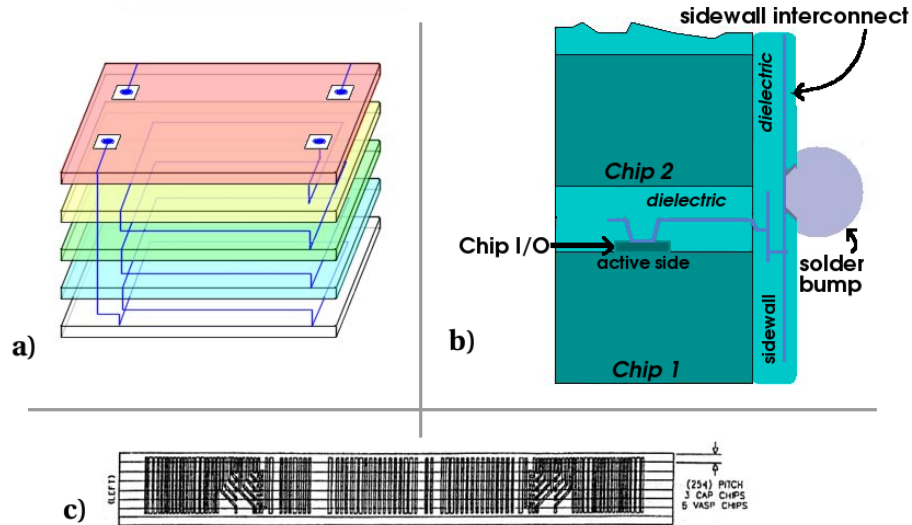


Figure 2.37: **a)** A schematic representation of a stack of integrated circuits connected to each other through interconnects extended to the edges and routed on the sidewall. **b)** A cross-section of a stack of two chips connected vertically through a T-connection formed on the sidewall, with outside connections enabled by a solder bump, adapted from [64]. **c)** The schematic representation of the sidewall connections for a stack of VLSI associative string processor chips, from [62].

Among the circuit and system applications proposed and fabricated using these approaches are microcameras, biomedical systems, sensors [3], massively parallel processors [62], high-density (possibly DRAM or SRAM) memory stacks and cubes [3, 64, 63] and artificial neural networks [65].

2.4.2.2 Chip Stacks with Through-Wafer and Through-Die Connections

The next approach to creating vertical connections between chips layered on each other is to form vias, similar in form and function to vias between different metal layers in a planar process, which go through the chip substrate and/or the supporting material between the chips.

Bulk Technology 3-D Integration through Vertical Vias

Starting from individually fabricated wafers, vias passing through the entire wafer can be created and filled with metal to connect a top-level wafer to aligned bonding pads on a lower-level wafer [3]. Figure 2.38 illustrates a stack formed in this manner. With this approach, it is possible to develop processes for wafer-to-wafer, die-to-wafer and die-to-die stacking. A starting step in the stacking process is face-to-face bonding of the lowest tier to the second tier, followed by via formation, and thinning the substrate of the upper tier in preparation to bonding the next tier up [66]. Different bonding methods between tiers are used, including oxide-to-oxide bonding, copper-to-copper bonding and utilizing a dielectric adhesive.

As an alternative to forming the vertical through-vias after the bonding process, it is also possible to create "buried interconnections," which delve a certain level into the substrate, before stacking. This way, when the substrate is thinned to the appropriate level to reveal the tips of the buried metal posts, a vertical connection to the next level can be formed [68]. In an example process demonstrated

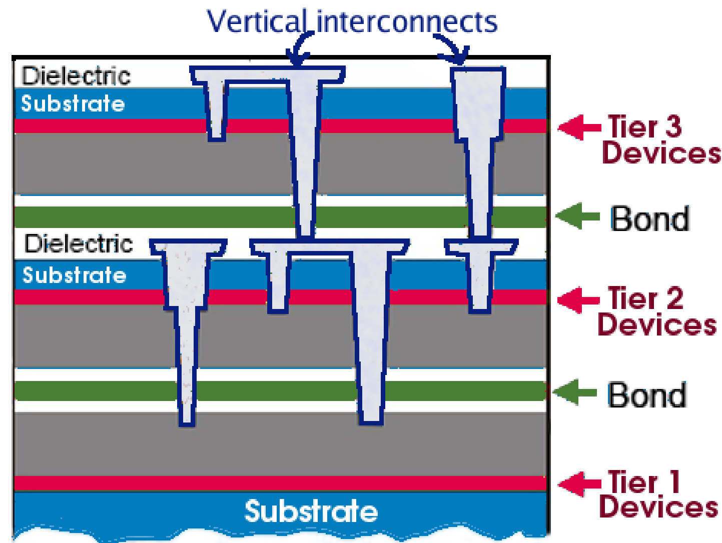


Figure 2.38: A schematic representation of wafer-level 3DI. The substrates of the upper tiers are thinned, but not removed entirely, and the process in question is a bulk process. The bottom tier has the active side facing up, while the higher tiers have the active devices facing down. (Adapted from [66].)

in 2006, the buried interconnects are created by etching deep trenches, lining them with an insulator thin film, and filling the vias with the conducting metal.

Among the circuit and system applications proposed and fabricated using these approaches are shared-memory chips and parallel processors.

SOI Technology 3-D Integration through Vertical Vias

It is also possible to vertically stack and interconnect individually-fabricated planar SOI technology ICs. In this case, the vertical vias can avoid the active device islands, delving entirely through oxide or similar insulating dielectric layers, with no need to line the via trenches with insulators. An example such process [69, 44] was

used in the production of the self-powering SOI CMOS system [45, 46] presented in the bulk of this Chapter; the process is specifically described in Section 2.2.2. Applications fabricated using this process include ring oscillators, visible imagers, and a 3-D laser radar (LADAR).

2.4.2.3 Alternate Vertical Connection Methods

Instead of bonding wafers or individual dies vertically to each other and forming direct electrical connections by means of vertical through-vias or peripheral connections, alternative methods of forming connected 3-D structures from ICs have been proposed. One such technique involves embedding thinned dies in a layer of benzocyclobutane (BCB) deposited over another (host) chip substrate, which also houses already-fabricated circuits [70]. Intra-chip connections are formed by patterned interconnection lines throughout the BCB.

Connections that are not directly conductive between vertically-stacked ICs have also been considered. The idea of optically coupled data links between the chips in an 3DIC was proposed in the 1980s [1]. Recently, a successful implementation of capacitively-coupled data and power in an SOI 3-D system has been reported [71].

2.4.3 An Analysis of State-of-the-Art in the Advantages and Problems of 3-D Integration

2.4.3.1 Gains in System Size

The first physical advantage of 3-D integration is packing density increase: Two stacked tiers effectively packs two transistors in the semiconductor footprint area of one. Considering the different interconnect routing requirements, a two-layer CMOS system does not quite achieve the full 50% area reduction intuitively expected over a planar implementation of the same system, but it comes close. The area gain falls with increasing layers, for instance to 75% over the planar circuit for four layers, but it is substantial nonetheless [73]. Lower area-per-die requirements are also expected to increase fabrication yield [66].

Considering the effect on packaging requirements, 3DI causes an advantage in terms of packaged system weight as well. 3DI systems can be 5-6 times smaller in volume than MCM approaches, which themselves are around 4 times smaller than discrete approaches. Likewise, the packaged system weight can be as much as 13 times lighter than an MCM implementation [1, 3].

Further, by analyzing the interconnect length distribution in multi-layer ICs by extending established stochastic analysis methods for 2D systems into 3D, a decrease in the average wire length is found. The decrease can be around 21% for a three-layer system as compared to a planar system, assuming both implementations are of well-partitioned designs, and can even rise to 47% if interconnections between any

two layers are enabled [8]. The length distribution of horizontal wires falls uniformly as more layers are stacked. While subject to some exceptions, the average length of vertical wires also falls [9]. This reduction has implications in speed and power consumption which we will focus on later.

2.4.3.2 Enabling Higher System Complexity

Another effect of 3DI on the interconnect network is a sharp increase in accessibility. Depending on the maximum trace length and stack thickness, a design unit in a 3D configuration may have access to many more neighbors within the same interconnect distance, as illustrated schematically in Figure 2.39. Intuitively this is reasonable, as accessibility now depends on the area of the unit rather than its peripheral length [3].

3D configurations are natural for parallel processing and parallel data transfer applications such as image sensors [1, 74]. It has been noted that in computing systems where multiple processors are trying to access the same memory, 3D technology

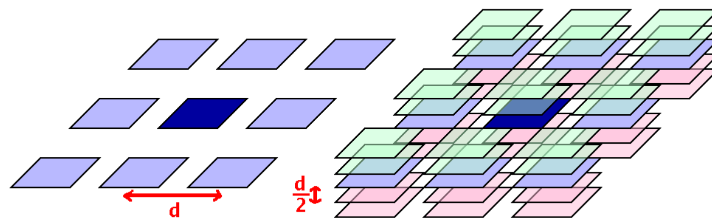


Figure 2.39: Using the same maximum interconnect linear length, a central unit in a three-dimensional configuration can access many more neighboring units, depending on the relative sizes of the central unit vs. layer-to-layer distance in a stack.

can help reduce memory contention blockages. Similarly, the ability to directly integrate both the first-level cache and the second-level cache with the microprocessor is afforded by 3DI, along with a larger area for the first-level cache [72].

3DI further enables higher system complexity by facilitating hybrid integration. An example is RF systems. While research is ongoing in RF electronics to migrate everything to CMOS, currently the best performance is still obtained by different technologies for different functions—SiGe BiCMOS for RF and Si CMOS for digital [66]. Similarly, by patterning thin-film microstrip lines over a dielectric layer over the semiconductor circuit substrate, 3-D MMICs have been developed [7]. As another example, systems using silicon thin film image sensors vertically integrated with a reader chip have been demonstrated [6], yielding nearly 100% fill factor. Also in image sensing, where a different material is optimal for the target wavelength range, vertical integration between the silicon readout circuit and the sensor plane of photosensitive material is desirable; an example is focal plane arrays of HgCdTe detectors for short- and medium- to long-wave infrared imaging [67].

2.4.3.3 Gains in System Speed and Power Consumption

The first implication of shorter average interconnect lengths of 3DICs is lower delays in signal propagation between functional blocks [3]. This is thanks to lower parasitic capacitive and inductive loads[1], and the ability to use more compact networks with shorter average delay in a 3-D layout instead of widespread, long-delay nets [9]. In microprocessor-type systems with the processor, memory and cache

[1]An example study of this effect we performed is presented in Section 5.6.

components are integrated in 3-D instead of placed in separate chips, significant bandwidth and access time improvements are possible, with reduced averages time per instruction execution [66, 72]. 3DICs can thus operate at higher clock speeds, although if the process cost limits the number of connections between tiers and the number of interconnect layers per tier, this improvement stalls after a certain number of tiers [8].

Reduction in interconnect length also contributes to power conservation in two ways: By decreased parasitic power consumption [3] and by requiring less power-consuming I/O units in general. 3-D systems need fewer buffers, or repeaters, while moving data between functional units in a complex system [9]. For the remaining buffers, circuits able to source less power will suffice [75]. Depending on the implementation technology, the power dissipation through interconnects can be reduced as much as 35% by just using two layers of integration [10].

2.4.3.4 Noise and Crosstalk Problems and Proposed Solutions

Certain kinds of digital noise, such as crosstalk between interconnects, simultaneous switching noise, reflection noise and susceptibility to electromagnetic interference, can be suppressed by using shorter interconnects. Therefore, 3-D integration may ameliorate problems of this kind if the designer takes advantage of its connectivity advantages [3]. Furthermore, with 3DI it is possible to integrate digital and analog subsystems in one "chip" without these subsystems sharing a substrate and thus being vulnerable to substrate noise coupling.

However, especially in technologies where the interconnect layers of one tier is close to the active layers of the next tier, the possibility of tier-to-tier crosstalk still exists [61, 76]. Electromagnetic modeling and experimental studies suggest that a grounded layer of metal between layers, designed at an optimum size as a ground plane, alleviate inter-tier coupling.

2.4.3.5 Thermal Problems and Proposed Solutions

A major potential disadvantage of a 3DIC is higher vulnerability to self-heating problems [106]. As will be covered in Chapter 4, in planar ICs most of the excess heat is dissipated through the substrate to the package towards the ambient. In 3-D stacks, the low thermal conductivity of the dielectric layers between the active layers insulates the devices near the middle of the stack from the ambient temperature, hindering thermal dissipation. This causes a variety of problems in both analog and digital circuits, with circuit performance degradation or even just unexpected changes in operation, such as operation frequency drift, higher leakage current rising with temperature and causing a positive feedback loop, and matching problems when two devices, designed to operate identically, are not at the same temperature.

The inter-tier dielectrics, whether used for stack bonding or just as the insulator for the interconnect layers, are the greatest problem. The interconnects themselves, with their higher thermal conductivity, can be an asset in carrying away excess heat. But the directionality of the low-thermal-resistivity paths is significant; therefore vertical heat conductors perform better [61]. The use of vertical

vias to carry out excess heat build-up in the middle layers has also been suggested elsewhere [10, 56], although both approaches increase layout area and complexity.

2.4.3.6 Layout Issues

Layout tools with features incorporated to handle the extra requirements of 3DIC design have been developed. Depending on the particular 3DIC process, such tools need to account for the existence of multiple tiers of devices, connectivity relationships between tiers as well as between layers, extra layout design rules for the 3-D vias and similar structures, and the alignment requirements for inter-tier connections [55, 77].

3DICs might also require enhancements in the automatic routing tools now in existence to make optimum use of the increased accessibility and shorter-average-length nets they offer. There has already been studies on the development of new interconnect network hierarchies, taking advantage of the chance in 3DICs to implement higher-level connections between subunits of a system in a reasonable area with shorter interconnects [78].

2.4.3.7 New Process Requirements

Three-dimensional integration requires specialized processes. Beyond the extra stacking and bonding steps, which contribute to the process cost, 3DI techniques have tighter requirements than the individual tier fabrication processes. Some extra process requirements of 3DI techniques are:

- Good planarization, of both the covering oxide and thinned handle substrates and similar, for successful tier stacking and bonding [1, 55];

- Formation of high-aspect ratio via holes and vias with good conductive qualities between tiers for through-hole 3DI techniques [1, 69];

- Wafer-to-wafer alignment with both lateral and rotational precision requirements, defined by the fabrication process [44, 66];

- A reliable wafer-to-wafer, die-to-wafer or die-to-die bonding technique, at low-enough temperatures so as not to damage the already-fabricated tiers [66, 44].

This increased process complexity does raise the fabrication intricacy and cost. However, these raises are somewhat balanced by the advantages of 3DI such as smaller system area, higher potential yield, the higher potential complexity of its system-on-a-chip type applications which could be contained in a single package, and lower power consumption.

2.5 Recent Progress in Self-Powering Methods

Developments in low-power electronics on one hand and ideally-autonomous wireless sensor nodes on the other [57, 58, 59] drives investigations into self-powering systems. Among the approaches considered are long-life, small-size power sources and on-chip systems harvesting power from the ambient. Here we present a brief overview of the approaches to this question.

2.5.1 Photoelectric Methods

Since photodiodes are readily integrable into semiconductor circuits, it is in a sense natural to pursue the idea of integrated circuits powered by on-chip optical power harvesters [45, 46]. One question that must be addressed here is the means for high-energy density storage, which is solvable by integrated capacitors [82]. Another is making sure that the photodiodes receive enough incident optical intensity and are efficient converters. Besides attempts to use the photogenerated current, converted into stored energy, to supply applications designed using conventional circuit technology and applications, there has also been suggestions for a new circuit architecture that depends on either optical signals or optical power to form novel analog or digital building blocks [60].

2.5.2 Piezoelectric, Vibrational and Thermoelectric Methods

It has been suggested that MEMS fabrication technology comprises a suitable platform for harvesting vibrational energy and converting it to electrical energy by means of piezoelectric materials and structures. One demonstrated approach uses silicon beams covered by a piezoelectric material thin film, whose AC voltage output is then rectified and converted to a DC level by means of on-chip rectifier and voltage regulator circuits [83]. Another example incorporates a cantilever-system, which uses not only ambient vibrations, but also the kinetic energy from β-particle emissions of a thin film of ^{63}Ni pad integrated under the cantilever tip [84]. Another method is to take advantage of the inertia of a partially-mobile mass in the structure

to power a microgenerator [58].

Careful design to maintain a temperature difference between two sides of a silicon-based structure has been demonstrated to generate power in the microwatts through the Seebeck effect. Such thermoelectric generation may be sufficient for applications designed for low power consumption [81]. Even a commercial wristwatch utilizing this form of harvested energy has been introduced [58].

2.5.3 Rectifying Antennas

Wireless microwave power transmission was the initial drive behind the development of rectifying antenna, or rectenna, systems. These systems were often designed to operate in the 2.4-2.45 GHz region, which is in the center of a industrial/scientific/medical band and not sharply attenuated by the atmosphere. Systems to operate at higher frequencies, such as 5.8 GHz and 35 GHz, were also suggested [85, 86]. Over time, rectifying antenna systems were also investigated for harvesting ambient energy.

The typical rectifying antenna system consists of an antenna, whose output is given to a rectifier (which could be a pn-diode or transistor at low frequencies) connected to an output bypass/storage capacitor and a load resistor [87]. At high frequencies, the rectifying element needs to be able to switch rapidly, so Schottky diodes with short transit times are preferred [88]. A low turn-on voltage for the diode is also an advantage. At frequencies and circuit sizes where matching is necessary, the antenna design might proceed iteratively and depend on the diode plus the rest

of the circuit. A transformer might be used for impedance conversion, along with more conventional matching circuits [89].

To illustrate the operation principle, we designed and built a simple rectifying antenna, which works at the 465-467 MHZ region. Figure 2.40 shows photos and a schematic of this system, which consists of a simple walkie-talkie antenna, a Schottky diode such as a BAS40 from Infinion or an MBD101 from ON Semiconductor as a rectifier, and optional capacitors and resistors as storage elements and load between the output of this diode and ground. The Schottky diodes were chosen for low turn-on voltages, which translates to a low voltage drop across the diode. No matching circuits were included, thanks to the low frequency and small system size.

Figure 2.41 shows the voltage measured across the load of this system in two cases. In the first measurement, the RF source is operated for twenty seconds (starting at t=5 sec.) and charges a 2.2 mF capacitor, which then discharges across a 100 KΩ load when the power is turned off at t=25 seconds. The slope of the charging curve indicates that during the charge, the rectifying antenna is supplying an average of 52.5 μA to the capacitor. In the second measurement, the storage element is a 10 nF capacitor loaded with a red LED. The RF power is being turned on and off; the LED turns on and clamps the voltage at 1.725 V when the RF source is on. Since the storage capacitor is very small, the load voltage falls rapidly to zero when the RF power is turned off.

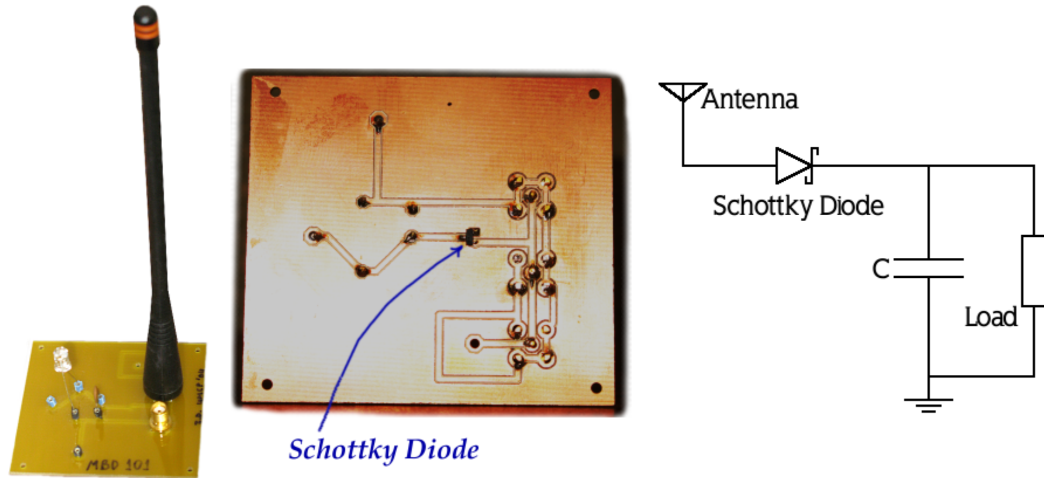


Figure 2.40: Photos of the rectifying antenna system: Left, the full board with an LED load; middle: close-up of the back side of the board with the Schottky diode. Right: Schematic of the circuit.

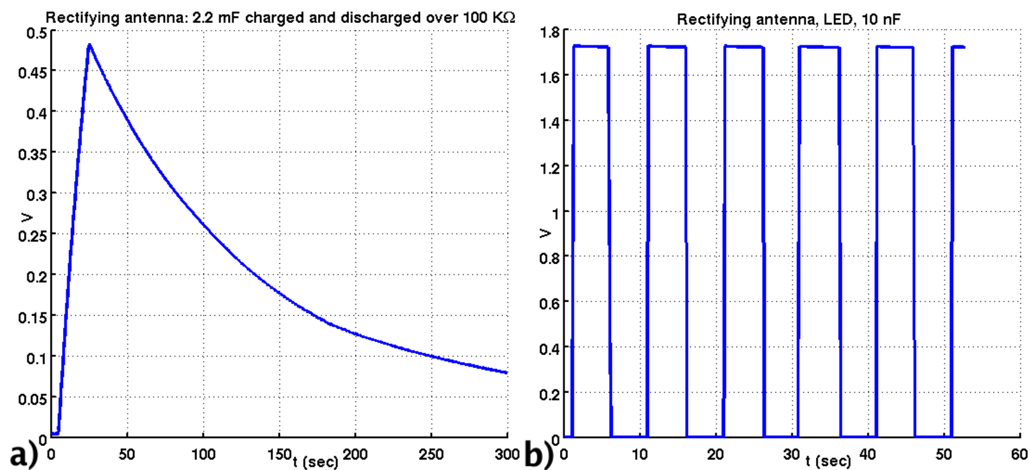


Figure 2.41: **a)** The rectifying antenna of Fig. 2.40 charging a 2.2 mF capacitor over twenty seconds, with ∼ 5mW of RF power provided at ∼467 MHz. **b)** The voltage drop across an LED load being suppled by the rectifying antenna as the RF power is turned on and off.

2.5.4 Small Batteries

Although not strictly a self-powering approach, this option to powering miniature systems has to be mentioned for completeness, and also because it is possible to design rechargeable miniature batteries in conjunction with one of the methods described above. Batteries can be capable of higher energy density than capacitors, and there has been strong interest in the development of small batteries in parallel with the advent of MEMS technology [90]. The ideal battery to power a smart dust scale system would be rechargable, small, capable of a large number of charge-discharge cycles, present low internal resistance and environment dependencies, and easy and cheap to produce. Research on such devices focuses on novel materials as electrodes and electrolytes and on form factors and packaging for small size. Several approaches, such as thin film batteries and flexible devices [91] have been demonstrated. Recently, here at the University of Maryland, we have participated in a collaboration to develop high energy density, small hybrid battery/capacitor structures [92] and adapt them for use in wireless systems.

2.6 Summary

Three-dimensional integration is a promising new approach to VLSI, with potential advantages over planar integration. In this Chapter, we described the physical design, analysis and testing of a self-contained three-dimensional integrated circuit. To our knowledge, this is the first demonstration of a self-powering SOI 3DIC using a photodiode array.

After presenting an overview of our design and the fabrication process, we related the design process of the photodiode arrays, which supply power to our circuit. We demonstrated that while the very small thickness of the SOI technology photoactive material is a handicap, we do have some extra gain from multiple reflections from the underlying layer structure and that the amount of photocurrent we expect to obtain is enough, by simulations, to run our functional circuit. We explained the self-regulating operation of the self-powering system qualitatively and analyzed it quantitatively. We also showed that for the expected photocurrent and resulting rail voltages, the diodes forward-biasing themselves would not be a crippling problem. We then presented the measurement results from the circuit operation.

We have designed a second-generation system for the next run of this 3-D process. The fabrication facility has informed us that at the time of writing, the individual tiers have been fabricated and the 3-D integration is underway. Here we presented our expanded design, which features modified diodes taking advantage of a new no-silicide layer, modified diode arrays featuring two diodes in series in each branch, an analog amplifier as a new functional unit, and integrated externally-powered output buffers for ease of testing.

Finally, we reviewed the extant literature on 3-D integration, presenting the state-of-the-art fabrication techniques, advantages and problems. We also presented a brief review of self-powering methods for small, autonomous applications such as smart dust sensor nodes.

Chapter 3

Physical Aspects of VLSI Systems: A Transient Spatially-Dependent Green's Function Approach to Modeling 3-D On-Chip Interconnect Networks

3.1 Introduction, Motivation and Background

Process scaling, the development of newer processes with more metal layers available to the designer, and novel approaches to hybrid and 3-D integration all raise the complexity of integrated circuit design. With all these contributing factors, the interconnect network on chip-scale systems emerges as a detailed electromagnetic system [13]. As it carries power and signals from one part of the circuit to the other, the interconnect network is vulnerable to variable delays, unintended noise coupling from different parts of the circuit and from a "dirty" power source, and from external electromagnetic sources.

We present a methodology for investigating the response of a complex on-chip interconnect network to external and internal noise, through the development of a 3-D solver based on creating a lumped element model of an interconnect network and solving for its impulse responses [96, 97]. Our method exhibits a lower computational cost than SPICE and allows the user the flexibility to work on a wide variety of interconnect network geometries as well as to include as many parasitic effects as

desired for the application.

While full-wave electromagnetic solutions have been used to model on-chip interconnect networks [93, 94, 95], these are computationally intensive for such structures on a semiconducting substrate. The wide variations in the significant dimensions such as layer thicknesses, conductivities and dielectric constants require the creation of a complex mesh with associated numerical problems. For a problem like determining which points on a chip are particularly vulnerable to noise coupling, the ability to repeat simulations on a certain network quickly, for many input distributions, would be advantageous. Our method provides this ability without having to solve for the entire system repeatedly, saving storage and operation counts.

This chapter opens with an outline of our methodology. We give the details of our numerical modeling technique by going over the network construction technique and the solver algorithms. We present the results of our simulations, along with some comments on numerical issues such as computational cost and convergence tests. The chapter concludes with our method applied to an example interconnect network and resultant commentary on layout design practices.

3.1.1 Methodology: Linear Time-Invariant Systems and Green's Function

We model on-chip interconnect networks as lumped element networks comprised of unit cells. Figure 3.1 gives possible unit cell "seed"s for a two-metal process. The methods to define equivalent circuits for these unit cells and their

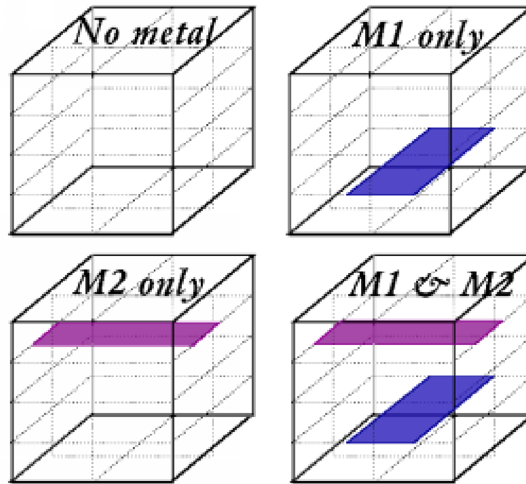


Figure 3.1: Unit cell examples for a two-metal process.

couplings will be detailed later.

Next, we set up a lumped-element network using our specific interconnect network layout and determine which output nodes are of interest. For instance, input transistor gate in a low-noise amplifier could be picked as a particular point of vulnerability. The responses to impulses induced at or injected into likely input points in the network are then calculated. At this step, the full system is solved. These impulse responses are all we need to get the output for any random signal distribution [96]. Repeating this process to obtain the responses to different random input distributions does not require solving for the response of the full network again.

This method has two main computational advantages. First, we need to use full-wave solutions only to obtain equivalent circuits for small unit cells if desired. For a preliminary-analysis type approach, a full-wave solution might prove unneeded and unit cells may simply be constructed by inspection.

Second, for selected network points of interest, we need the impulse responses at these points only to have been calculated and stored to get the response to a general input. Especially when exploring effects of random interference, this method has speed and storage advantages over repeating full lumped-network solutions for all possible input distributions.

3.2 Numerical Modeling

3.2.1 Theory

For notational simplicity, we present the mathematical background for a single spatial dimension, x. Consider a linear time-invariant system and let $h_i[x,t]$ be the system's time-dependent Green's function response at every point of the output domain to a unit impulse at point x_i and time t=0, $\delta[x-x_i]\delta[t]$:

$$\delta[x-x_i]\delta[t] \longrightarrow h_i[x,t]. \tag{3.1}$$

We can define an input function $f[x,t]$ (see Figure 3.2), whose value at point x_i over all time t is given by

$$f[x_i,t] = f[x,t]\delta[x-x_i]. \tag{3.2}$$

Sampling this function at time points t_j, we can write it as a summation of impulses marching over time, assuming the time points are sufficiently close and the sampling frequency sufficiently high:

$$f[x_i,t] = \sum_j f[x,t_j]\delta[x-x_i]\delta[t-t_j]. \tag{3.3}$$

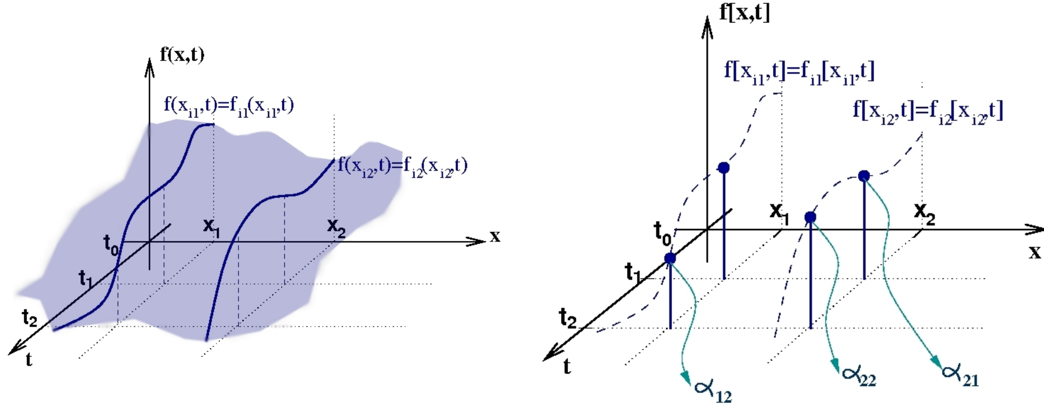


Figure 3.2: Left: A continuous time-and-space dependent function. Right: A discretized time-and-space dependent function, which can be written as the sum of its samples using the coefficients α_{ij}.

This will then cause the time-dependent system response $F_i[x,t]$ at all points in the system:

$$f[x_i,t]\longrightarrow F_i[x,t] = \sum_j f[x_i,t_j]h_i[x,t-t_j]\ . \tag{3.4}$$

Thus $F_i[x,t]$ is the contribution to the system response by an input applied at point x_i over the time points t_j. Figure 3.3 shows two such responses F_1 and F_2 to two inputs f_{i1} and f_{i2}, which are applied at times t_{j1} and t_{j2} respectively.

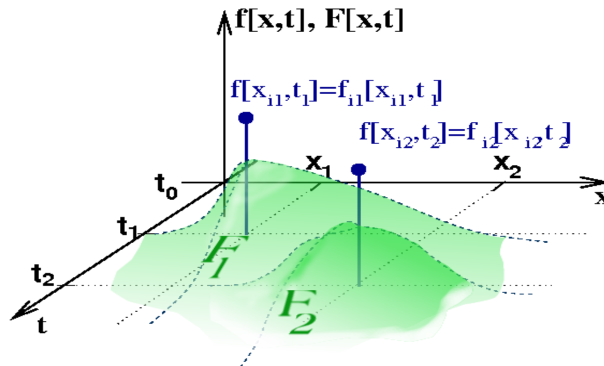


Figure 3.3: Time-dependent responses over all space to two individual impulse-type inputs applied at different x-points at different times t_j.

The principle of superposition [98] gives the response of a linear system to the full input $f[x,t]$ as the sum of all F_i:

$$F[x,t] = \sum_i \sum_j f[x_i, t_j] h_i[x, t - t_j] \tag{3.5}$$

Thus if the impulse responses $h_i[x_{out}, t]$ at point x_{out} to impulses at every possible input point are known, defining a space- and time-dependent random input distribution by the coefficients $\alpha_{i,j} := f[x_i, t_j]$ gives the system output at x_{out} by time-shifting and summation:

$$F[x_{out}, t] = \sum_i \sum_j \alpha_{ij} h_i[x_{out}, t - t_j] \tag{3.6}$$

This approach therefore allows us to test for the effects of random input distributions easily, allowing more flexibility than experimentation.

3.2.1.1 An Example Implementation using SPECTRE

Consider the network in Figure 3.4 with six nodes $\{x_1..x_6\}$. Assume we need the output at point 3, F_3, for a discrete-time input given by

$$\begin{aligned} f[x,t] = 2\delta[x - x_2, t] + 3\delta[x - x_2, t - 200\ ns] \\ +3\delta[x - x_5, t - 100\ ns] + \delta[x - x_5, t - 400\ ns] \ . \end{aligned} \tag{3.7}$$

If h_{3_i} is the response at x_3 to an impulse at node i, theoretically this response to the input given in Eqn. 3.7 is

$$\begin{aligned} F_3[t] = 2h_{3_2}[t] + 3h_{3_2}[t - 200\ ns] \\ +3h_{3_5}[t - 100\ ns] + h_{3_5}[t - 400\ ns] \ . \end{aligned} \tag{3.8}$$

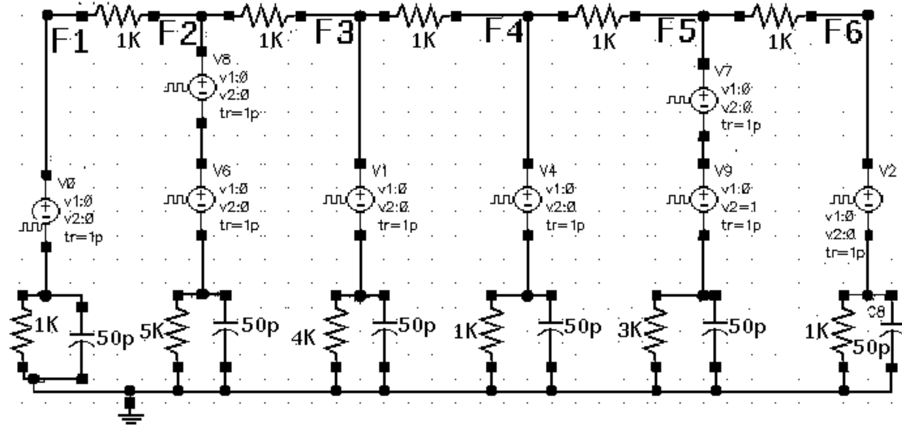


Figure 3.4: A linear network with six nodes defined.

The top half of Fig. 3.5 shows the impulse responses $h_{3_2}[t]$ and h_{3_5} as given by the circuit simulation program SPECTRE. The bottom half is the simulated response to the input defined by Eqn. 3.7.

We stored these response waveforms and implemented our convolution by using the calculator function of SPECTRE. Figure 3.6 shows the weighted, time-shifted sum of individual impulse responses (triangles) presented in Eqn. 3.8 superimposed upon the simulated result (circles) previously shown in Figure 3.5. The agreement is exact, validating the mathematical framework.

It should be re-emphasized that it was thus possible to reproduce the result of a full simulation simply by storing two time-dependent impulse responses calculated for node x_3 and performing arithmetical operations on these instead of running a full simulation again.

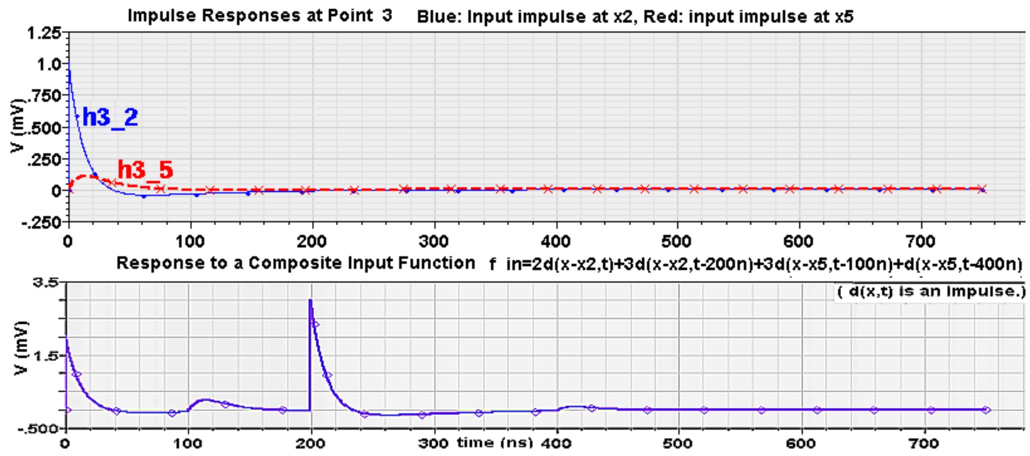


Figure 3.5: Top: Simulated impulse responses at node x_3 to impulses at x_2 and x_5. Bottom: Simulated output at node x_3 to the input of Eqn. 3.7.

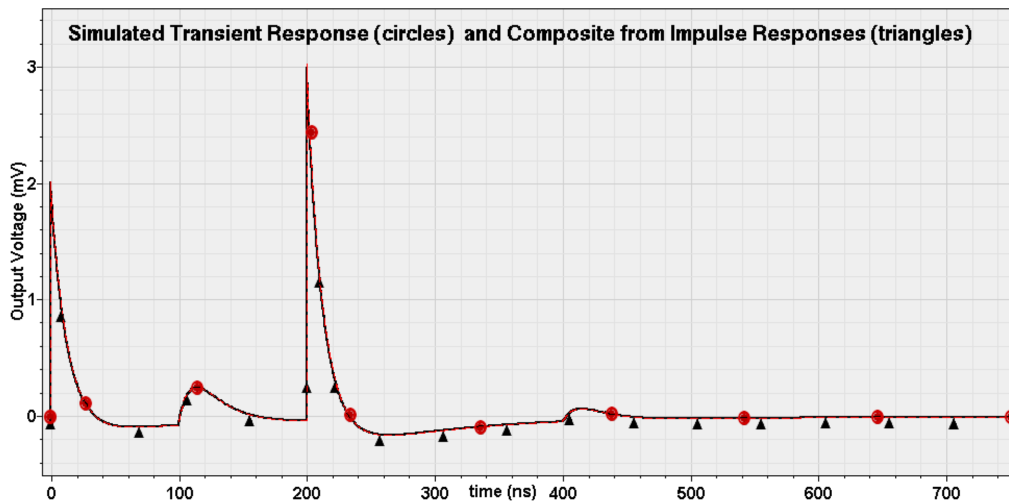


Figure 3.6: Response at node x_3 calculated using Eqn. 3.8 (black triangles) together with the simulated output (red circles), exact agreement.

3.2.2 Computational Cost

Once the impulse responses $h_i[x,t]$ have been calculated, to find the response to a time-dependent input $f_i[x,t]$, we proceed by adding the output contribution from each input time step:

$$V_{out}[t] \longleftarrow V_{out}[t] + f_i[t_n] \times h_i[t - t_n] \tag{3.9}$$

For an input applied at a single spatial point and lasting for t_{in} temporal points, this multiplication/addition is carried out $t_{in}t_h$ times, where t_h is the number of timesteps required for the impulse response due to that particular input point to decay. Assuming the input is spread across N_{in} spatial points and the longest it lasts at any of these points is for t_{in}, the maximum number of times the multiplication/addition operation needs to be made is $N_{in}t_{in}t_h$ times. Thus if $t_{in} << t_h$, the operation cost of the time-shifting/summation depends largely on t_h. These scalar-vector multiplications, vector shifting and summations of Eq. 3.6 are needed only once per input and output points of interest. Assuming there are N_{in} likely input points (e.g. points vulnerable to EM coupling or noise injection as from a power rail) and N_{out} important output points we are interested in, obtaining the entire solution thus requires $N_{in}N_{out}\mathcal{O}(t_h)$ operations.

Repeating the calculation for different random inputs does not require solving the response of the entire network again, whereas SPICE solves the entire network matrix equation at each timestep [99]. Its operation cost per timestep grows as N^m, where N is the number of mesh points and $m > 1$ depends on the matrix equation solution method. Typically, $N >> N_{in}$ or N_{out}, and the number of timesteps de-

pends on how fast the impulse responses decay, paralleling the $\mathcal{O}(t_h)$ component of the cost of our method.

3.3 Implementation

3.3.1 Interconnect Network Construction

Possible metal segment combinations in a small area define our unit cells. Various equivalent circuit models, including semiconducting substrate effects, for coupled interconnect segments have been proposed [19, 102, 100]. The requisite element values can be derived by parameter extraction from full-wave simulations or S-parameter measurements [21].

For the purposes of our application, the equivalent lumped element network needs to be linear and time-invariant. The common RLCG-type models for single and coupled transmission lines conform to this requirement.

Figure 3.7 shows some possible simplified unit cell equivalent circuit definitions in a two-metal technology.

To sum up, combining unit cells with ground connections and on-chip device loads according to the layout, we obtain a full lumped-element network. We can also define unit cells describing the interaction between stacked (3-D) chips' interconnect layers by treating inter-tier vias as analogous to inter-metal layer vias. While the cubic increase in the number of mesh points would limit repeated full analysis of such interconnect networks with SPICE, and although our impulse response calculations are bound by the same limit, once these have been completed finding the

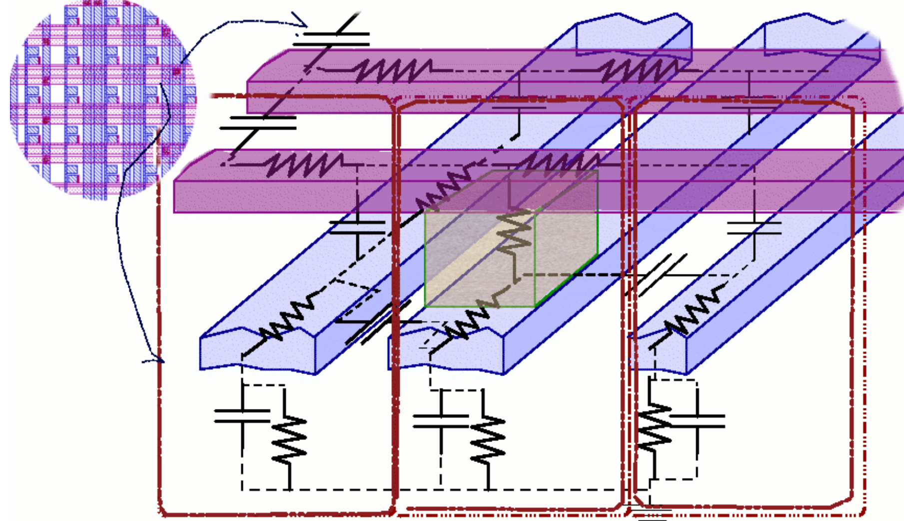


Figure 3.7: Two-metal network; a simplified equivalent RC network shown. Possible unit cells partitioned with dashed outlines. Note that where two metal lines on different layers overlap with a via, the connection between the mesh points on the overlap is resistive, whereas at locations without a via the same connection is capacitive.

full response of a 3-D system to random different inputs with our Green's Function based method is only slightly more complex or computationally intensive than the same operation on a 2-D system.

For use in our self-developed code suite, we have implemented a network-creation routine. This routine scans the solution domain mesh point by mesh point and writes, in an output file, the resistance and capacitance leaving in each direction for each mesh point in a matrix with one row per mesh point. Each columns of this matrix is reserved to denote a certain element connected to the mesh point of that row from a certain direction: For instance, for the 3-D version of this program, the first seven columns are x-mesh-index, y-mesh-index, z-mesh-index, resistance to the

"north," capacitance to the "south," resistance to the "east," capacitance to the "east," and so on. For the RC-network version of our mesh, Figure 3.8 illustrates all the elements that need to be defined per mesh point.

The shape of the metal interconnect network is pre-determined by the user and is reflected in the final network by whether there is a capacitive, resistive, or no connection between mesh points related to each other in a certain direction. The code uses a dummy element value at places to indicate a capacitive or resistive open circuit connected a mesh point in a given direction.

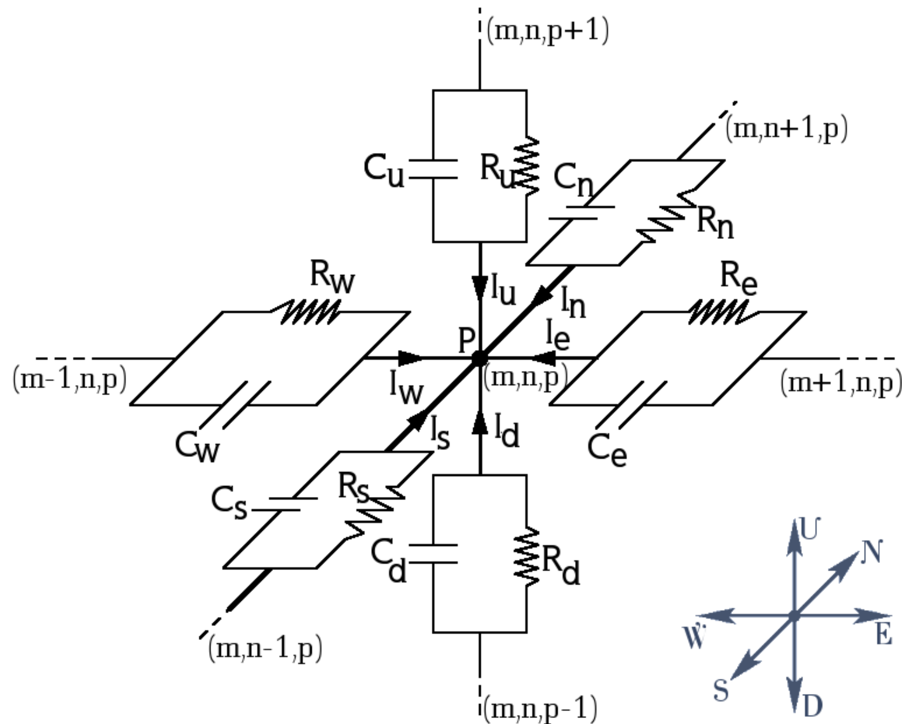


Figure 3.8: The element values that need to be defined per mesh point by the mesh creator code.

3.3.2 Impulse-Response Solver

3.3.2.1 KCL Network, Equations and Discretization

For a point in our mesh as depicted in Fig. 3.8, the sum of all currents entering point P ((x, y, z) indices (m, n, p)) is zero:

$$I_d + I_s + I_w + I_e + I_n + I_u = 0 \tag{3.10}$$

The current from each branch towards the node is the sum of the current through the resistor of the branch and the current through its capacitor; for example

$$I_n = I_{Rn} + I_{Cn} = \frac{V_{m,n+1,p} - V_{m,n,p}}{R_n} + C_n \frac{d(V_{m,n+1,p} - V_{m,n,p})}{dt} \tag{3.11}$$

Using a first-order discretization for the differential, this can be rewritten as

$$I_n^t = I_{Rn}^t + I_{Cn}^t = \frac{V_{m,n+1,p}^t - V_{m,n,p}^t}{R_n} + \frac{C_n}{\Delta t}[(V_{m,n+1,p}^t - V_{m,n,p}^t) - (V_{m,n+1,p}^{t-1} - V_{m,n,p}^{t-1})] \tag{3.12}$$

where t donates the current time step, $t-1$ the previous time step, and Δt the time step size. Writing Eqn. 3.12 for each branch and bringing them together in Eqn. 3.10 yields

$$\begin{aligned}
&\frac{V_{m,n,p-1}^t - V_{m,n,p}^t}{R_d} + \frac{C_d}{\Delta t}[(V_{m,n,p-1}^t - V_{m,n,p}^t) - (V_{m,n,p-1}^{t-1} - V_{m,n,p}^{t-1})] + \\
&\frac{V_{m,n-1,p}^t - V_{m,n,p}^t}{R_s} + \frac{C_s}{\Delta t}[(V_{m,n-1,p}^t - V_{m,n,p}^t) - (V_{m,n-1,p}^{t-1} - V_{m,n,p}^{t-1})] + \\
&\frac{V_{m-1,n,p}^t - V_{m,n,p}^t}{R_w} + \frac{C_w}{\Delta t}[(V_{m-1,n,p}^t - V_{m,n,p}^t) - (V_{m-1,n,p}^{t-1} - V_{m,n,p}^{t-1})] + \\
&\frac{V_{m+1,n,p}^t - V_{m,n,p}^t}{R_e} + \frac{C_e}{\Delta t}[(V_{m+1,n,p}^t - V_{m,n,p}^t) - (V_{m+1,n,p}^{t-1} - V_{m,n,p}^{t-1})] + \\
&\frac{V_{m,n+1,p}^t - V_{m,n,p}^t}{R_n} + \frac{C_n}{\Delta t}[(V_{m,n+1,p}^t - V_{m,n,p}^t) - (V_{m,n+1,p}^{t-1} - V_{m,n,p}^{t-1})] + \\
&\frac{V_{m,n,p+1}^t - V_{m,n,p}^t}{R_u} + \frac{C_u}{\Delta t}[(V_{m,n,p+1}^t - V_{m,n,p}^t) - (V_{m,n,p+1}^{t-1} - V_{m,n,p}^{t-1})] = 0
\end{aligned} \tag{3.13}$$

We need to solve this equation at each mesh point while progressing through time. Rearranging yields the KCL matrix equation to solve:

$$
\begin{aligned}
V^t_{m,n,p-1}[\frac{1}{R_d}+\frac{C_d}{\Delta t}]+V^t_{m,n-1,p}[\frac{1}{R_s}+\frac{C_s}{\Delta t}]+V^t_{m-1,n,p}[\frac{1}{R_w}+\frac{C_w}{\Delta t}]\\
+V^t_{m+1,n,p}[\frac{1}{R_e}+\frac{C_e}{\Delta t}]+V^t_{m,n+1,p}[\frac{1}{R_n}+\frac{C_n}{\Delta t}]+V^t_{m,n,p+1}[\frac{1}{R_u}+\frac{C_u}{\Delta t}]\\
-V^t_{m,n,p}[\sum_k\frac{1}{R_k}+\frac{C_k}{\Delta t}]\\
=V^{t-1}_{m,n,p-1}\frac{C_d}{\Delta t}+V^{t-1}_{m,n-1,p}\frac{C_s}{\Delta t}+V^{t-1}_{m-1,n,p}\frac{C_w}{\Delta t}\\
+V^{t-1}_{m+1,n,p}\frac{C_e}{\Delta t}+V^{t-1}_{m,n+1,p}\frac{C_n}{\Delta t}+V^{t-1}_{m,n,p+1}\frac{C_u}{\Delta t}-V^{t-1}_{m,n,p}[\sum_k\frac{C_k}{\Delta t}]
\end{aligned}
\tag{3.14}
$$

If the mesh has $N_X \times N_Y \times N_Z$ points, this yields a matrix equation $\mathbf{A}\vec{V}^t = \mathbf{B}\vec{V}^{t-1}+\vec{b}$, where $\mathbf{A}$ and $\mathbf{B}$ are matrices of size $N_X N_Y N_Z \times N_X N_Y N_Z$. They are sparse, with 7 nonzero terms in each row, and banded thanks to our indexing scheme. $\vec{b}$ is a source term which modifies Eqns. 3.13 and 3.14 at the time and point when and where the input impulse is applied to the system.

The only boundary condition present in the system is the ground node being held at zero potential, which is implemented by an extra row and an extra column added to the left-hand matrix $\mathbf{A}$ and a zero value set as the last element of the right-hand vector. The elements in the last-column value of each row of $\mathbf{A}$ are derived from the impedance values between each mesh point and ground.

Using the resistance and capacitance values connected to each mesh point as specified by the mesh creator, our implementation code fills the left-hand side matrix A and the matrix B required for the right-hand side vector only once at the beginning of the time progression if a fixed time step scheme is used. In the case an adaptive time step method is applied, the matrices need to be recreated when the

time step changes.

3.3.3 Convolution Routine

Among the inputs of the program are two lists of points: The input (impulse) locations and output (target) locations. The program goes through the first list, solving for the full transient impulse response of each impulse individually. However, at each time step, it only saves the results at those points specified in the target location list.

When the impulse responses for each specified input location are calculated and saved at each specified output location, if an adaptive time step method has been used, the program then uses a spline method to interpolate between unequal time steps to get a fixed time step impulse response.

The input required by the convolution part of the program is the transient input waveforms applied at specified points. The input waveforms do not need to be localized at single points; in other words they could be spatially distributed over the mesh. It is only important that all points in this distribution to have been included in the list of impulse locations used before. With these waveforms as the $f_i[t_n]$ term and the response to the impulse applied at that input point as the $h_i[t-t_n]$ term in Eqn. 3.9, we implement Eqn. 3.6 to obtain the system's response to the given input(s). The computational cost per input and output points was derived previously in Sect. 3.2.2.

3.4 Simulation Results

3.4.1 Comparison with SPECTRE

We start by comparing both the impulse response and full input response solutions of our code with the outputs computed by SPECTRE for a 5x5x3 mesh (N=75), in which each node is connected with an R//C to its six nearest neighbors. The bottom layer nodes are down-connected to ground and top layer nodes have no up-connections. In this network the north-south resistors are 5 Ω, east-west resistors are 10 Ω, inter-layer resistors are 10 KΩ, and all capacitors are 10 mF. The impulses are injected at points (1,1,1) — bottom layer — and (3,3,2) — middle layer. The full input consists of a 4 ms, 4 A impulse injected at (1,1,1) and 3 ms later a 5 ms, 8 A impulse at (3,3,2); the rise and fall times are 50 μs. Fig. 3.9 demonstrates the agreement between the SPICE results and our solver. For this mesh, SPECTRE requires 0.24 msec per timestep, while our solver requires 0.0124 msec per timestep at each output point to calculate the output there for the full combined input from the impulse responses.

3.4.1.1 Convergence Test

We have performed a convergence analysis on the results of our code to demonstrate that it is first-order convergent, as expected from the time-discretization scheme that was used to set up the KCL equations featuring capacitors.

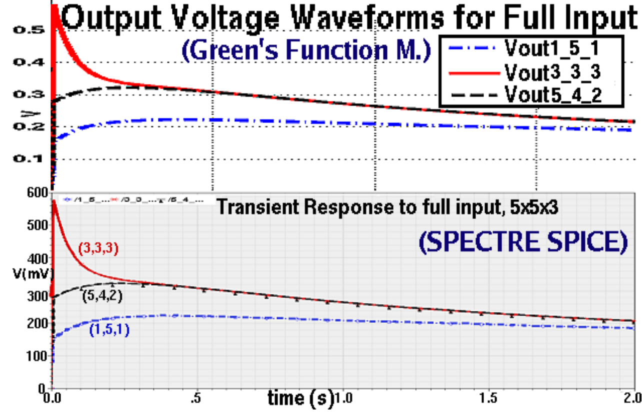


Figure 3.9: A 5x5x3 mesh full response simulation. Top: Our solver combines impulse responses at three points using the input data. Bottom: SPECTRE performs full mesh solution at each timestep for the same input pattern. The solutions match.

The time-discretization scheme we use has a first-order error component:

$$\frac{df}{dt} = \frac{f(t+\Delta t) - f(t)}{\Delta t} + \epsilon \Delta t \tag{3.15}$$

Let h_{exact} be the "true" solution to the impulse response problem. Let $h_{\Delta t1}$, $h_{\Delta t2}$ and $h_{\Delta t3}$ be the numerically-calculated impulse responses when the time steps used are Δt_1, Δt_2 and Δt_3 respectively. Then we can write

$$\begin{aligned} h_{\Delta t1} &= h_{exact} + \varepsilon \Delta t1 \\ h_{\Delta t2} &= h_{exact} + \varepsilon \Delta t2 \\ h_{\Delta t3} &= h_{exact} + \varepsilon \Delta t3 \ , \end{aligned} \tag{3.16}$$

where ε is a constant coefficient. If we subtract these three solutions from each other in pairs, we can define three values a, b and c:

$$a = h_{\Delta t3} - h_{\Delta t2} = \varepsilon(\Delta t3 - \Delta t2)$$

$$
\begin{aligned}
b &= h_{\Delta t3} - h_{\Delta t1} = \varepsilon(\Delta t3 - \Delta t1) \qquad (3.17)\\
c &= h_{\Delta t2} - h_{\Delta t1} = \varepsilon(\Delta t2 - \Delta t1) \ .
\end{aligned}
$$

Then knowing the relative magnitudes of the different time steps allows us to define the ratios of these three values to each other:

$$
\begin{aligned}
\frac{a}{b} &= \frac{\Delta t3 - \Delta t2}{\Delta t3 - \Delta t1}\\
\frac{a}{c} &= \frac{\Delta t3 - \Delta t2}{\Delta t2 - \Delta t1} \qquad (3.18)\\
\frac{b}{c} &= \frac{\Delta t3 - \Delta t1}{\Delta t2 - \Delta t1} \ .
\end{aligned}
$$

Thus we can run our simulation with different time steps, sample the solutions at the same output times, calculate the differences between the solutions, i.e. a, b and c of Eqn. 3.18 and compare their ratios to what we expect from Eqn. 3.19. Table 3.1 gives the result of such a convergence test applied to the output of a 21x21x5 mesh, whose impulse response results themselves are presented in the following section. The program was run with three time steps: Δt = 50 ps, 75 ps and 100 ps. The outputs h_{50}, h_{75} and h_{100} were taken at two output times, $t = 6.5$ ns and 100 ns.

t_{out} (ns)	h_{50}	h_{75}	h_{100}	$a = h_{75} - h_{50}$	$b = h_{100} - h_{75}$	$c = h_{100} - h_{50}$	a/b	a/c
6.5	6.2484	6.2414	6.2344	0.007	0.007	0.014	1.0	0.5
100	0.9215	0.9220	0.9225	0.0005	0.0005	0.001	1.0	0.5

Table 3.1: Example convergence test results. The ratios a/b and a/c are expected to be 25/25=1 and 25/50=0.5 respectively, which is what the test results yield.

3.4.2 Example 3D Network Simulations

Having verified the operation of our code with SPECTRE, we expanded the 5x5x3 mesh presented in the previous section to 21x21x5, using the same element values. The unit impulse is given at point (1,1,1), at the lowest layer. Figure 3.10 presents the spread of the Green's Function response on Layer 5 to this input over 25 ns. At 1 ns the impulse response has reached Layer 5 and begun to spread by coupling between the mesh points. It spreads more in the north-south direction, with its smaller resistances, than in the east-west direction. At around t=13ns, the peak has decayed considerably and reached the opposite edge, at which point it starts coupling back in the direction it came from. Once again the signal spreads more in the y-direction.

Moving on to a more uneven network designed to emulate a real chip interconnect network more closely, we take the three-metal-layer network presented in Figure 3.11. This network is 50x50x3 (N=7500). The two bottom layers are envisioned as bus lines. The top layer is designed after a clock H-tree. Metal 1 (M1) segments are assumed to point north-south (towards the up-right/down-left of the figure) and Metal 2 segments are assumed to point east-west; Metal 3 segments can point either way, depending on where they are in the H-tree.

Between the layers, there is resistive coupling where there are vias and capacitive coupling at other overlapping areas. At the first level, for points where there are no M1 segments, there are still "dummy" nodes to provide capacitive coupling between the upper metal layers and grounded substrate. There are similar nodes at

the second level where such coupling between M1 and M3 is needed.

The equivalent lumped element network values were obtained by using experimental data for a three-metal 0.5 μm process, published by the MOSIS fabrication facility [101]. Our network is comprised of 10 μm long, 4 μm wide metal segments, with a 6 μm separation.

Fig. 3.12 shows signals induced at four nodes of this network by a combination of a lowest-level, 0.1 mA noise spike injected at (5,15,1), and a top-level, 5 mA interference pulse injected at (25,25,3), the center of the H-tree. The closest output point (29,50,3) has the highest response to the (25,25,3) input. But it is worth noting that the other three, equidistant points exhibit different responses to this input due to the effects of the lower metal layers and a via on the path to (9,1,3). The lower left point (9,1,3) is found to be most sensitive to the ground-level pulse because of proximity and the presence of a close resistive path. The different end points of the H-network respond with different time constants and sensitivities to this input. With our method it is easy to explore the location-dependent response to other input signals.

Figures 3.13 and 3.14 display the impulse responses, over time, that the combined input response of Fig. 3.12 was calculated from using our method. Figure 3.13 displays the responses at level 1 and 2 to the impulse applied at (5,15,1) (on Level 1) and Fig. 3.14 displays the responses at level 2 and 3 to the impulse applied at (25,25,3) (on Level 3).

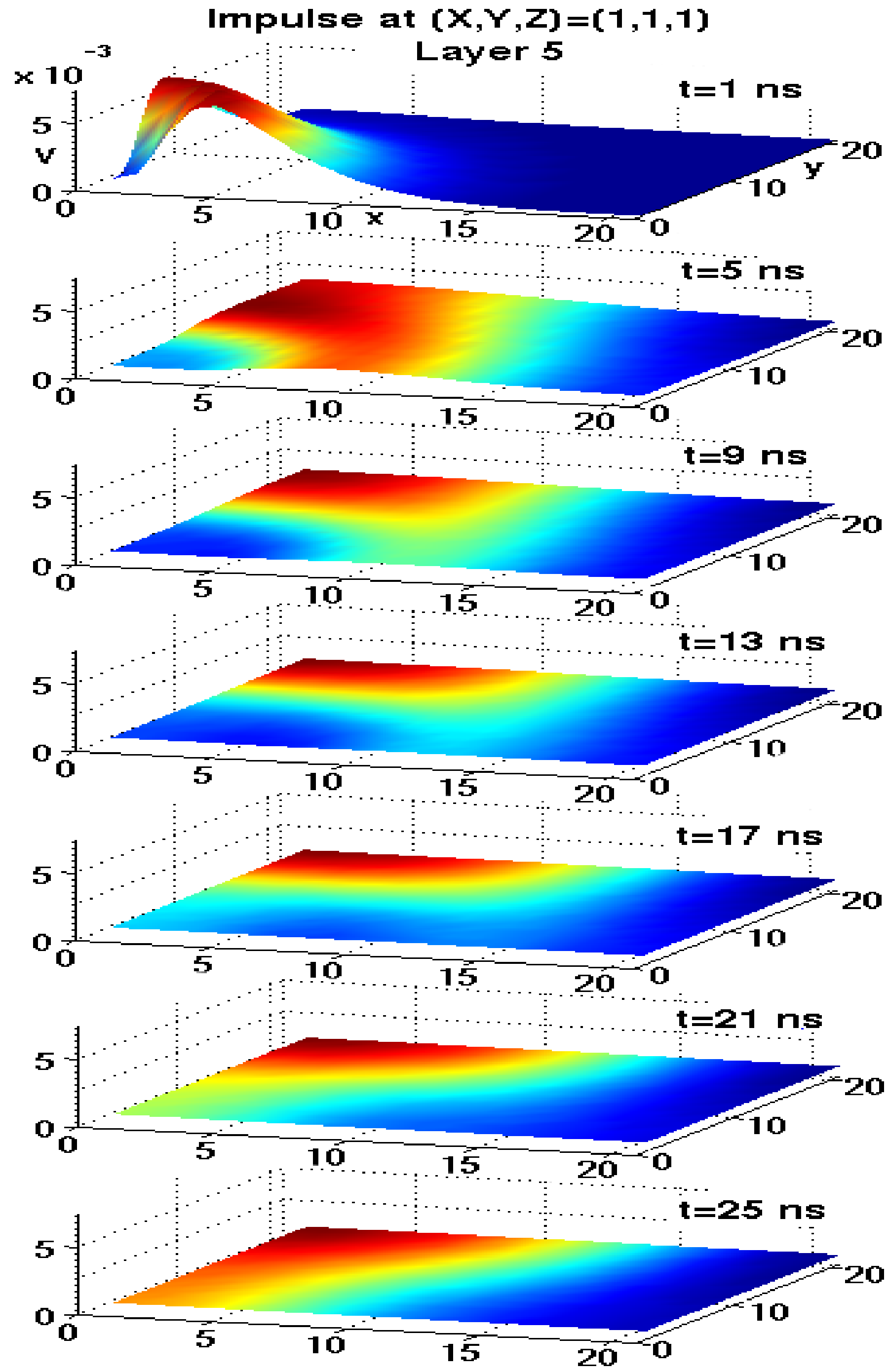


Figure 3.10: The time evolution of the Green's Function of a 21x21x5 system on the topmost leayer (layer 5). The unit impulse was injected at t=0 into point (1,1,1): The lower left corner of the lowest level, layer 1. The uneven spread is due to the lower resistivity of the network in the N-S direction.

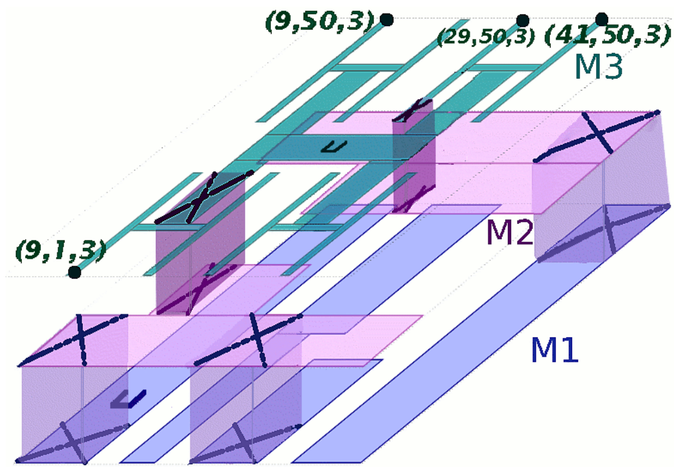


Figure 3.11: 3-metal interconnect network. Vias: **X**. Inputs: ⊔. Outputs: •.

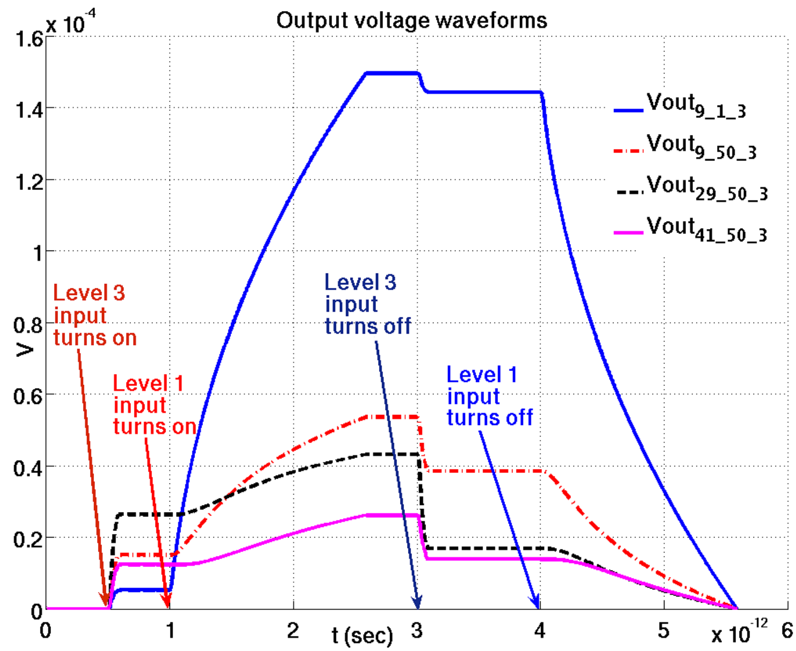


Figure 3.12: Responses of the network in Fig. 3.11 to a combined input signal injected at points (5,15,1) and (25,25,3). Distance as well as the effect of inter-level coupling and presence of paths to ground affect the responses.

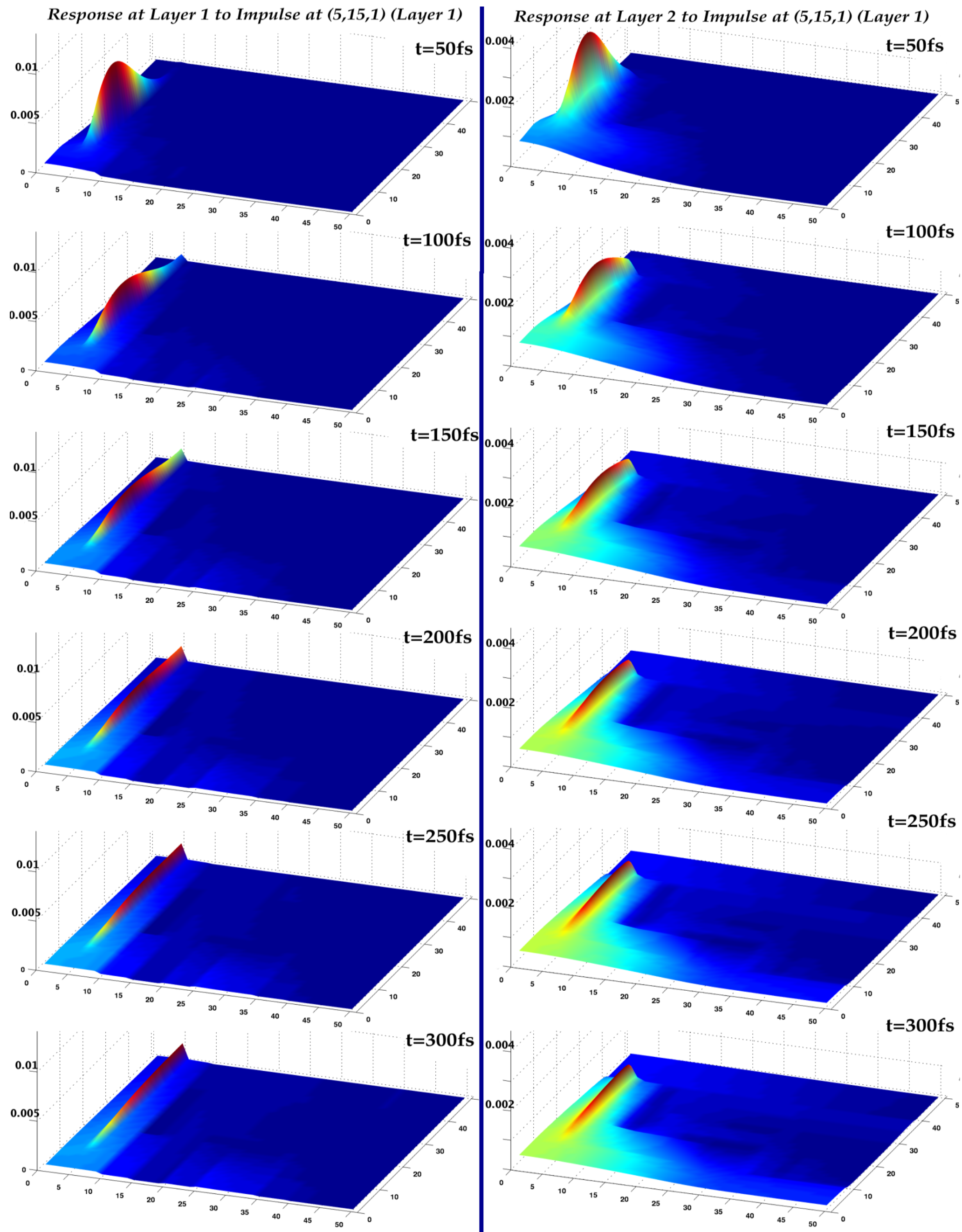


Figure 3.13: Responses of the network in Fig. 3.11 over time to an impulse injected at point (5,15,1). Left: Response at Level 1 (Metal 1 network) over 300 femtoseconds. Right: Response at Level 2 (Metal 2 network) over 300 femtoseconds.

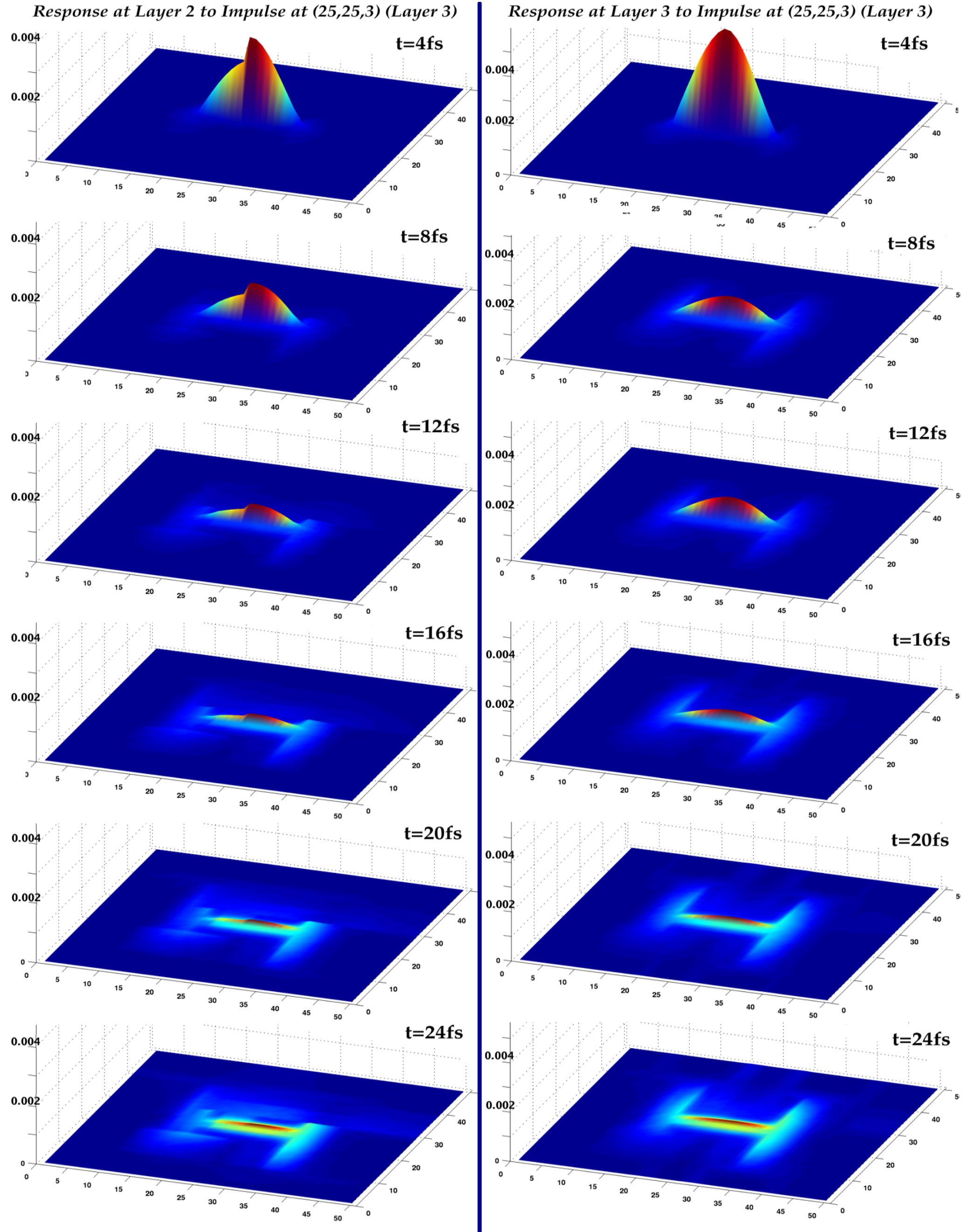


Figure 3.14: Responses of the network in Fig. 3.11 over time to an impulse injected at point (25,25,3). Left: Response at Level 2 (Metal 2 network) over 24 femtoseconds. Right: Response at Level 3 (Metal 3 network) over 24 femtoseconds.

3.5 Summary

In this chapter, we presented the development and results of a computationally-efficient Green's Function based method to investigate the response of a complex on-chip interconnect network of a planar or 3DIC to internal and external noise and signals. Our method divides the interconnect network into unit cells and models it as an RC network based on the corresponding RC "seed" unit cells. After presenting the numerical background and computational cost of our method, we described our implementation in terms of the mesh structure used and equation system to be solved. To demonstrate the validity of our simulator, we compared its solution for a given input into a small (5x5x3) network to the solution provided by Spectre. We also showed that the implementation is first-order convergent, as was expected from our time-domain discretization.

We used our simulator to investigate the time-domain behavior of a realistic 3-metal interconnect network, which incorporates a clock tree in the top layer and bus lines in the bottom two layers. Our simulation results revealed, among other things, that the presence of non-identical layouts in the lower metal layers affects the spread of signals in the top layer, causing a non-symmetric response there although this layer by itself has a symmetric layout. This is an example of how this method is useful for investigating interconnect network response details. Since the meshes we create are already three-dimensional, the work is extendable to model the coupled interconnect networks of the different tiers of a stack of chips in a 3DIC.

Chapter 4

Physical Aspects of VLSI Systems: On-Chip Heat Generation and Dissipation

4.1 Introduction and Motivation

In this chapter we consider thermal effects on integrated circuits and methods of modeling them.

As device scaling progresses in CMOS technologies, power densities increase. Combined with the rising system sizes and complexities, higher device densities and clock frequencies, and the effects of process innovations such as SOI technologies and 3-D integration, chip and device heating becomes a more urgent problem [27, 105, 106, 109, 110, 111]. On a different front, for electronic systems meant to be used in extreme conditions such as space applications, controlled heating to keep certain parts of the system at set temperatures could be required [112]. Either way, accurate modeling of heat generation and dissipation on integrated circuits is necessary. The ability to seamlessly include heating and cooling effects into the integrated circuit design flow is also desirable [113].

There is a considerable amount of complicated interaction of thermal and electrical processes in integrated circuits. Many electrical and material properties of semiconductors, for instance mobility, density of states, and thermal velocity all

depend on temperature. Thus for accurate modeling of device behavior it is essential to take the operating temperature into account [43]. Since during operation devices consume power and dissipate heat, they in turn have an effect in setting up the thermal conditions they operate in. Self-consistent modeling and analysis of device and chip-level heating is thus necessary and has been the subject of previous work [103, 104, 105, 106].

Hardware- and software-based and hybrid approaches have been proposed to deal with excessive heating in integrated circuits. Addressing the particular problems cropping up in digital [26, 27] and analog [28, 29] circuits are solutions implemented to scale down power consumption temporarily, compensate for specific electrical effects of temperature increase, and to carry away excess heat more efficiently. In this vein, cooling schemes involving dedicated cooling metal-level layouts on planar chips [61, 114] and vertical metal vias incorporated into 3-D stacks have been suggested [10, 32, 54, 56, 105]. We examine this particular idea in further detail in part of the present work. We then study the effectiveness of a new approach to heat selected IC regions using via frames.

We open this chapter with a review of the background physics and the pre-existing work. We next detail the design of integrated circuits meant for controlled heating and high-resolution temperature sensing in a commercial semiconductor process system. After experimental results and their interpretation, we present an analysis on cooling integrated circuits or heating them in a controlled manner using high-thermal-conductivity networks.

4.2 Background

4.2.1 The Heat Equation

An electrical component with the potential ϕ dropping across and the current density J flowing through it generates heat through *Joule* or *ohmic heating* [22]:

$$H(\vec{x},t) = \vec{J}(\vec{x},t) \cdot \vec{\nabla}\phi(\vec{x},t) \tag{4.1}$$

The generated heat H is measured in W/m^3.

In the physical system of an integrated circuit, many components generate heat, which must be dissipated. Current flow across interconnects and passive devices dissipates power and generates heat. Power consumption by active devices switching or in small-signal operation also generates heat. Even the unwanted power consumption by, for instance, MOSFET leakage currents is a source of heat [24, 25].

The equation that governs the flow of heat energy in a medium is a continuity equation [108]:

$$\frac{\partial U(\vec{x},t)}{\partial t} + \vec{\nabla} \cdot \vec{J}_T(\vec{x},t) = H(\vec{x},t) \tag{4.2}$$

Here $U(\vec{x},t)$ is the heat energy content per unit volume. J_T is the heat flux vector, measured in W/m^2. It is related to temperature T by the Fourier Law [23]:

$$\vec{J}_T = -\kappa\vec{\nabla}T \tag{4.3}$$

κ, the thermal conductivity, is a material property. Depending on the material system and temperature ranges, κ can be location- and temperature-dependent (and therefore time-dependent in a non-steady-state system. As the systems we study

initially do not have a wide temperature range, we will disregard the temperature dependency of κ.

Changes in U are related to the changes in temperature through another material property, the specific heat C_p:

$$\frac{\partial U(\vec{x},t)}{\partial t} = \rho C_p \frac{\partial T(\vec{x},t)}{\partial t} \quad , \tag{4.4}$$

where ρ is the material density.

The heat generated and stored in the integrated circuit system is also dissipated through connections to the ambient. Using the above definitions in Eqn. 4.2, we arrive at the heat conduction equation:

$$\vec{\nabla} \cdot (\kappa(\vec{x})\vec{\nabla}T(\vec{x},t)) + H(\vec{x},t) = \rho C_p \frac{\partial T(\vec{x},t)}{\partial t} \tag{4.5}$$

Assuming a constant heat source and unchanging system, the right-hand side of this equation can be set to zero to obtain the time-invariant heat flux equation, whose solution gives the temperature distribution of the system in steady state.

The boundary conditions necessary for solving Eqn. 4.5 are related to the mechanisms of heat transfer: Radiation, convection and conduction. In the case of thermal conduction in solids, the possible methods [108] are specifying insulating surfaces (implying no heat flow perpendicular to this surface), setting other Neumann boundary conditions (constant heat flux normal to the surface), setting Dirichlet boundary condition (constant T) surfaces or accounting for radiation from a surface. While modeling thermal distribution of ICs, we apply a variant of the Dirichlet boundary condition approach, detailed further in the following section.

4.2.2 A Simplified Solver and the Effects of Non-Isotropic Thermal Conductivity

The time-invariant heat-flux equation becomes

$$\vec{\nabla} \cdot (\kappa(\vec{x})\vec{\nabla}T(\vec{x})) = -H(\vec{x}) \; . \tag{4.6}$$

To help with the discretization, first we replace the heat flux vector:

$$\vec{\nabla} \cdot (\vec{J_T}) = -H \tag{4.7}$$

Figure 4.1 shows an example mesh for a two-dimensional thermal network to be solved. The temperature, which we solve for, is defined at the nodes just like the potential in an electrical network. The heat flux and the thermal conductivities are defined on the branches between mesh points.

If (m, n) are the index numbers in the x- and y-directions respectively, we can discretize Eqn. 4.7 as

$$\frac{J_T(m+\frac{1}{2},n) - J_T(m-\frac{1}{2},n)}{\Delta x} + \frac{J_T(m,n+\frac{1}{2}) - J_T(m,n-\frac{1}{2})}{\Delta y} = -H(m,n) \tag{4.8}$$

At this point, it is obvious that this equation can be considered an effective KCL equation, with the heat source the equivalent of a current source. Consequently, one can consider the heat flux to be the equivalent of current and temperature as that of potential, with the thermal conductivity relating them.

Let κ_E, κ_S, κ_W and κ_N be the thermal conductivities in the branch directions east, south, west and north with respect to each node, similar to the system used in the interconnect modeling network of Section 3.3.1. Then putting the definition

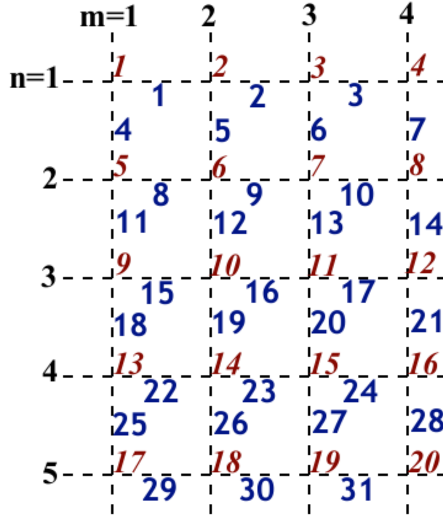


Figure 4.1: An example mesh for a 2-D thermal network. The temperature is measured at nodes, numbered with the red (italic) numbers. The thermal conductivity is the property of the material on the branches, numbered with the blue (sans serif) numbers, between the nodes.

of the heat flux vector back into the equation we obtain

$$\frac{\kappa_E \frac{T_{m+1,n}-T_{m,n}}{\Delta x} - \kappa_W \frac{T_{m,n}-T_{m-1,n}}{\Delta x}}{\Delta x} + \frac{\kappa_S \frac{T_{m,n+1}-T_{m,n}}{\Delta y} - \kappa_N \frac{T_{m,n}-T_{m,n-1}}{\Delta y}}{\Delta y} = -H(m,n) \ . \tag{4.9}$$

This is a system of equations, cast in the form of a matrix equation, with the right-hand side incorporating the heat source. The resulting matrix equation is sparse and has been solved by Gaussian elimination here.

To solve this two-dimensional system for temperature, we need either the temperature or the gradient of the temperature defined at the boundaries. With the assumption that the only heat transfer to outside the system is conductive, we set an ambient temperature T_{out} defined at virtual boundary nodes outside the mesh and

define “package” conductivities κ_{pack} at each boundary. In the case where radiative and convective heat transfer from the die surface is considered , [106] suggests the implementation of a heat sink at the surface mesh points to account for these.

We use the method presented in this section to obtain a feel for how thermal distribution changes in two-dimensional integrated circuits each with different, nonunform lateral thermal conductivities, as described below. For all examples, the mesh has 11x13 nodes and the heater is in the location $x \in [3..5]$, $y \in [8..10]$.

Uniform thermal conductivity. κ is the same on all branches. Figure 4.2, left side, displays a sample T distribution in our 11x13 mesh in this case. The thermal conductivity to ambient is the same in all directions. The nonuniformity of the temperature gradients is due to the asymmetry of the heater location: From the heater to the west- and south-boundaries, there is a lower “resistance,” therefore less of a “voltage drop”: That is to say, a lower temperature drop.

Uniform thermal conductivity and nonuniform boundary conductivity to the ambient. The same as above, but with nonuniform boundary conditions, as the southern boundary is defined to be more thermally conductive than the rest. The right side of Figure 4.2 gives the results. The peak temperature is lower with the lower overall effective resistivity to the ambient. There is a faster temperature drop towards the south edge, causing a sharper gradient in that direction.

Nonuniform thermal conductivity. κ is different at different spatial locations. The resulting T distribution is given by Fig. 4.4 when the thermal conductivity changes along the mesh as given in Fig. 4.3. Comparing this result to that of

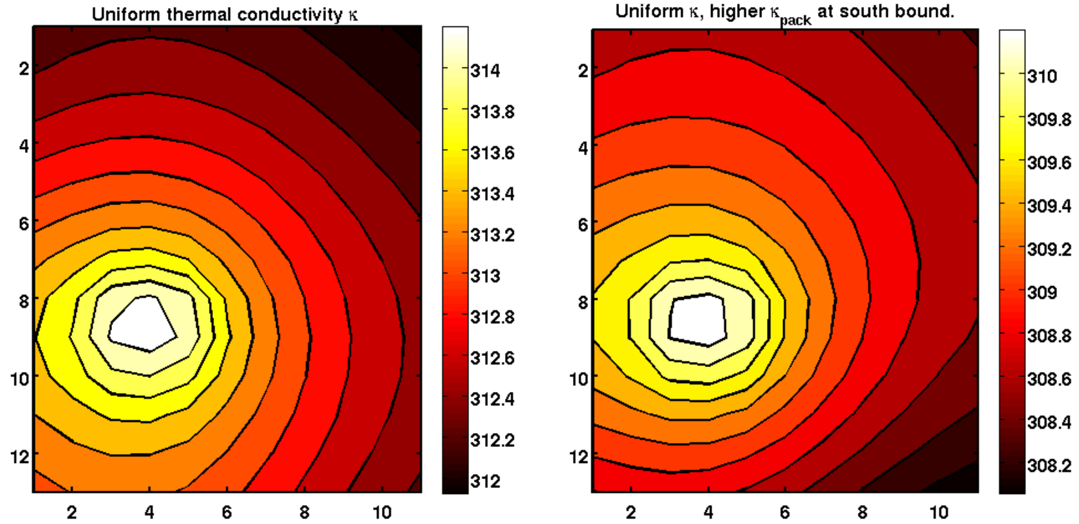


Figure 4.2: Left: Resulting temperature distribution with a uniform κ distribution. Right: Resulting temperature distribution with the same uniform κ distribution, with the southern boundary three times as thermally conductive to ambient than the other boundaries.

Fig. 4.2, one can immediately see that the gradient of temperature is much lower in the direction of the higher thermal conductivity, which is expectable considering the heat flux equation.

Directionally nonuniform thermal conductivity. κ changes with location, and is also different depending on whether it is defined on a north-south branch or an east-west branch. Figure 4.5 displays the κ-distribution for our example here. In the more highly-conductive areas of our mesh, the horizontal branches are only three times more conductive than the low-conductivity areas, but the vertical branches are still five times as conductive. Figure 4.6 shows the resulting temperature distribution. Compared to the result in Figure 4.4, there is higher

thermal spread towards the north border than the east border of the chip area in this case, as the thermal conductivity is higher in that direction.

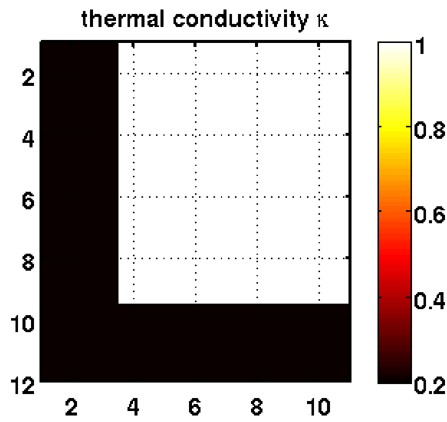


Figure 4.3: A non-uniform thermal conductivity distribution. The light-colored areas are five times more thermally conductive than the dark areas.

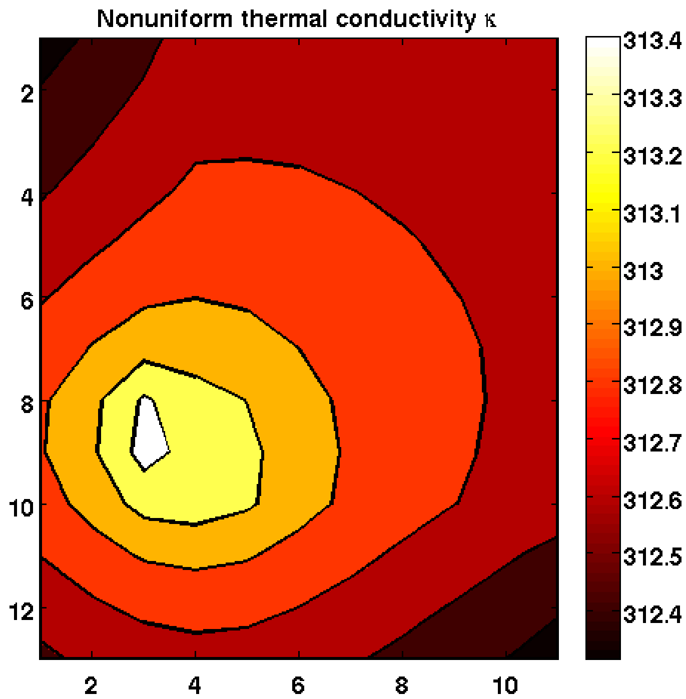


Figure 4.4: Resulting temperature distribution with the nonuniform κ distribution given in Fig. 4.3.

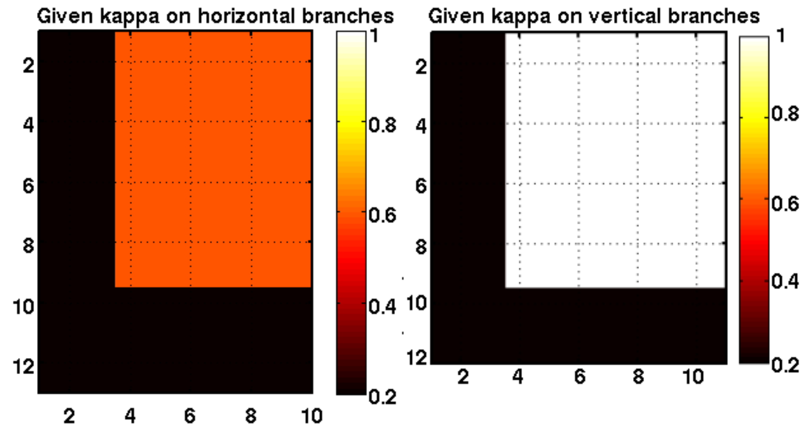


Figure 4.5: A directionally non-uniform thermal conductivity distribution. The left-side and right-side shows the κ distribution of the E-W branches and N-S branches respectively. The gray (orange) areas are three times, the light-colored areas are five times more thermally conductive than the dark areas.

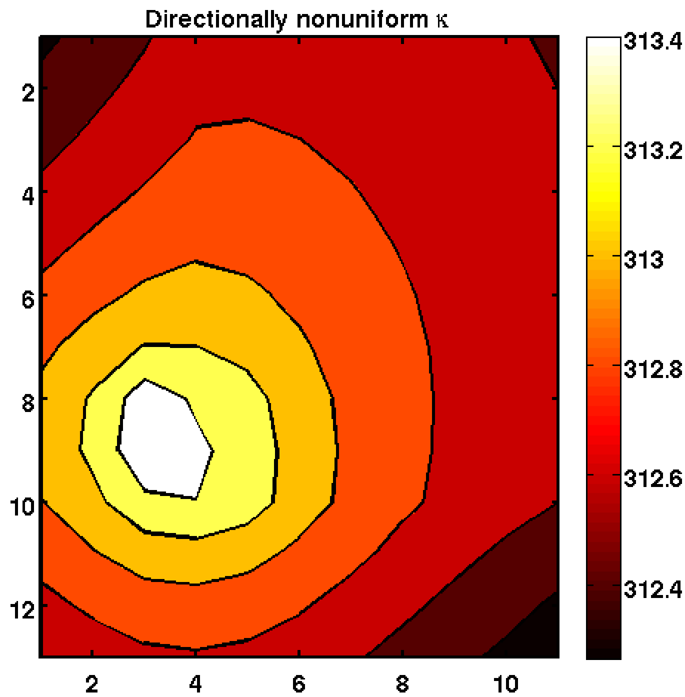


Figure 4.6: Resulting temperature distribution with the nonuniform κ distribution given in Fig. 4.5.

4.2.3 A Thermal Solver with Individual IC Layers Represented

The integrated circuit is a multi-material system. For a bulk silicon process, the silicon substrate is the highest-volume component. Its thermal conductivity is taken as 1.5 K/W·cm [106]. Next in volume are the silicon dioxide layers, which are highly thermally insulating since the conductivity, at 0.015 K/W·cm, is a hundred times lower than that of silicon. Finally, we use a conductivity of ∼3 K/W·cm for the metal layers.

Previously, our research group has demonstrated a methodology to solve self-consistently for heating and device performance in 2-D and 3-D integrated circuits [103, 104, 105, 106]. The thermal solver subcomponent of the implementation of this methodology models heat generation, storage and transfer as thermal current sources, capacitors and resistors respectively. Thus the KCL equations corresponding to this thermal network are cast and solved.

In previous applications of this method, the thermal conductivity of the integrated circuit was taken as a an averaged constant with the contributions of different layers. For this work, we created a finer 3-D mesh, with nodes placed in every individual layer (from substrate up: substrate, oxide, polysilicon, oxide, metal, oxide, metal, oxide) [107]. The thermal resistances connected to every mesh point in each layer were set independently according to their location in the layout. For instance, in a metal layer, the resistances going in the lateral directions to the nodes in the same layer were set using the metal thermal conductivity where there are metal interconnects in the layout and with the oxide conductivity in locations without

metal features. Similarly, in the metal layers and in the interlayer oxides, the vertical resistances going from layer to layer were set with the metal conductivity at the via or contact locations and with the oxide conductivity elsewhere. This let us create a layout-dependent thermal network.

To calculate the thermal resistivities between mesh points, we divided the materials in question into rectangular prisms, width (W) and length (L) determined by the mesh size and thickness (d) taken from the process parameters. Then the lateral thermal resistance between mesh points in the same material layer and the vertical thermal resistance connecting the layer to the layers above and below are obtained, respectively, by

$$R_{th,lat} = \frac{W}{\kappa L d} \quad R_{th,ver} = \frac{d}{\kappa L W} \quad . \tag{4.10}$$

We will present the layout details of the chips used in experimental and this modeling work in the next section. One thing that is important to point out here is that even though metal thermal conductivity is taken as twice that of silicon thermal conductivity, the lateral substrate thickness is nearly two hundred times that of a typical metal interconnect layer thickness in a bulk process. Therefore, the lateral thermal resistance between nodes in a metal layer is in fact higher than the same in the substrate layer. This implies that in a bulk technology, the substrate tends to dominate the thermal transfer. We will examine this point in more detail ahead.

4.3 Experimental Design

We have designed and fabricated two integrated circuits in AMI's 1.5 micron technology, to explore heat generation and distribution in a typical silicon chip. In this section, we list the features on these chips and present their operation details. The 1.5 micron technology was selected because we could have chips with larger surface areas fabricated through this process. The layouts/microphotographs for these two chips are shown in Figure 4.7. Both chips have lateral dimensions of 2200 μm by 2200 μm.

Both of the chips include the following features:

- *Temperature Sensor Array*: We designed an array of pn-junction diodes as localized temperature sensors.

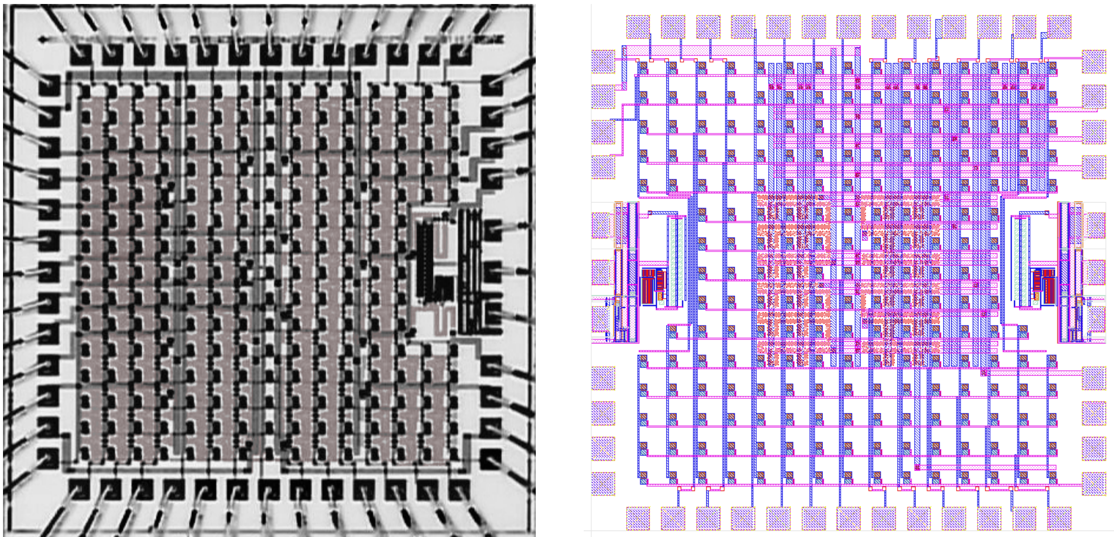


Figure 4.7: *Left*: The A-type heater/temperature array sensor chip, microphotograph. *Right*: The B-type heater/temperature array sensor chip, layout.

- *Heater Array*: We implemented an array of polysilicon resistors as micro-heaters.
- *Ring Oscillator*: Each chip features a ring oscillator as a sample active-device application.

4.3.1 Heaters

Serpentine-design polysilicon resistors are utilized as microheaters in our chips. Figure 4.8 displays such an example resistor. The design is intended to weave through the temperature sensor diode network, as the figure illustrates.

The A-type chip (Fig. 4.7) has 16 of these heaters in a 4x4 array, covering all of the diode array, named **R11** through **R44** . The B-type chip has 4 heaters in a 2x2 array in the center, covering only the middle of the chip area, named **R11** through

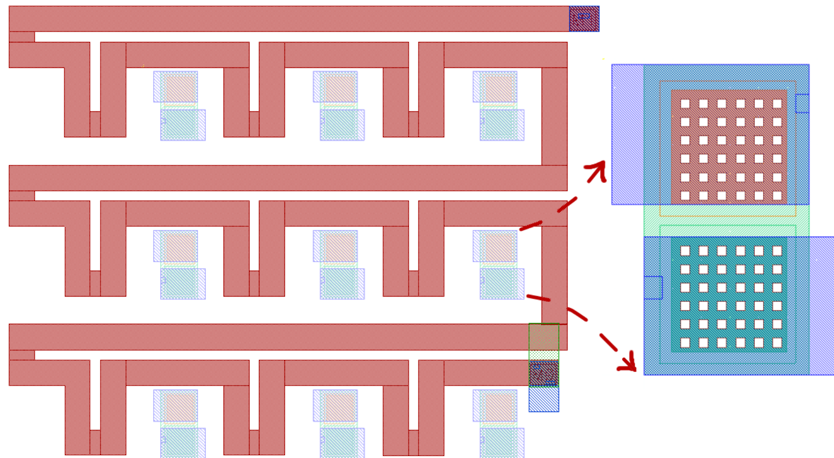


Figure 4.8: An example polysilicon microheater resistor. The poly layout shape is designed to surround temperature-sensor diodes, which are shown faintly within the microheater layout, with one example layout to the right.

R22. One end of each of these heaters, which are effectively resistors ranging from 3.1 KΩ to 6.15 KΩ, are shorted together to be grounded. The other ends go to dedicated bonding pads, thereby allowing us to individually turn each heater on and off.

4.3.2 The Temperature Sensor Array

The temperature-dependent characteristics of a pn-junction diode lets us use these diodes as temperature sensors [115]. Here we first give the derivation of the method we use to extract temperature information from the IV curves of diodes measured at two different temperatures. Next we present our diode array design.

4.3.2.1 Temperature Dependence of the Diode Current

Here we outline our method to extract the temperature of a diode from its measured I-V curve.

The current flowing through a pn-junction diode is often described by the following analytical equation [43]:

$$I_D = A_D q n_i^2 (\frac{D_p}{N_D L_p} + \frac{D_n}{N_A L_n})(\exp(\frac{qV_D}{nkT}) - 1) \tag{4.11}$$

where A_D, I_D and V_D are the diode area, current and voltage, $D_{p/n}$ and $L_{p/n}$ are the hole and electron diffusion constants and lengths, N_A and N_D are acceptor and donor densities, q is the electronic charge, n_i is the intrinsic carrier density, n is the diode non-ideality factor, k is the Boltzmann constant and T is the temperature in degrees Kelvin.

We approximate the diode non-ideality factor from the measured diode IV curves by calculating the slope of the function $\ln I_D$ with respect to V_D, which gives $1/nkT$ in the operation region where the diode is on and the -1 in the last parenthesis can be ignored.

Apart from the explicit dependence in the argument of the exponential, the intrinsic carrier concentration in Eqn. 4.11 is temperature dependent [43]:

$$n_i^2 = N_C N_V \exp(\frac{-E_g}{kT}) \tag{4.12}$$

where the effective densities of states at the conduction and valence band edges, N_C and N_V, both have the same of dependency on temperature:

$$N_C = c_{Nc} T^{3/2} \qquad N_V = c_{Nv} T^{3/2} \tag{4.13}$$

Here, c_{Nc} and c_{Nv} are proportionality constants. Therefore, defining $\alpha = c_{Nc} c_{Nv}$, we write

$$n_i^2 = \alpha T^3 \exp(\frac{-E_g}{kT}) \tag{4.14}$$

We take the bandgap energy E_g as temperature-independent for our purposes, since over a range of 40 degrees K centered around room temperature, it changes linearly by about 10 meV, i.e. about 0.9%. Similarly, back in Eqn. 4.11, we assume that the terms $D_{p/n}$ and $L_{p/n}$ are temperature independent. We further examine this assumption in Appendix A.

Now we can define another constant, β:

$$\beta = q(\frac{D_p}{N_A L_p} - \frac{D_n}{N_D L_n}) \tag{4.15}$$

Using this definition and Eqn. 4.14 in Eqn. 4.11, we obtain

$$I_D(T) = \alpha\beta T^3 \exp(\frac{-E_g}{kT}) \exp(\frac{q(V_D/n)}{kT}) = \alpha\beta T^3 \exp(\frac{q(V_D/n) - E_g}{kT}) \qquad (4.16)$$

Now let (V_{D1}, I_{D1}) be the applied diode voltage and resultant diode current at temperature T_1, and (V_{D2}, I_{D2}) be the same at temperature T_2. Taking the ratio of the two currents we obtain

$$\frac{I_{D2}}{I_{D1}} = (\frac{T_2}{T_1})^3 \exp(\frac{E_g - q(V_{D1}/n)}{kT_1}) \exp(\frac{q(V_{D2}/n) - E_g}{kT_2}) \qquad (4.17)$$

Defining a function $f(T_2)$ as

$$f(T_2) = (\frac{T_2}{T_1})^3 \exp(\frac{E_g - q(V_{D1}/n)}{kT_1}) \exp(\frac{q(V_{D2}/n) - E_g}{kT_2}) - \frac{I_{D2}}{I_{D1}} \qquad (4.18)$$

with the assumption that we know T_1, (V_{D1}, I_{D1}) and (V_{D2}, I_{D2}), we can use a numerical method to solve the nonlinear equation

$$f(T_2) = 0 \qquad (4.19)$$

to obtain T_2 for each diode. We use Newton's Method, modified with an update-limiting coefficient [116]. The basic Newton's Method algorithm uses the following update scheme: Starting with an initial guess $T_2^{k=0}$,

$$T_2^{k+1} = T_2^k - \frac{f(T_2)}{f'(T_2)} \quad . \qquad (4.20)$$

Here, $f'(T_2)$ is given by

$$\begin{aligned} f'(T_2) &= \frac{df}{dT_2} \qquad (4.21) \\ &= (\frac{1}{T_1})^3 \exp(\frac{E_g - q(V_{D1}/n)}{kT_1}) \exp(\frac{q(V_{D2}/n) - E_g}{kT_2})[3T_2^2 + T_2\frac{E_g - q(V_{D2}/n)}{k}]. \end{aligned}$$

Equations 4.18 and 4.22 are used in a modified version of Eqn. 4.20, wherein the update is size-limited to prevent overshooting and slow convergence in case of a bad initial guess:

$$T_2^{k+1} = T_2^k - 0.25\frac{f(T_2)}{f'(T_2)} \quad . \tag{4.22}$$

4.3.2.2 Diode Array for Distributed Temperature Sensing

Both chip designs presented in Figure 4.7 incorporates a 15×15 array of pn-junction diodes, with 10 diodes cut out of the right side of the A-type design to make space for a ring oscillator (B-type design has two cutouts of 9 diodes each, symmetrically placed). Therefore we have a total of 215 diodes on each A-type chip and 207 diodes on each B-type chip. Instead of requiring individual p- and n-side connection pads for each diode, in both types we have shorted the p-sides of the diodes in each of the 15 columns and the n-sides of the diodes in each of the 15 rows, therefore using only 30 pads to address each diode individually. A 5x5 section of the array is schematically presented in Fig. 4.9. The columns contact the p-sides and are therefore labeled ****p**; the rows, similarly, ****n**, where ****** varies between "01" and "15".

The diode layout itself is given in Figure 4.8. In this design, 10 μm by 10 μm p- and n-active regions are laid inside an n-well in the single-well, p-type substrate process we are using. The junction is formed between the p+ active area and the n-well.

Figure 4.10 shows the IV curves of two diodes, both at the same location (row

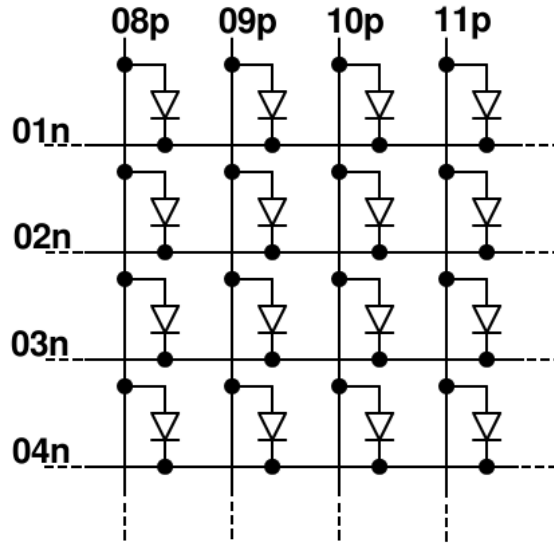


Figure 4.9: Schematic representation of five columns and five rows of the diode array. For reference purposes, for instance, the diode on the top left corner of this schematic would have the label **d08p01n**.

8, column 6), on the Type A and Type B chips. The first (dashed) set of curves are measured with no power applied to the heaters; the solid lines are with the heaters right above these diodes turned on.

4.3.3 Ring Oscillators as Temperature Sensors and Heat Sources

We included ring oscillators in both types of chip design to verify the effect of temperature changes on their operation and to evaluate temperature changes caused by these structures. Our designs are 31-stage ring oscillators designed from minimum-size inverters, the PMOS size scaled for a mid-range shift. Their output is connected to a five-stage output buffer with geometrically increasing transistor sizes to drive off-chip loads. While switching, the transistors that comprise the oscillator

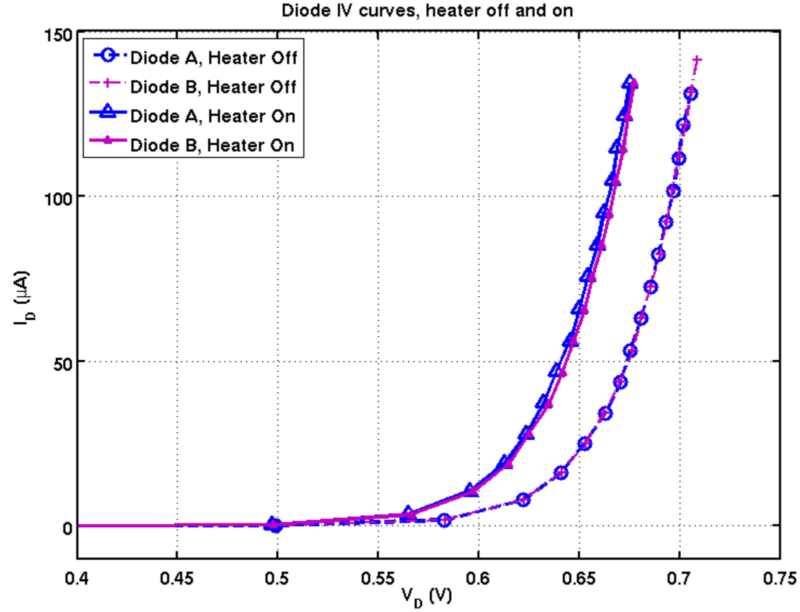


Figure 4.10: IV curves of two diodes on the Type A and Type B chips. The dashed and dot-dashed lines with circle and square markers are measurements taken at room temperature. The solid lines with triangle markers are measurements taken with the heater over these diodes turned on.

and buffers consume power and dissipate heat. In the meantime, the current-voltage characteristics and therefore the power consumption of the transistors themselves depend on their local temperature.

As will be covered in more detail in Section 5.6.1, the operating frequency of a ring oscillator is given by [12]

$$f_{osc}(T) = \frac{1}{N(t_{PLH}(T) + t_{PHL}(T))} \tag{4.23}$$

where N is the number of inverters and $t_{PLH/PHL}$ are the low-to-high or high-to-low propagation delays. The delays depend on the gate capacitance and the effective resistances of the transistors during switching, and these resistances are proportional

to the bias (rail) voltage and inversely proportional to the on-state drain current. The temperature dependency of oscillation frequency stems from this drain current. In saturation, the drain current of an NMOS can be given by the standard threshold-voltage based approach:

$$I_D(T) = \mu_n(T)\frac{W}{L}\frac{\epsilon_{ox}}{t_{ox}}(V_{GS} - V_{thn}(T))^2(1 + \lambda(V_{DS} - V_{DS,sat})) \tag{4.24}$$

where μ_n, ϵ_{ox} and t_{ox} are the channel electron mobility, oxide dielectric constant and oxide thickness, V_{GS} is the gate-to-source voltage and V_{thn} is the n-channel MOSFET threshold voltage. λ is the channel-length modulation coefficient.

Here, μ_n and V_{thn} are the main factors dependent on temperature, and the full $I_D(T)$ dependence is nontrivial. $\mu_n(T)$ can be modeled by a power law [118]. We explore the threshold voltage dependency by starting from the standard elements that go into deriving the threshold voltage according to the approach in [12]. There, the factors contributing to the final threshold voltage are shown to be the gate potential required to invert the surface potential, the built-in potential difference between gate and the body, and the initial depletion region charge. The built-in potentials both depend ot kT/q and on the intrinsic carrier concentration, which itself has the temperature variation mentioned in Section 4.3.2.

Combining these effects in the derivation given in Appendix B, we arrive at a description of the temperature dependency of the NMOS drain current as follows:

$$I_D(T) = c_1 T^{-2.5}[V_{gs} - aT - b\sqrt{cT - 1.5T\ln(T) + d} - e\,]^2(1 + \lambda(V_{DS} - V_{DS,sat}))\ , \tag{4.25}$$

where a, b, c, c_1, d and e are constants dependent on physical constants and fabrica-

tion parameters. The initial power-term is the effect of mobility variation, while the other terms bring together the effects of surface and gate potential dependencies of the threshold voltage.

It should be remarked that sub-threshold or linear region operation is not considered here. The further complexity that such considerations require illustrates the value of physics-based semiconductor modeling, particularly in the context of self-consistent thermal/electrical solutions [30]. This $I_D(T)$ dependency governs the effective switching resistance, which in turn yields a nontrivial temperature dependency for the oscillation frequency by using the following in Eqn. 4.23 [12]:

$$R_{effN}(T) = \frac{V_{DD}}{I_{DN}(T)} \; ; \; t_{PHL}(T) = 0.7 R_{effN}(T)(2.5C_{oxN} + 2.5C_{oxP}) \qquad (4.26)$$

for the NMOS (output high-to-low transition), and the corresponding expression for the PMOS (output low-to-high transition).

Overall, oscillation frequency goes down as temperature goes up.

Figure 4.11 displays simulation results of oscillation frequency vs. rail voltage with changing temperature for the 31-stage oscillator included in our designs. It is interesting to observe that as the operating temperature increases, the slope df_{osc}/dV_{rail} decreases. Hence at already higher temperatures decreasing the operating voltage of such a clock circuit becomes less effective for decreasing its frequency than at lower temperatures, which may have implications in the design of control-theory based dynamic thermal management systems.

We also need to consider the ring oscillator as a heat generator during operation, as the power consumed during switching is dissipated as heat. This power is

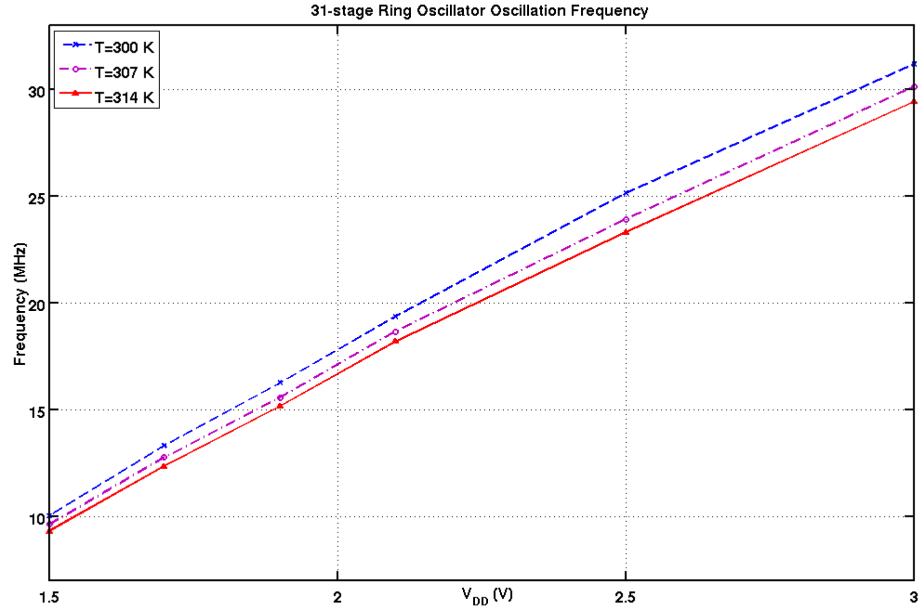


Figure 4.11: SPECTRE simulation results for 31-stage minimum-size ring oscillator oscillation frequency vs. rail voltage for T=300 K, 307 K and 314 K, using the AMI 1.5 μm technology.

proportional to the rail voltage and the drain current, whose temperature dependence was considered above. The exact effect of this heat source on the final chip temperature depends on other factors such as the layout area and location.

4.3.4 The Metal Layers

As part of our study of thermal effects of metal layers in integrated circuits, we designed two integrated circuits whose layouts differ mainly in the metal interconnect networks they incorporate. After observing experimentally what, if any, effect these different networks have on the resulting temperature distribution, we expected to be able to use the experimental results for verification and calibration of our simulator,

which is capable of calculating the temperature individually in each material layer.

AMI's 1.5 micron process is a two-metal process. In our chip layouts, generally, the north-south directed interconnects are laid out using metal 1 (M1) and the east-west directed interconnects are laid out using metal 2 (M2).

Due to the diode array addressing scheme, there is already an almost-uniform metal network comprised of minimum-width lines overlaying the entire chip area. However, there are wider metal lines interwoven with this design:

- On Chip A, 16 such wide lines, laid out on M2, connect one terminal of each individual heater resistor to a bonding pad. These lines are mostly uniformly distributed around the chip.

- On Chip A, 4 wide lines, laid out on M1, connect the other terminals of all heater resistors to a single bonding "ground" pad.

- On Chip B, 4 wide lines, laid out on M2, connect one terminal of each individual heater resistor to a bonding pad. These four lines are all concentrated on the northeast quadrant of the chip.

- On Chip B, 2 wide lines, laid out on M1, connect the other terminals of all heater resistors to a single bonding "ground" pad.

- On Chip B, the entire upper right (northeast) quadrant of the chip is covered with a net of wide M1 interconnects going north-south and M2 interconnects going east-west. One such north-south line is extended to the southern boundary of the chip, where it connects to an east-west M2 line going only east.

In conventional layouts, wider metal lines are used to conduct higher currents to avoid electromigration-related problems. They also present higher thermal conductivity than narrower metal lines. Furthermore, the thermal conductivity of the metal interconnect material is higher than that of the surrounding oxide. We shall explore the effects of this difference in Section 4.5, where simulation results of real and theoretical integrated circuits with different metal layouts is presented.

4.4 Measurements of Temperature Distribution in Integrated Circuits

4.4.1 Experimental Setup

The chips are packaged in LCC-52 surface mount type packages from Kyocera. They are mounted on printed circuit boards designed and fabricated for this application. These boards allow us to access the 15 diode rows and columns, as well as all the resistors and the ring oscillator voltage rail, the ground (and p-substrate tap), and output, through pin adapters. 15 adapter pairs are arranged for inserting external serial resistors common to each row of diodes. The PCB can be completely enclosed in a large isolation box, or just the chip can be placed in a small isolation box. A photograph of the measurement setup is given in Fig. 4.12.

A programmable voltage source (HP E3631A) and a digital multimeter (Agilent 34401A), controlled by a voltage sweep-current measurement program coded in Agilent VEE, are used to retrieve IV curves of the diodes. The pin connecting the measurement setup probes to the diode under measure is moved from column-

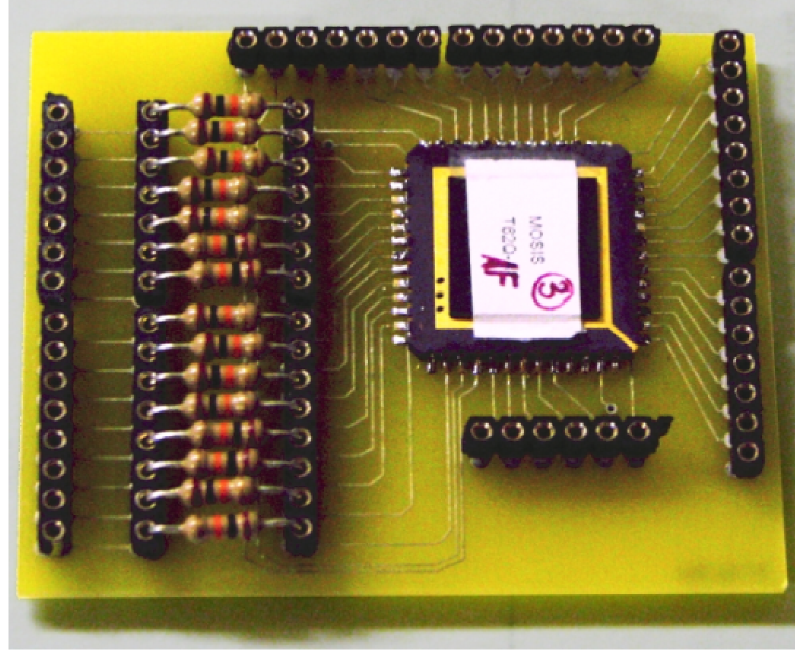


Figure 4.12: A Type-B measurement chip mounted on the experimental PCB.

to-column (p-side) and row-to-row (n-side) to address each diode one by one. The entire diode array is measured this way, first with no heaters turned on (to obtain the (V_{D1}, I_{D1}) pair to be used in the function of Eqn. 4.18), then with one heater turned on after a period of waiting (to obtain the (V_{D2}, I_{D2}) pair).

The ideal IV-curve measurement circuit, with no parasitics shown, is presented in Fig. 4.13. The sweep program records the voltage applied across the diode/resistor and the current. A post-processing program we developed, possessing a database of all the serial resistors, extracts the voltage drop across the diode.

4.4.2 Parasitic Effects

We observed two parasitic effects which might introduce errors into the temperature values obtained from our measurements.

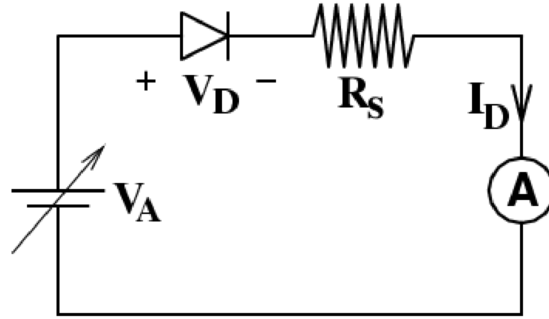


Figure 4.13: The circuit for diode IV curve measurement.

- *Row/column charging.* In the beginning of a no-heat or heated measurement, as diodes in the same row or the same column were being measured, the obtained IV curves would show a clear trend in the measurement direction. We theorized that this effect was due to the parasitic capacitances within the diode array being gradually charged during the first few measurements. To prevent this, 1 nF capacitors were placed between each row and column not under measurement and ground.

- *Parasitic BJTs.* Figure 4.14 displays the parasitic pnp BJTs that form between nearby diodes and the substrate, which are operating in the forward active region to turn on and draw some of the diode current as the IV measurement was being taken. To turn these BJTs off, each row under measurement was shorted to the psubstrate tap. The equivalent circuit of the diode array and the effect of this shorting is also shown in Fig. 4.14.

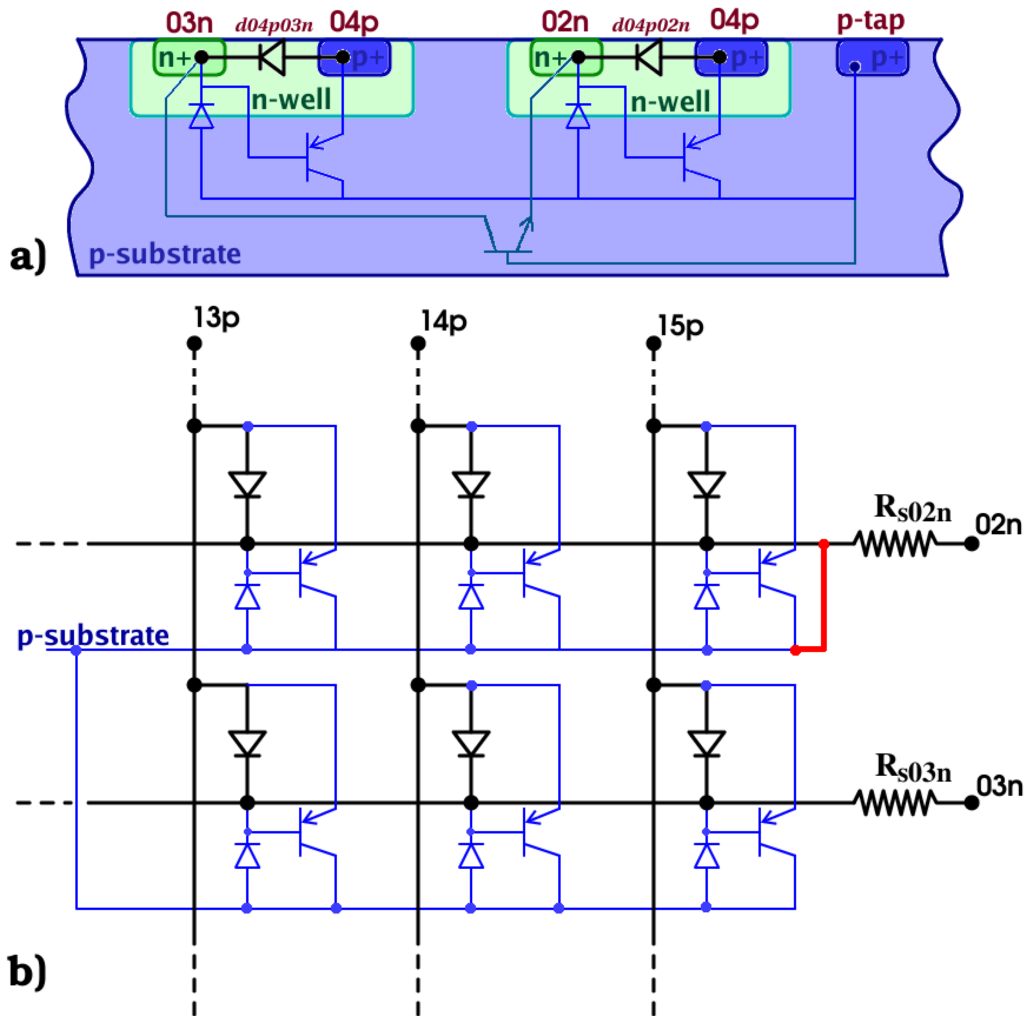


Figure 4.14: **a)** The formation of parasitic BJTs in the diode array. **b)** The simplified equivalent circuit of a small (3x2) diode array, with the parasitic BJT and n-well/substrate diodes shown. The red (thick) wire indicates the shorting of the n-side of the diodes of the row under measurement, Row 2 in this example, to the p-substrate.

4.4.3 Measurement Results

Chip Type A Data Figure 4.15 shows the measured temperature distribution on the Type A chip when the resistor in the marked position is turned on. A potential difference of 25V is applied across this 3110Ω polysilicon heater, which should thus be dissipating approximately 0.2 W of power. The peak temperature is 315.13 K and the lowest measured temperature is 312.1 K. As expectable, the peak temperature is in the substrate area right underneath the powered heater.

Chip Type B Data Figure 4.16 shows the measured temperature distribution on the Type A chip when the resistor in the marked position is turned on. The peak temperature is 314.36 K and the lowest measured temperature is 311.69 K.

The different minimum temperatures of the two chips while they are under equal heating power is due to variations in package and mounting thermal resis-

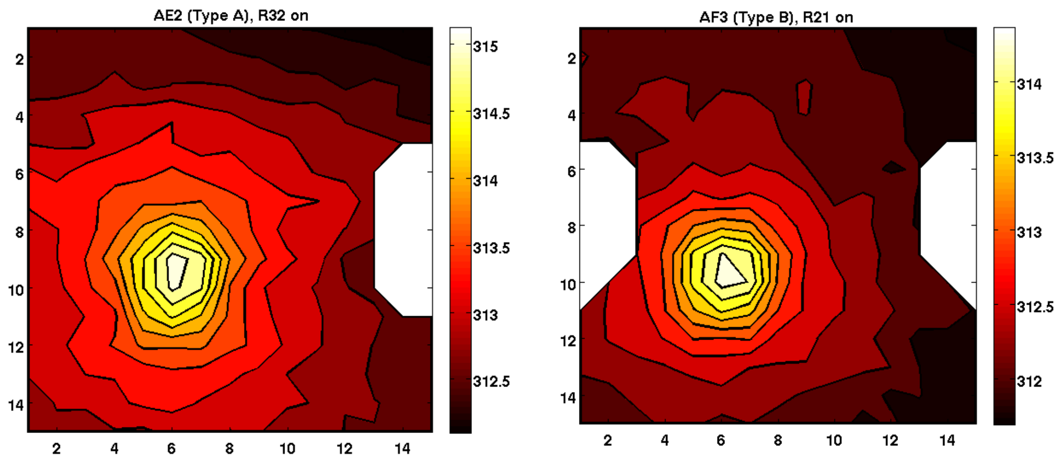


Figure 4.15: Measurement results from the Type-A chip with the uniform metal network.

Figure 4.16: Measurement results from the Type-B chip with the asymmetric metal network.

tances. The wider temperature spread of the Type A chip can be explained by its lower lateral thermal resistance to the package, as it features more bondwires. Both of these observations are verifiable by simulation.

One thing to note is that beyond measurement errors, the temperature distribution shape does not seem to be significantly affected by the presence of the asymmetric metal network on the Type B chip. This is because, as per our earlier remarks, the major lateral heat transfer medium in the IC system of a bulk process is the substrate itself, and the substrate geometries of both chips are extremely similar. We will return to this point in the simulation section.

We measured ring oscillator frequency vs. rail voltage under different heating levels as well. The results in Fig. 4.17 corroborate the observations of Section 4.3.3.

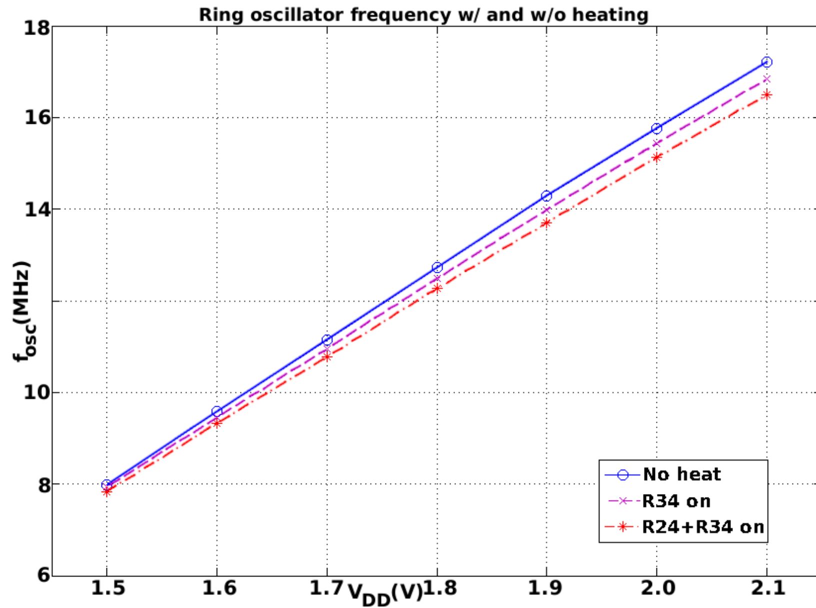


Figure 4.17: Measured ring oscillator frequency vs. rail voltage from chip Type A, when one (R24) or two (R24+R34) of the heaters next to the oscillator are operating.

4.5 Simulations: Controlled Heating and Cooling on ICs

4.5.1 Simulator Validation

Using the capability of our solver to model each material layer in our ICs as described in Section 4.2.3, we emulate our experiments by applying power at the resistor layouts defined in the polysilicon layer and observing the temperature distribution in the substrate. One advantage of simulation is that we can further observe the temperature distribution within every layer individually.

We start by validating our simulation and calibrating its thermal resistivity parameters by matching simulation results to measurements. To obtain the different material thermal resistivities to be used in our simulator-created mesh described in Section 4.2.3, we apply the following steps:

- For the oxide and metal layers, we calculate a resistivity-per-meshpoint based on estimated layer thicknesses and the thermal conductivities of 0.0015 K/W·cm for SiO_2 and 3 K/W·cm for metal.
- For the substrate, initially we assume a thickness of 200 μm and a thermal conductivity of 1.5 K/W·cm to calculate the lateral thermal resistances.
- To ascertain the thermal resistances to the package, from below the substrate, through the bondwires on the sides and from above, we use an iterative calibration process. It is these values that govern the peak temperature shift from the ambient.
- During this process, we also modify the lateral substrate resistivity, in order

to calibrate the observed temperature range. It is this parameter to a greater degree, and the package resistances to a lesser degree, which control the observed temperature range. This calibration step accounts for the uncertainity we have in the thicknesses of the substrate and the conducting epoxy bonding the die to the package, and the thermal conductivity of the latter.

Figures 4.18 and 4.19 show simulated temperature distributions with matching peak temperatures and temperature ranges to those obtained by measurement. The same heaters as were running during the experiments are being powered with 25 V for the simulations as well. The profiles geometrically differ somewhat from the measurements since the spatial bondwire and conductive epoxy variations around the chip are being implemented as averages in the simulator: As the results in Section 4.2.2 have shown, having a different thermal conductivity to one side of the IC die makes an observable difference in the thermal gradients.

The matching thermal conductivity parameters for chip Type B indeed did yield a slightly higher lateral thermal resistivity to the package. While this, along with a slightly lower lateral conductivity in the substrate, makes the Type A chip display a lower temperature gradient near the peak, its minimum and maximum temperatures are lifted higher than that of Type B by its extracted higher vertical resistivity to the package. We believe that packaging and mounting variations somewhat account for the differences.

Matching the measurement data with reasonable parameters has thus allowed us to verify our simulator operation. We now take advantage of this to evaluate our

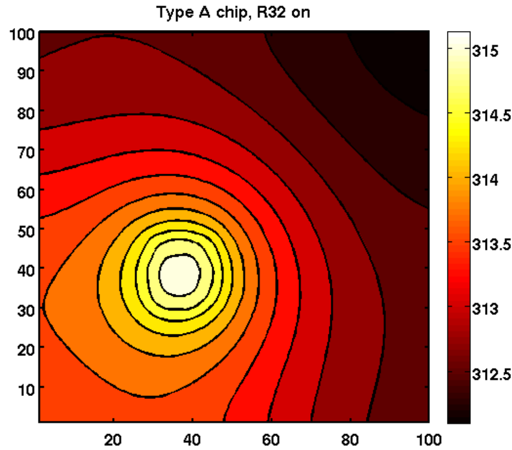


Figure 4.18: Simulation results for the Type-A chip (uniform metal network), displaying the temperature distribution at the substrate level, matching Fig. 4.15.

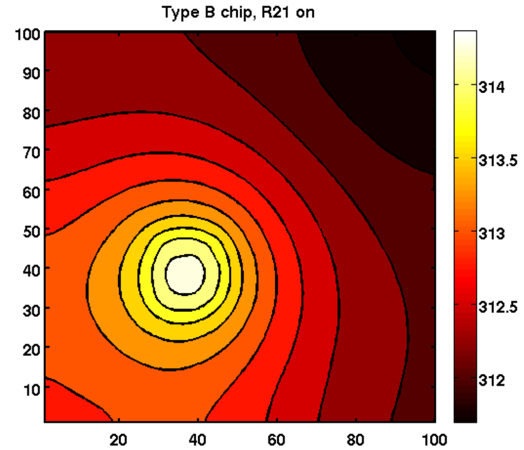


Figure 4.19: Simulation results for the Type-B chip (asymmetric metal network), displaying the temperature distribution at the substrate level, matching Fig. 4.16.

claim in the previous section that the presence of the extra lateral metal network is not very influential in the resulting temperature profile. Rerunning the simulator with unrealistically low values for metal thermal resistivity, which would have required tens of microns thick interconnects to achieve, we obtain very little change in the resulting temperature profile or range in the substrate. The substrate is a uniform thermal conductor with a high conductivity which indeed dominates heat transfer in the planar bulk integrated circuit system.

4.5.2 Controlled Heating and Cooling with Thermal Via Frames

Above the substrate, the oxide layers impede vertical heat transfer. This may especially be the case in 3-D systems, with isolating separator layers between the circuit tiers. Vertical thermal vias have been suggested as cooling aids [10, 56].

To study the effects of vertical thermal vias in a planar bulk technology, we implemented a thermal via ring around the operating heater in our simulator. This ring consists of metal posts which connect the ambient to the substrate, passing through the overlaying oxide, metal and polysilicon layers. The thermal conductivity of the ring material is taken to be the same as that of the metal layers.

Connecting this vertical thermal via ring directly to the ambient has two effects: Bringing down the peak temperature by 8.65 K from the measurement given in Fig. 4.15, and changing the temperature distribution profile. These effects on the temperature distribution of the Type A chip while the heater is again dissipating 0.2 W is shown in Figure 4.20. We observe that the temperature gradient near the peak gets sharper. To investigate possible exploits of this effect for controlled, localized heating in integrated circuits, we increase the dissipated power to 0.46 W to obtain a temperature distribution with the same peak temperature in the heated area as the measurement of Fig. 4.15. The result is shown in Fig. 4.21. In this case, while the peak temperature in the substrate area below the heater is comparable to that obtained before, the temperature in the outlying areas is about 1 degree Kelvin lower, in spite of the higher heater power. Compared to 4.18, the temperature distribution features higher gradients within the frame.

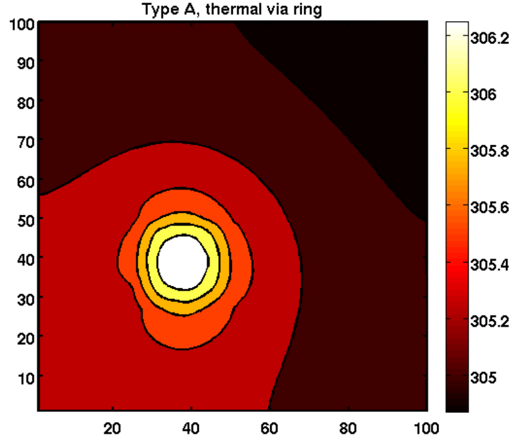


Figure 4.20: Simulation results for the Type-A chip with a vertical thermal via frame consisting of metal poles from the topmost layer all the way down to the substrate placed around the operating heater.

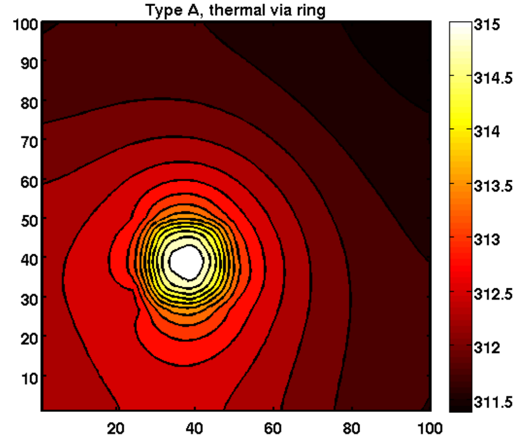


Figure 4.21: Simulation results for the Type-A chip, with the same vertical thermal via frame placed around the heater, operating at 0.46 W to obtain a comparable temperature range as the original measurement.

Just to emphasize the effect, we have simulated the case where the vertical via frame is a solid metal frame instead of single poles. This results in even higher thermal gradients and a more-highly limited temperature spread. The resulting temperature distribution in the substrate is given in Fig. 4.22. The same at the polysilicon layer level is given in Fig. 4.23, in which the via frame location is visible, as well as the outline of the heater. To get a substrate peak temperature just 0.2 K higher than the results with the via frame consisting of metal poles, given in Figure 4.21, we increased the heater power further to 0.59 W, since the continuous via frame would decrease the peak temperature further by itself.

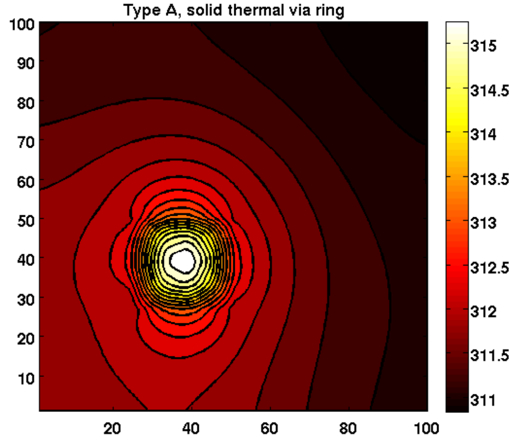


Figure 4.22: Simulation results for the Type-A chip substrate temperature, with an unbroken vertical thermal via frame from the topmost layer all the way down to the substrate placed around the operating heater.

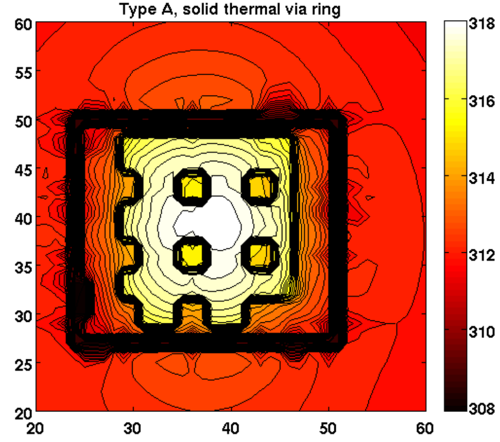


Figure 4.23: Simulation results for the Type-A chip polysilicon layer temperature with the same frame as Fig. 4.22. Note that the axis of this figure are zoomed in from the full-chip scale.

We next ran the simulator with a heater outside the via frame operating. This creates the interesting effect that the coldest part of the substrate now is the section that is shielded from the heater location by the via frame, which acts like a low-resistive path to "ground" (the ambient) for the thermal "currents". The result is displayed in Fig. 4.24.

4.5.3 Effects in SOI Technologies

In SOI technologies, the silicon substrate is much thinner. As an example, the SOI technology used in the fabrication of the system in Chapter 2 has an 50 nm

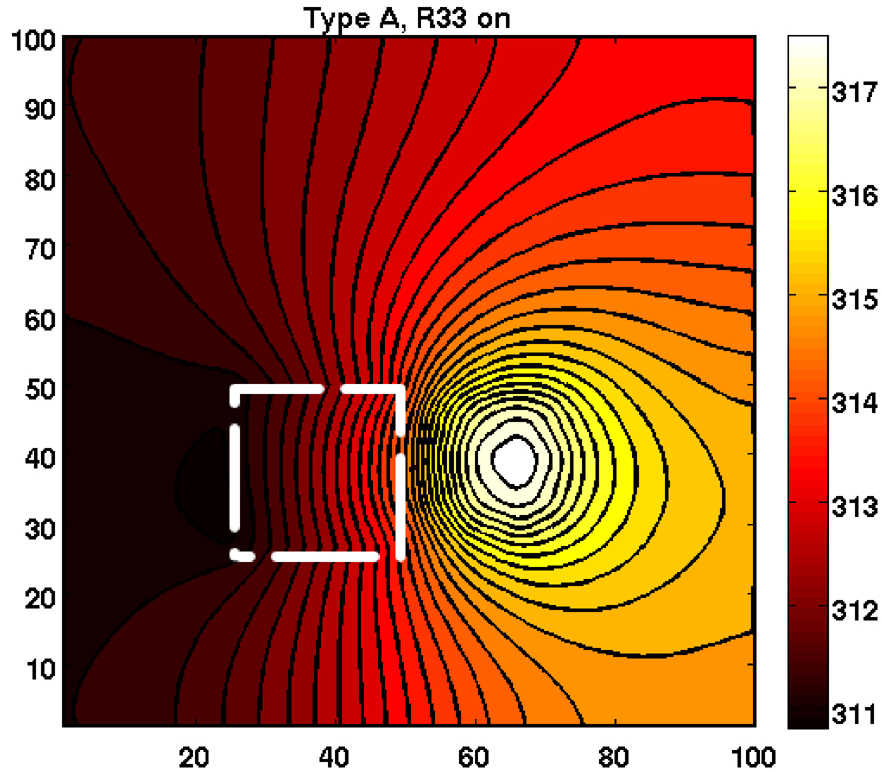


Figure 4.24: The temperature distribution when a heater outside of and to the east of the via frame is operating. The via frame, consisting of poles, is marked with a dashed white outline.

thick active layer, as opposed to bulk silicon die which could be 200 μm or thicker. While this reduces the vertical thermal resistance to ground through the substrate, it increases the lateral resistance at the same time. Since the thickness in our example here changes by a factor of 4000, the resulting increase in lateral resistivity might mean that the semiconductor is no longer the dominant lateral thermal conductor in the system.

To investigate this question, we modified our simulator to implement the Type A and B chips as though they were SOI chips. We used the same layout, but modified

the lateral and vertical resistivities used for the substrate. As expected, the lateral metal networks of the Type A and B chips now influenced the substrate thermal distribution. The simulation results are presented in Figures 4.25 and 4.26. The first thing to notice is the vastly increased peak temperature; with the same heaters operating at the same power level, the SOI version of the Type A chip exhibits a peak temperature of 656.29 K, and the Type B chip, 609.06 K. The next noticeable difference is that the asymmetry in the metal networks of the Type A and B chips now affect the substrate thermal distribution. In both results, the heat spreads in the low-resistance paths, which in the relatively uniform Type A layout is towards the nearest edges to the heater (west and south) but in the Type B is towards the north and east, as the extra metal network overlays that region, as marked as a shaded overlay in Fig. 4.27.

We can conclude that in SOI technologies, ordinary lateral interconnect metal layouts will in fact affect temperature distributions and their geometry needs to be taken into account for thermal modeling.

Finally, when the same via frame used for the simulation of Figs. 4.20 is implemented in the SOI version, the peak temperature falls to 572.59 K and the temperature rise is nearly completely confined to within the via frame. This result is presented in Fig. 4.28.

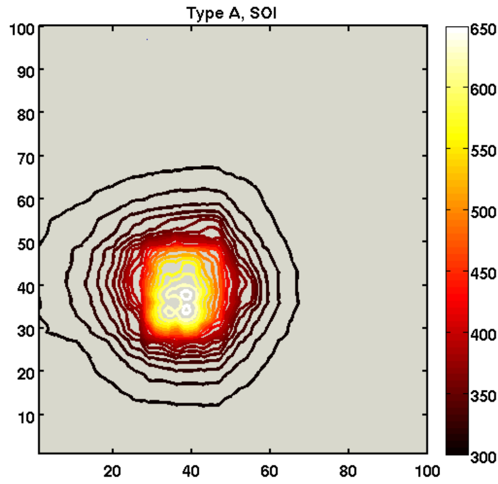


Figure 4.25: Simulation results for the substrate temperature in an SOI version of the Type-A chip (uniform metal network), with the same heater operating as the results in Fig. 4.18.

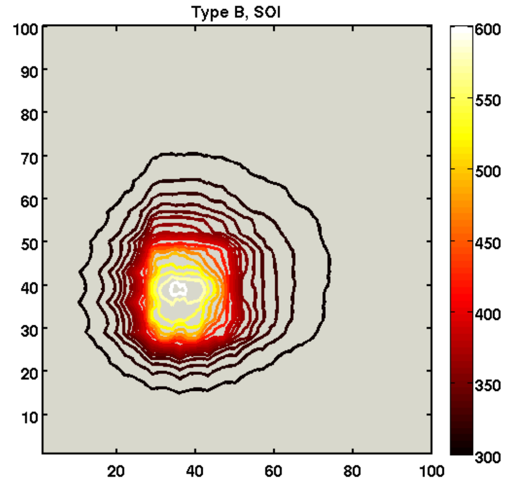


Figure 4.26: Simulation results for the substrate temperature in an SOI version of the Type-B chip (asymmetric metal network), with the same heater operating as the results in Fig. 4.19.

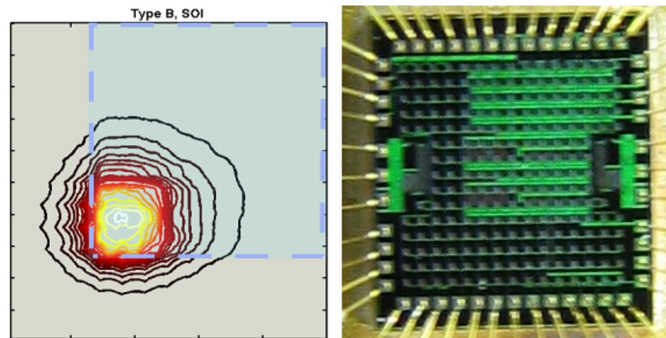


Figure 4.27: The lighter-colored region is the area on which there are the extra metal interconnects in Chip Type B, also visible in the chip photograph. It is clear from the contour lines that the thermal gradients change at the boundaries of this region.

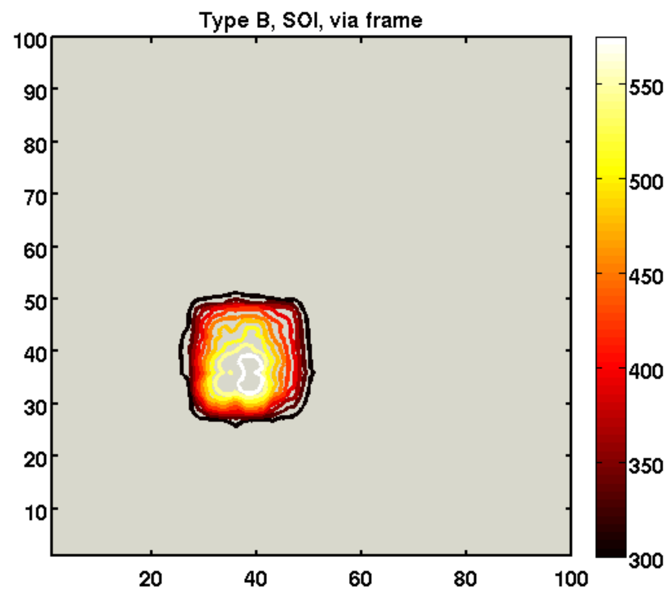


Figure 4.28: The SOI version of the Type-A chip simulated with a vertical via frame consisting of poles connecting the substrate to the ambient. The peak substrate temperature for the area under the substrate falls to 572.59 K and the areas outside the via frame experience no temperature rise.

4.6 Summary

In this chapter we focused on thermal effects on integrated circuits, which are critical to model for planar and three-dimensional ICs alike: To control overheating in modern systems with high power densities or little thermal conductivity to the ambient, or to ensure temperature-stable operating conditions for sensitive circuitry in a wide range of ambient conditions.

Given the heat source, the temperature distribution on an integrated circuit system can be determined by solving the heat flux equation. Discretizing this equation gives us a system of equations equivalent to that obtainable from applying KCL on a RC network (R-and-sources-only network in the steady-state case) which we can then solve. With this approach we examined the effects of non-uniform, non-isotropic thermal conductivity in chips and expanded our thermal modeling code to implement the layouts of all the different material layers of the physical system of integrated circuits. The code is just as readily expandable to include the individual layers of each tier of a stack of chips in a 3DIC.

To validate the simulator, we designed and had fabricated integrated circuits with individually-addressable on-chip heater and sensor arrays. We matched our measurements with simulation results by calibrating the thermal conductivities through the package to the ambient and the substrate lateral thermal conductivity.

We then investigated methods of controlled heating and cooling with thermal via frames. We found that thermal via frames connected directly to the ambient are effective as cooling factors both in bulk and SOI technologies. Furthermore,

while a lateral interconnect network does not have much effect on temperature distribution of bulk ICs, the vertical via frame somewhat contains thermal distribution through the substrate to within the area it encloses. In SOI technologies the lateral metal network does have a directional effect on thermal distribution throughout the substrate. The vertical via frame has the same containing effect in this case.

Chapter 5

Physical Aspects of VLSI Systems: Reactive Structure Design and Parasitic Capacitive Load Effects in Integrated Circuits

5.1 Introduction, Motivation and Background

Radio-frequency circuits have not been exceptions to the rule of miniaturization, nor to the trend of increasing integrated circuit complexity. Mobile communication devices drive an immense and growing industry, and lower power consumption with better portability are always the goals. Smart dust systems, intended to form local-area networks out of many such units for sensing, analyzing and surveillance purposes [59], also require integrated RF circuits for communication.

A specific challenge offered by radio frequency circuits is their need for passive elements like capacitors, inductors and transformers. While traditionally space-consuming, with rising operation frequencies it is increasingly possible to move these devices on-chip [34]. Research on the construction and modeling of such passive structures in has intensified since the latter half of 1990s [35, 36]. As 3-D integration opens up new geometric degrees of freedom in layout and is also suited for system-on-a-chip applications, it is to be expected that methods to migrate the construction of on-chip passives to 3DICs will be a topic of active research as well.

We start this chapter by presenting the basic concepts of on-chip inductors and

their physical features. We present and explain on-chip inductor behavior linked to the characteristics of the lossy metal-oxide-semiconductor system. We introduce design guidelines and present measurement results from devices we have designed and had fabricated. Next we bring up the subject of tunable LC tanks in relation to some experimental results we obtained and their electrical interpretation. Finally, we turn to parasitic capacitive load effects of off-chip connections in multi-chip systems, and demonstrate the extra power and speed costs of such loads, which might be avoided by three-dimensional integration.

5.2 Basics of On-Chip Inductors and Transformers

5.2.1 The Concept of Inductance

In circuit-theory terms, the *inductance* of an inductor links the potential "induced" across the device to the time rate of change of the current flowing through the device:

$$V(t) = L \frac{dI(t)}{dt} \,. \tag{5.1}$$

This is related to the physical concept of inductance between two current loops: A measure of the electromotive force induced in one loop is proportional to the rate of change of the total flux linkage set up over that loop by the current in the other loop [124, 125]. Thus if current I_2 in loop 2 causes the magnetic flux Φ_{12} in loop 1, $L_{12} = \Phi_{12}/I_2$ is the *mutual inductance* between these two current loops. This magnetic flux can be found by integrating the flux density $\vec{\mathcal{B}}$ over the flux linkage surface. The magnetic field of the current flowing in a loop also gives rise to a flux

linkage on the loop itself, which creates the *self-inductance* of the current loop. Part of the flux induced by the current flowing in a conductor is within the conductor, creating the *internal self inductance*.

Although we have so far focused on current loops, an important concept in the study of on-chip inductors is *partial inductance*, which can be calculated between current segments [124].

5.2.2 Physical Design

On-chip inductors and transformers are built by using the metal layers available in the semiconductor process. Like the components used in microwave electronics and built on a dielectric substrate, the prevalent geometric design is a spiral. Sometimes octagonal or other polygonal spirals are used, and another common style is a square spiral, as shown in Figure 5.1.

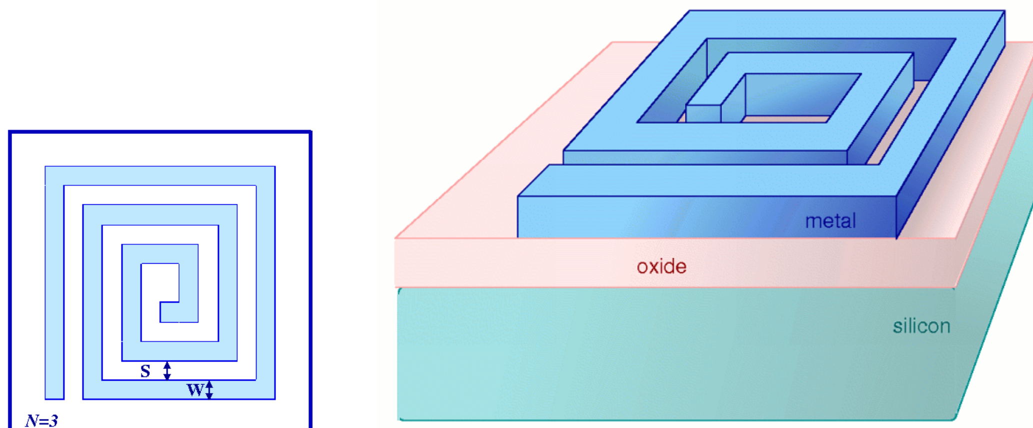


Figure 5.1: Top and side views of typical square spiral on-chip inductors.

This design style allows several degrees of freedom for a planar inductor:

- Number of turns,
- Total length of inductor traces,
- Length of the first or last segment,
- Width of inductor traces (W),
- Spacing between turns (S) (Setting three of the parameters listed up to this point uniquely determines the other two,)
- Metal layer used (This sets both the oxide thickness to substrate and the thickness of the metal strip,)
- Substrate doping (Selection of the in-substrate layer among the choices available in the process.)

An additional degree of geometrical freedom is obtained by using several metal layers for a stacked inductor geometry [35, 126, 127]. Finally, the insertion of a flat or patterned ground plane between the inductor metal layer and the semiconductor substrate has been suggested [128].

For on-chip transformer designs, common options are two spirals on different metal layers laid out on top of each other; three spirals, two in series, laid out on top of each other; spirals placed around each other or interwound spirals [40, 129, 130]. All of the above degrees of freedom apply.

5.2.3 On-chip Inductor Behavior and Design Guidelines

On-chip spiral inductors are nonideal inductors. Due to the metal layer resistivity, they exhibit non-zero resistance. They are capacitively coupled to the

substrate they rest on, and to other metal structures nearby. Their inductance arises from the inductive coupling between their turns and segments; however, these are also capacitively coupled to each other. Although not shown in Figure 5.1, an "airbridge"—a metal interconnect on a lower or higher layer—is required to connect to the end of the inductor at the spiral center. The capacitance between this bridge and the loops is the dominant parasitic capacitance for a planar inductor [131].

Unlike dielectrics that planar inductors are commonly laid on in microwave circuit technologies, the semiconductor substrate is lossy. Depending on the frequency, eddy currents are induced and some energy is lost due to substrate resistance; there is also substrate capacitance representing the high-frequency capacitive effects in the silicon. In these aspects on-chip inductor behavior arises from the lossy metal-oxide-silicon microstrip behavior, which is extensively studied [15, 19] and could be illustrated by a model such as that in Figure 5.2.

In the literature about on-chip RF passive inductors, the device inductance is usually defined as the imaginary part of the device impedance. The calculation method is to take two-port scattering (S-) parameter measurements of the fabricated device with a vector network analyzer, convert these to impedance and admittance parameters [132], and extract the inductance and the quality factor [40, 133]:

$$L = \frac{\mathcal{I}m\{1/Y_{11}\}}{\omega}\,, \tag{5.2}$$

$$Q = \frac{\mathcal{I}m\{1/Y_{11}\}}{\mathcal{R}e\{1/Y_{11}\}}\,. \tag{5.3}$$

As mentioned, the on-chip inductor features parasitic impedances and shows the effects of a number of series and shunt impedances and admittances. The induc-

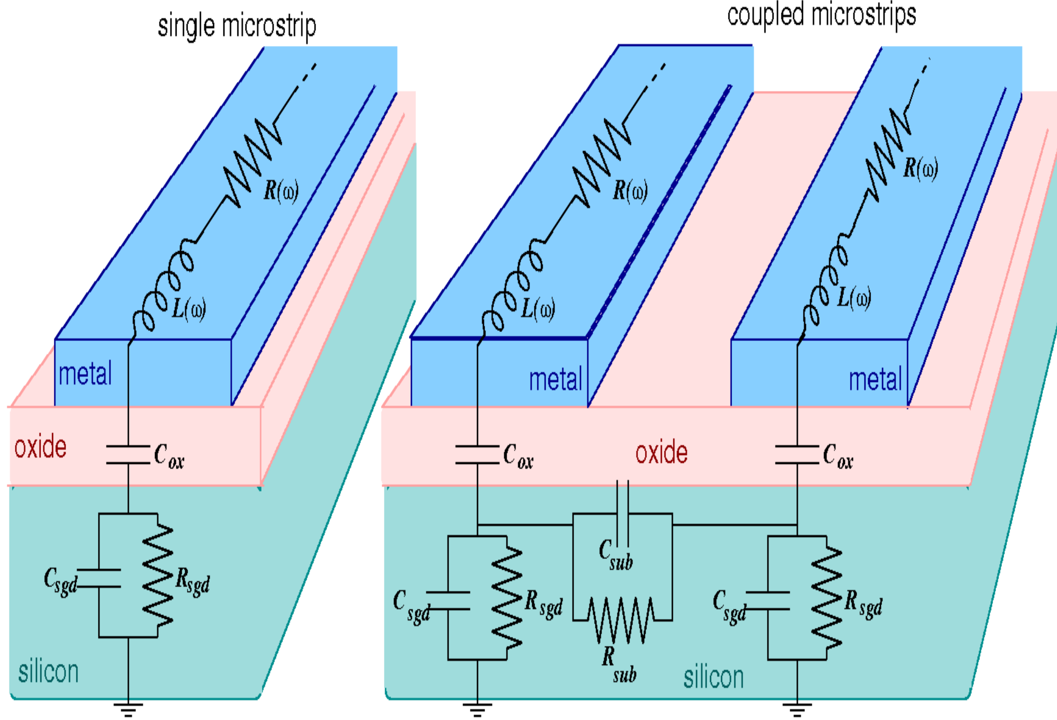


Figure 5.2: Lumped-element models for a single microstrip and coupled microstrips on a lossy substrate, which can represent the metal-oxide-semiconductor system of on-chip inductors.

tance of such a typical structure defined by Eqn. 5.2 exhibits the kind of behavior we measured and depicted in Figure 5.3: From low to high frequency, we have inductive, self-resonant, and capacitive regions. The $L = 0$ point where the imaginary part of the impedance is zero is the self-resonant frequency, or f_{sr}, of this device.

This response has several features important to circuit designers: Mainly, the low-frequency inductance and the self-resonant frequency (as the indicator of the frequency range in which the inductor is usable as an inductor). Related are other design concerns such as the device area and the parasitic coupling between the in-

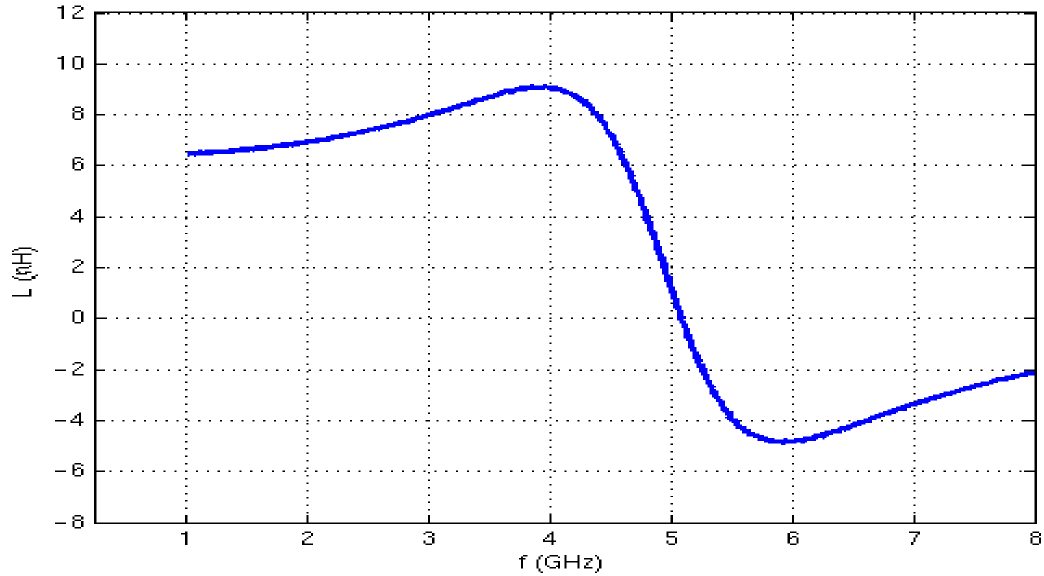


Figure 5.3: The measured inductance of a typical on-chip spiral inductor vs. frequency, as defined by Eqn. 5.2.

ductor spiral and nearby structures. The quality factor is also significant in several aspects of RF circuit design, depending on the application: filter bandwidths, oscillator frequency stability and noise sensitivity, and loss in matching networks are all affected by Q.

To gain insight into how the physical features of the inductor will affect the inductor behavior, we use a simple lumped-element model featuring a series inductance (L), series resistance (R), oxide capacitance (C) and substrate resistance (Rsub). Figure 5.4 displays the circuit, while Figure 5.5 shows how the inductance curve shifts with each changing parameter.

Briefly, the series inductance in the model determines the low-frequency inductance, the resonance peak and the self-resonant frequency (top left corner). The

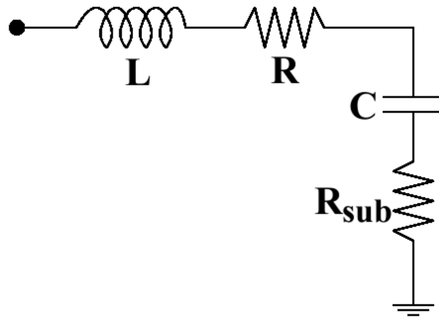


Figure 5.4: A very basic lumped circuit model which gives the same sort of impedance response with the same three regions of an on-chip inductor response.

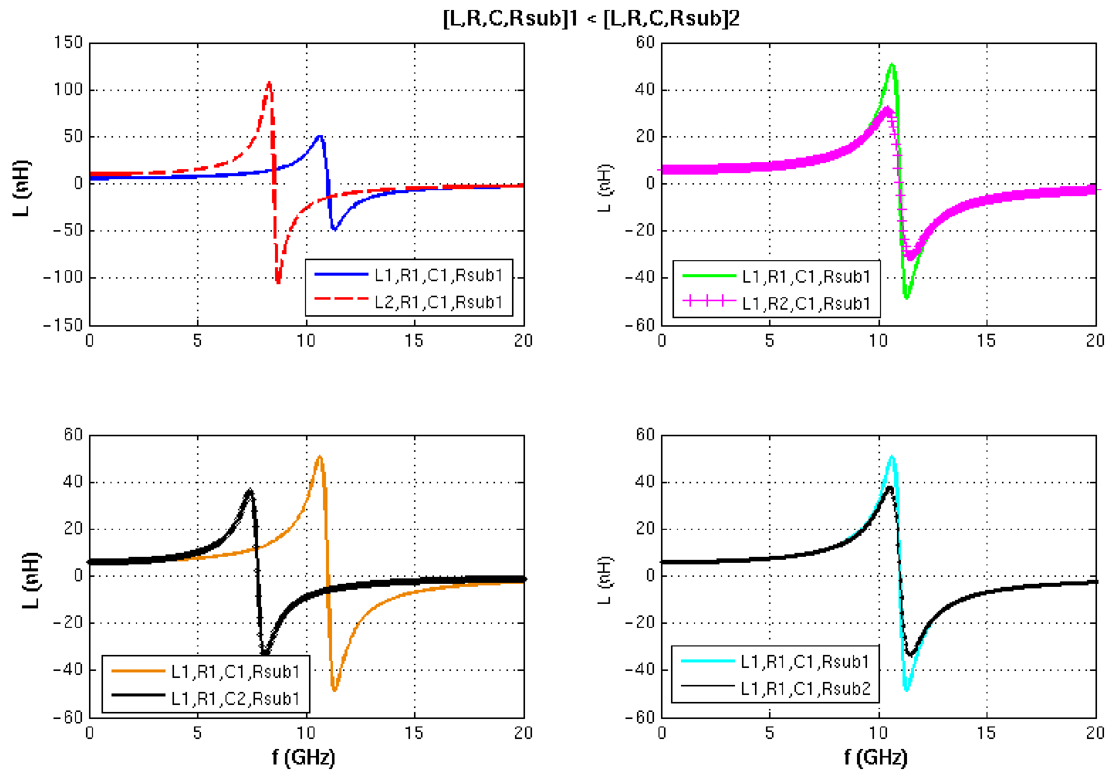


Figure 5.5: Changes in the elements of the simple lumped circuit of Fig. 5.4 reflected by changes in its inductance curve.

series resistance (of the metal) and the shunt resistance (of the substrate) both cause a softer resonance switch when increased (top right and bottom right corners). The shunt capacitance, when increased, reduces the self-resonant frequency and the resonance peak (bottom left corner). It is worth noting that the ratio of the self-resonant frequencies for the curve pairs is determined by the ratio $f_{sr2}/f_{sr1} = \sqrt{L_1C_1/L_2C_2}$.

In Section 5.4.2, we will briefly outline other modeling methods for on-chip inductors. Combining insights obtained from these various modeling methods and basics of inductance concepts, it is possible to come up with guidelines for inductor design. We give some samples here.

- Current segments in opposite directions yield negative partial inductances, while those in the same direction yield positive mutual inductances [124]. Furthermore, when designing an inductor with multiple loops the current directions should be arranged such that the resulting magnetic fluxes should point in the same direction, for higher mutual inductance.
- Longer conductors yield higher inductance but also higher serial resistance, which lowers Q. The same effect makes wider tracks preferable, as the serial resistance decreases; however see below.
- Increased oxide capacitance and increased capacitance between metal segments both reduce self-resonant frequency and Q. Therefore planar inductors built on higher metal layers are preferable, as they yield higher Q. Further, this creates a trade-off for increasing the track width as suggested previously.
- Small separation between metal tracks increases inductance due to higher mutual coupling; however, high frequency capacitive coupling between segments

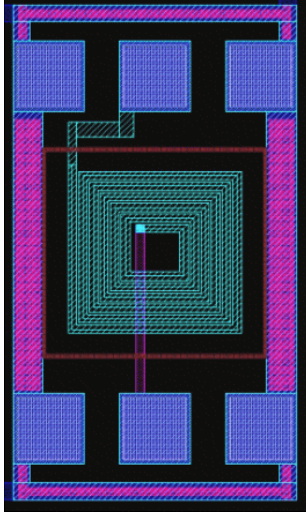


Figure 5.6: An on-chip inductor within a probe-pad structure designed for two-port S-parameter measurements.

reduces Q. Increased proximity effect may also cause the series resistance to rise more sharply with frequency and increase the resistive loss.

There is a sizeable body of literature investigating different configurations for both inductors and transformers [35, 36, 39, 40, 126, 127, 128, 129, 130]. We have concentrated on investigating the effects of some different geometrical shapes for these devices. Some results are presented in the sections ahead.

5.3 Experimental Design and Measurement Results

5.3.1 Experimental Setup and Methodology

As described in Section 5.2.3, we take S-parameter measurements using a vector network analyzer. To be able to do so using direct RF probing, we lay our structures out placed within probe pad structures designed for GSG (ground-signal-ground) type probes that are available to us. A typical inductor structure within such a pad setup is shown in Figure 5.6.

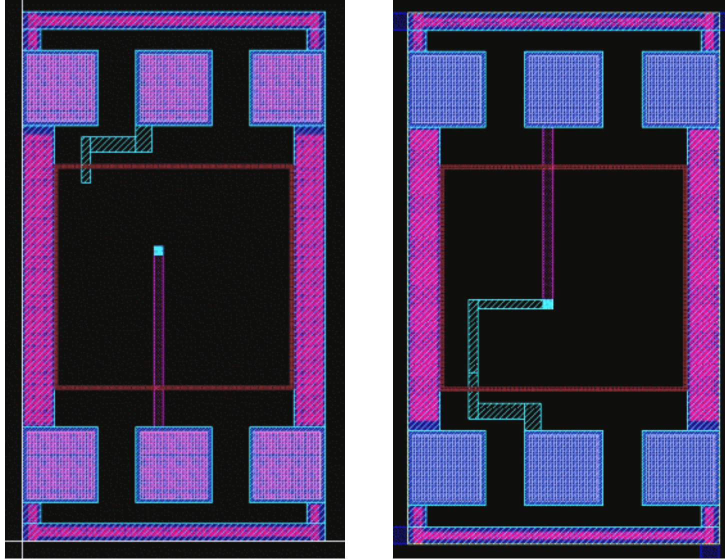


Figure 5.7: Open and thru de-embedding structures for on-chip measurements.

The probe size constraints mean that the parasitic impedance of these structures is considerable, and some de-embedding procedure must be followed. Therefore we also have "open" and "thru" pad structures (Figure 5.7) to be used in de-embedding. We have also written data analysis software to relatively automatize the de-embedding and inductance/Q extraction processes.

5.3.1.1 De-Embedding

The de-embedding process works by converting the scattering parameters of measured devices to impedance and admittance parameters as appropriate to extract the effect of the measuring pads [132, 134, 135]. A schematic illustration of the process is given in Figure 5.8.

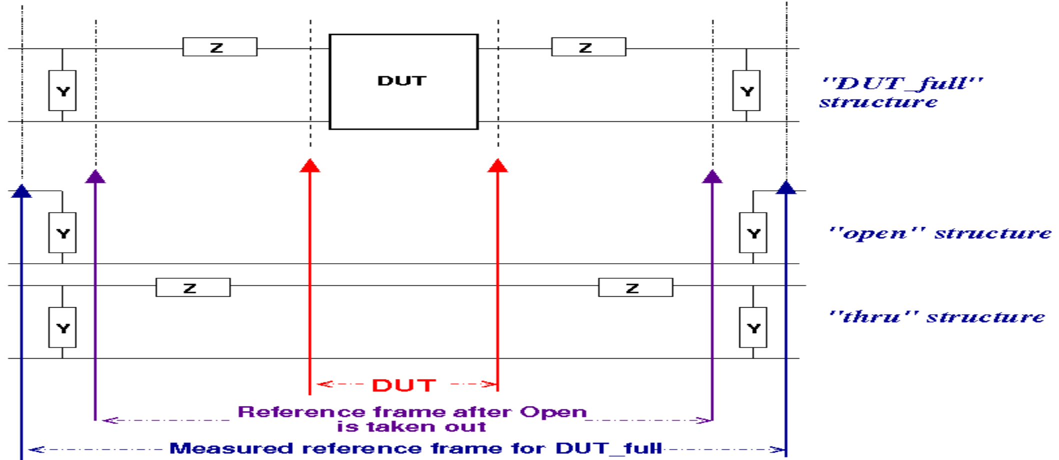


Figure 5.8: De-embedding.

5.3.2 Inductor Measurements

We have designed several chips and fabricated them through the MOSIS production facility. We designed the probe pads to be able to use RF probes to take on-chip two-port measurements with a vector network analyzer. We used Infinity probes from Cascade with a HP 8722D network analyzer. We will present some of our results here.

- First, the effects of de-embedding on the inductance plots is displayed in Fig. 5.9. As the shunt capacitance of the probe pads are subtracted, the self-resonance frequency shifts higher; it is also possible to observe the extra serial inductance of the pad connections and pads being removed.
- To illustrate one of the guidelines presented in Section 5.2.3, the measured inductance and Q-factors of three inductors with the same geometrical layout but on three different metal layers (M1, M2 and M3) are presented in Fig. 5.10. The trace geometries of all three are the same, yielding very close low-

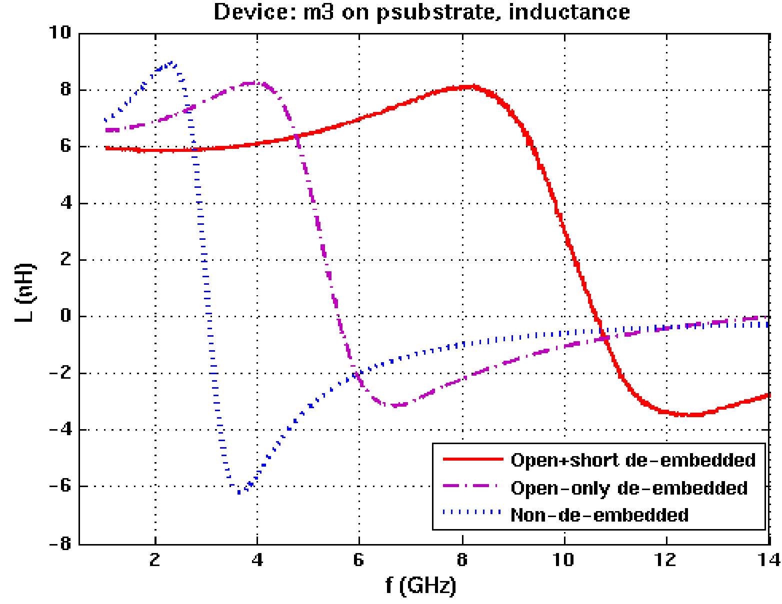


Figure 5.9: Results of the de-embedding process.

frequency inductances and serial resistances. They are built on the same type of substrate, causing similar substrate resistances as well, although the eddy currents induced necessarily differ somewhat. The most significant difference between the three is the capacitance to the substrate, clearly reflected in the resonant frequency f_{sr} shift to higher frequencies for the higher metal layers. As expected, the M3 inductor exhibits the highest quality factor.

- Figure 5.11 shows the planar inductor built on Metal 3 alongside a stacked inductor built using three metal layers. A side-view of this inductor and the current flow direction are also depicted. The total metal length is the same for both; however the planar inductor takes up 58072 μm^2 of silicon space for six turns whereas the 3-D stacked inductor takes 22500 μm^2 for nine turns. In addition to the greater number of turns, the stacked inductor has the advantage

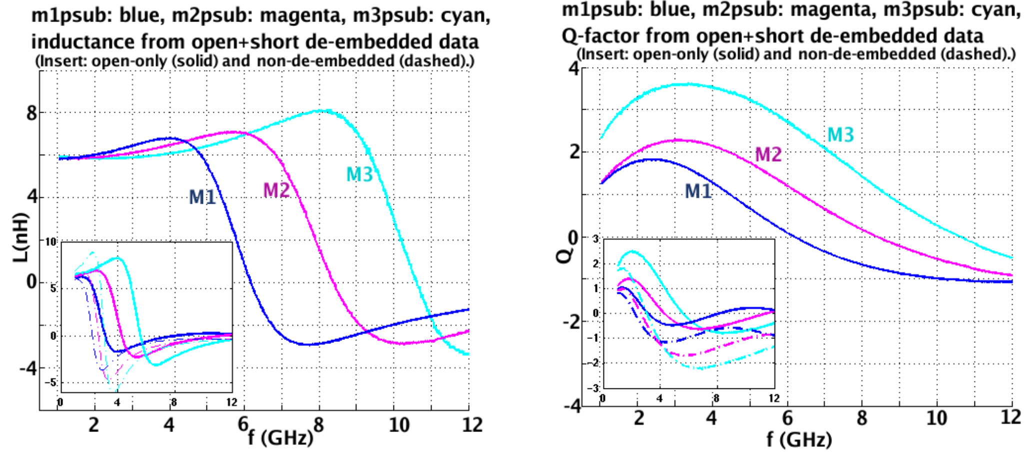


Figure 5.10: Inductance and quality factor plots for inductors built on p-substrate using different metal layers. (M3 (“m3psub”): Metal 3, M2 (“m2psub”): Metal 2 , M1 (“m1psub”): Metal 1.)

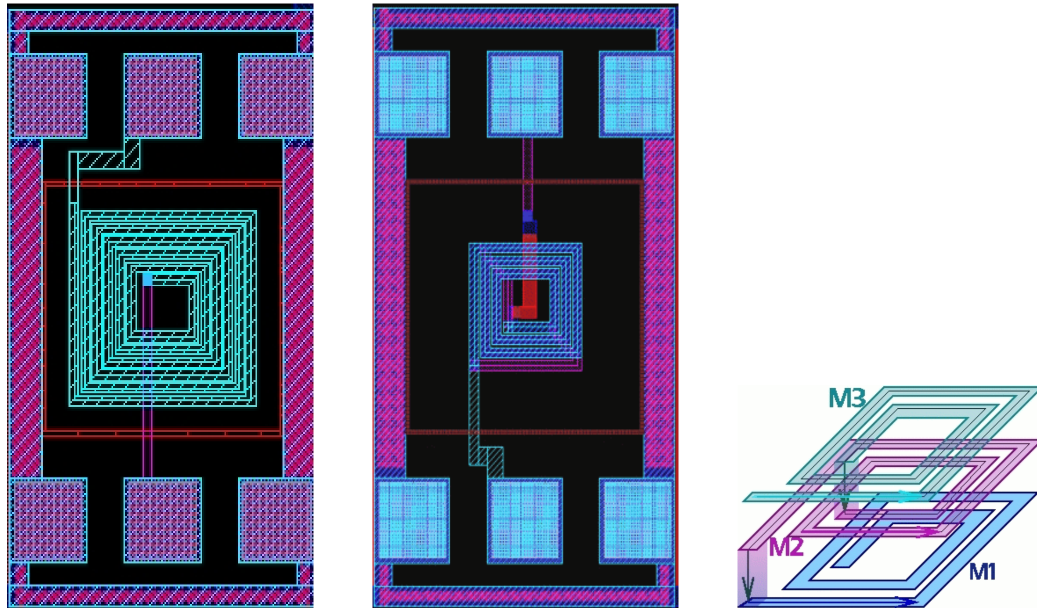


Figure 5.11: Layouts for planar and stacked inductors; schematic representation of the latter.

of clumping more segments with parallel current directions together, though this also raises the segment-to-segment capacitance. The net result, higher inductance with lower f_{sr}, is presented in Figure 5.12.

- Another way of exploiting multiple metal layers is attempting to emulate the

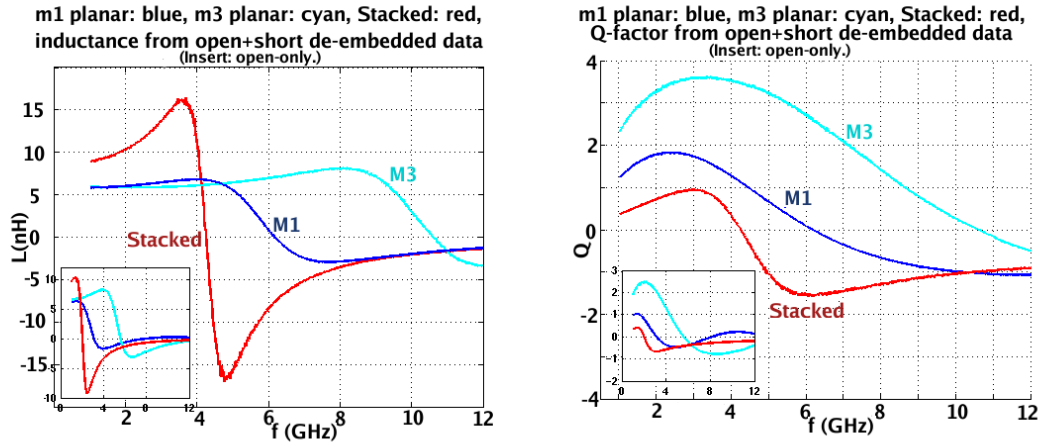


Figure 5.12: Inductances and quality factors of planar and stacked inductors. (M3, M1: Planar Metal 3 and Metal 1.)

macro inductor coil (wound-around-a-core) using metal lines in a fabrication process. The basic geometry we devised with this aim is the first example in Figure 5.13. Variants to increase single "turn" areas and bring positive mutual-inductance segments closer while keeping negative segments farther apart are the next examples in this figure, coil2 and coil3. Compared to the latter two, coil1 exhibits a very low inductance; therefore Figure 5.14 shows the coil2 and coil3 layouts only. Figure 5.15 displays the inductances and Q-factors for these two inductors, layout area 47800 μm^2, as compared to the metal 3 planar inductor (area=58072 μm^2). From the results we observe that while the coil serial inductances are about one-third of that of the planar inductor, and the equal metal trace lengths yield similar serial resistances and therefore lower quality factors, it is interesting to point out that this low L translates to a higher self-resonant frequency. Coil2 especially seems to exhibit a comparable parasitic capacitance.

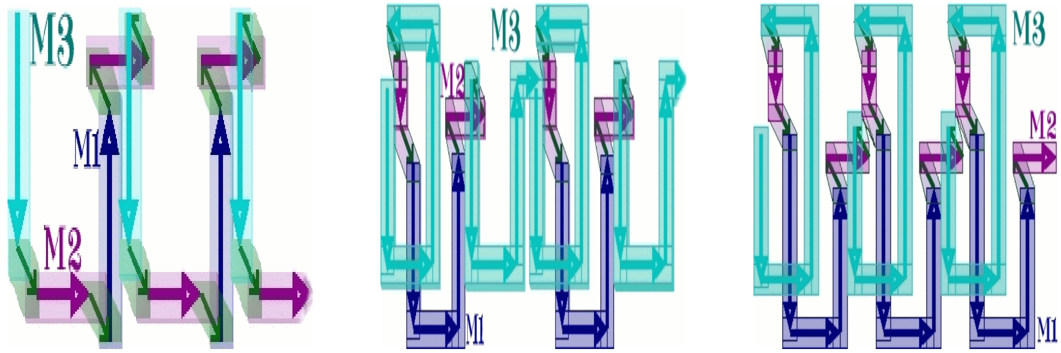


Figure 5.13: Schematic representations of the “Coil1,” “coil2” and “coil3”.

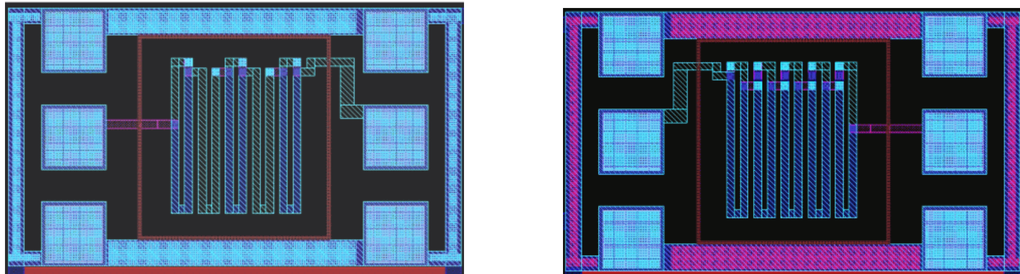


Figure 5.14: Layouts of the coil inductor variants coil2 and coil3.

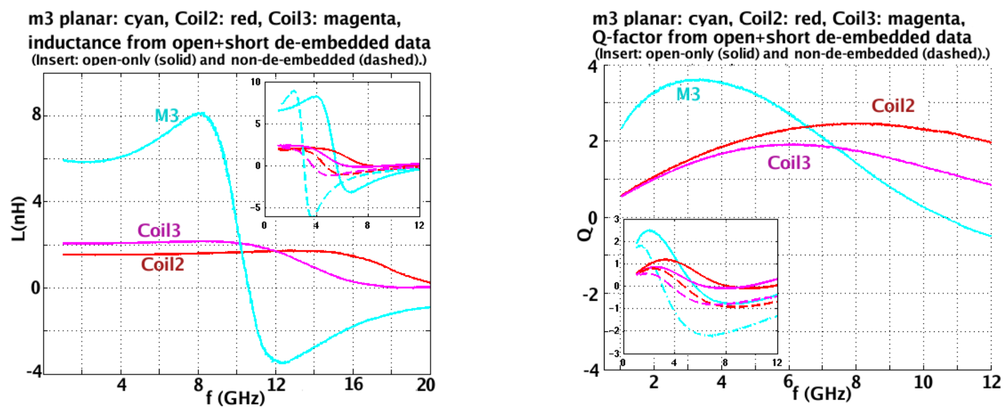


Figure 5.15: Inductances and quality factors for planar and coiled inductors.

(M3: Planar Metal 3.)

5.3.3 Transformer Measurements

Impedance conversions, feedback path couplers, and voltage/current gain are a few possible functions for on-chip transformers in RF circuits, whose design is also actively investigated in literature [39, 129, 130]. We have designed and fabricated some transformers with representative geometries.

Figure 5.16 displays the schematic representations for four different transformer designs: Two-metal stacked ("M2:M3"), three-metal stacked ("M2:M1M3"), interwound and spiral-within-spiral transformers, top left, right, bottom left and right respectively. Figure 5.17 shows the layout for the M2:M3 and the spiral-within-spiral transformers. We examine the reflection coefficients (S11), transmission coefficients (S12), effective "turn ratios" calculated by assuming a load of

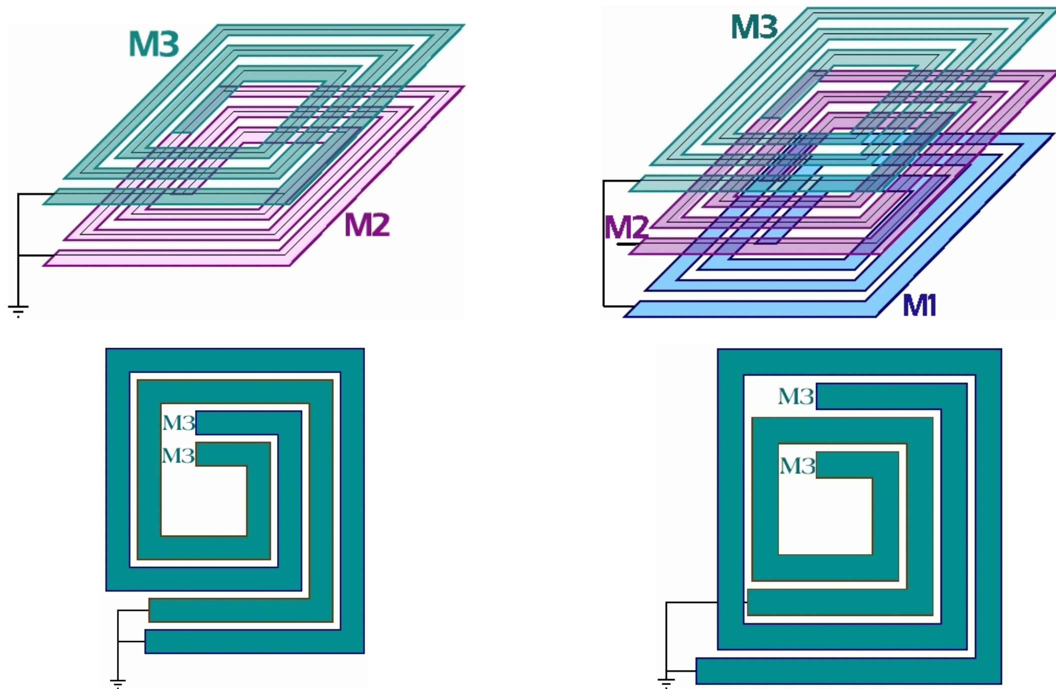


Figure 5.16: Different transformers: M2:M3 (metal 2 to metal 3), M2:M1M3 (metal 2 to metal 1+metal 3) , interwound, spiral-within-spiral.

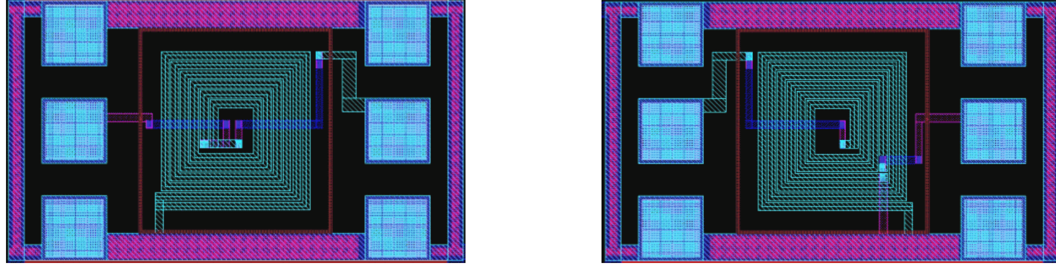


Figure 5.17: Layouts for the M2:M3 and spiral-within-spiral transformers.

50 Ω on the secondary side and using the expression $\sqrt{Z_{11}/50}$, and voltage gain calculated by starting from the impedance matrix for a transformer, replacing V_2 and I_2 by $(N_2/N_1)V_1$ and $(N_1/N_2)I_1$ respectively, and solving the resulting system of equations for $N := N_1/N_2$.

Figures 5.18 and 5.19 present these results for the M2:M3 and M2:M1M3 transformers. Figures 5.20 and 5.21 show the same for the interwound and spiral-within-spiral transformers. The most efficient device for coupling is the M2:M3 transformer, followed by the two planar transformers and the M2:M1M3 transformer, which last naturally exhibits the highest reflection. The coupling efficiency of the two planar transformers are more dependent on frequency than the stacked transformer geometries. The effective "turn ratio" is observed to be highest for this last device, however, as could be expected from comparing the designs. Therefore, if a high voltage gain is preferred over efficient coupling, this would be the design of choice.

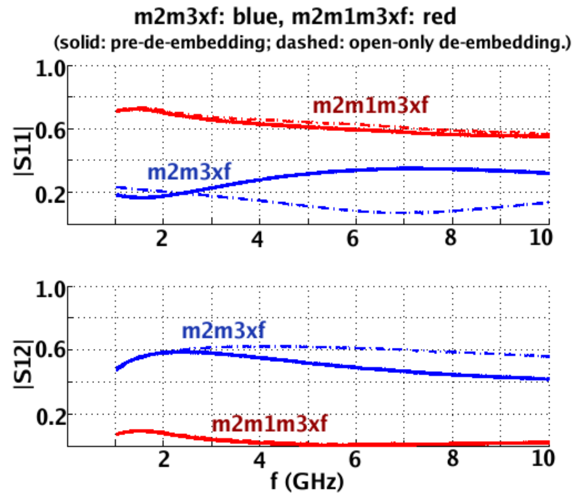


Figure 5.18: Reflection and transmission coefficients, for the M2:M3 (“m2m3xf,” metal 2 to metal 3, blue) and M2:M1M3 (“m2m1m3xf,” metal 2 to metal 1+ metal 3, red) transformers.

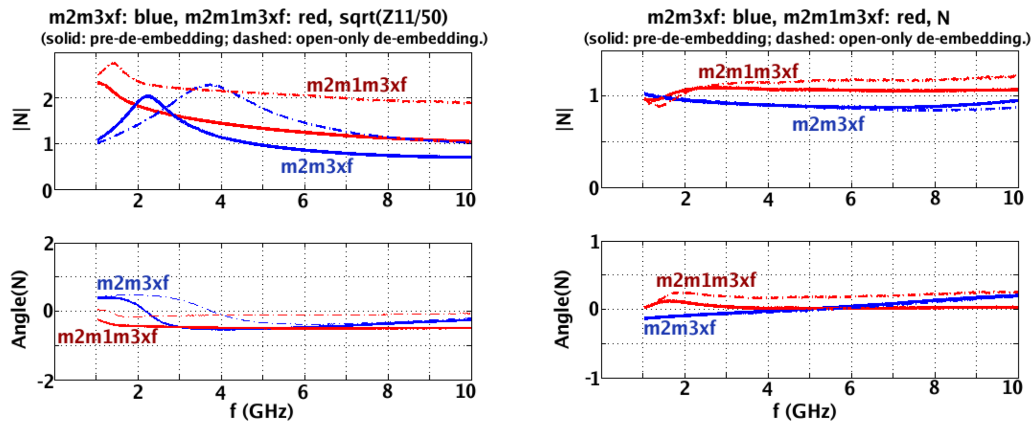


Figure 5.19: “Turn ratio” (left) and voltage gain (right) for the M2:M3 (“m2m3xf,” metal 2 to metal 3, blue) and M2:M1M3 (“m2m1m3xf,” metal 2 to metal 1+ metal 3, red) transformers.

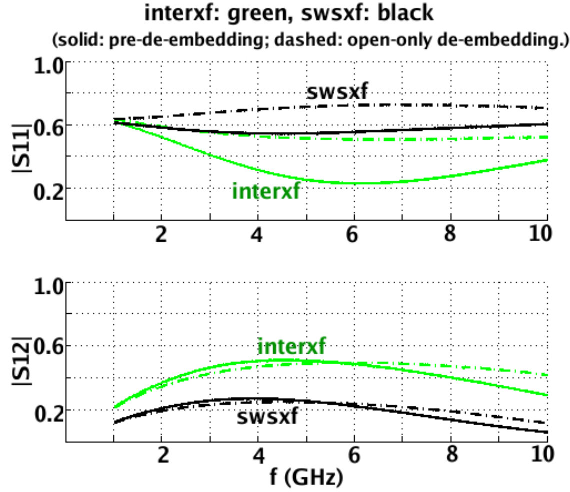


Figure 5.20: Reflection and transmission coefficients for the interwound (interxf, green) and spiral-within-spiral (swsxf, black) transformers.

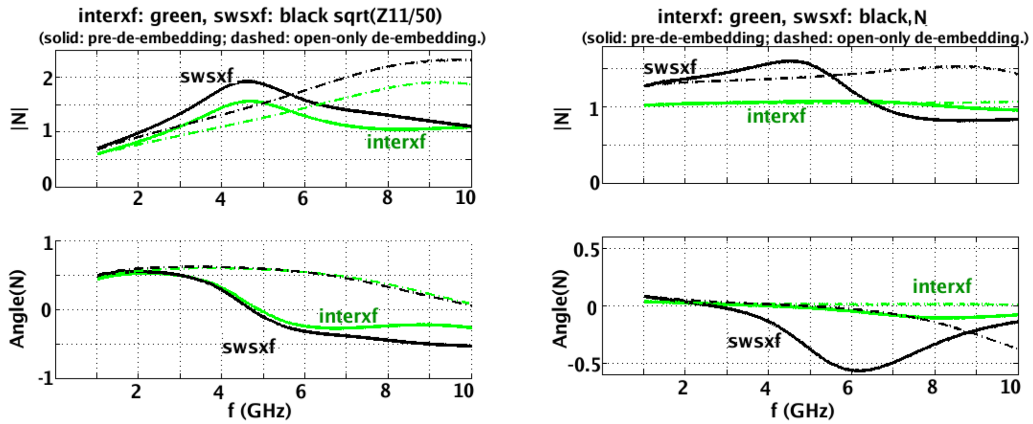


Figure 5.21: "Turn ratio" (left) and voltage gain (right) for the interwound (interxf, green) and spiral-within-spiral (swsxf, black) transformers.

5.4 Tunable LC Tanks: Photoelectric Effect

In many RF circuits, especially communication circuits like LNAs and mixers, a self-resonant LC tank is used as a tuned load for ensuring a certain frequency response [40, 136]. Just by themselves, LC filters have been long in use and relatively recently adapted to IC fabrication [137]. The use of tunable inductors, capacitors and therefore LC-structures for the implementation of tunable matching circuits,

VCOs etc. have been proposed [41, 138].

There has been some investigation on purposefully integrating an inductor-structure with a capacitor-structure in series or shunt configurations. But it should be noted that the on-chip planar inductor is by itself a self-resonant structure. There have been studies of coupling between nearby inductors on the same substrate [5, 139], and it has been suggested that this coupling, if intentionally designed, might be utilized to tune the L and Q characteristics of on-chip inductors [40].

The results presented in this Section point out another possibility, that of changing the characteristics of the semiconductor substrate as a modulation method for a tunable LC tank.

5.4.1 Experimental Results

One interesting and promising observation during our experiments was how the inductive structures on a semiconductor substrate reacted to ambient light. When we turn on the microscope light while taking a VNA measurement, the inductance curves of all our inductors undergo a characteristic shift, illustrated in Figure 5.22.

We can interpret the effects of light qualitatively referring back to the simplified model behavior outlined in Section 5.2.3: Zooming on the low-frequency response shows a very slight ($\sim$0.1 nH) increase in serial inductance. The inductance zero-crossing points, or f_{sr}, have shifted down by 150 MHz for the device on p-substrate and 200 MHz for the device on n-well (see Figure 5.23). Also, the peak impedance has increased, indicating higher Q (see Figure 5.24).

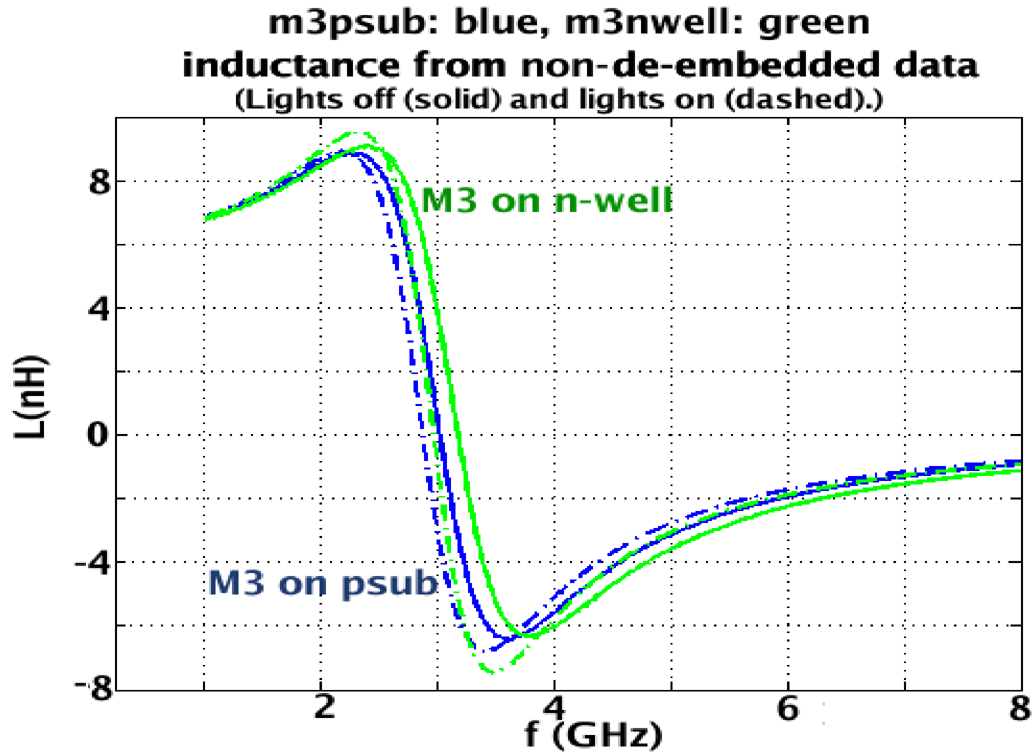


Figure 5.22: The measured inductances of two metal-3 on-chip spiral inductors, one on p-substrate and the other on an n-well. Solid lines: External light off. Dashed lines: Light on. This is data prior to de-embedding.

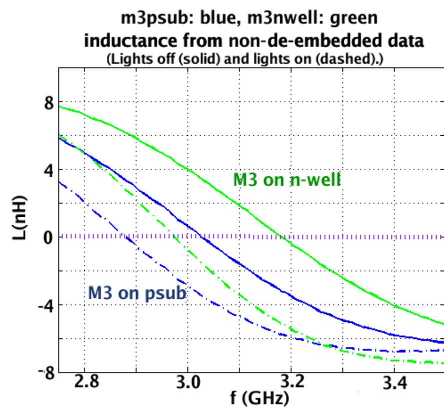


Figure 5.23: Zoom near the f_{sr} point of Figure 5.22.

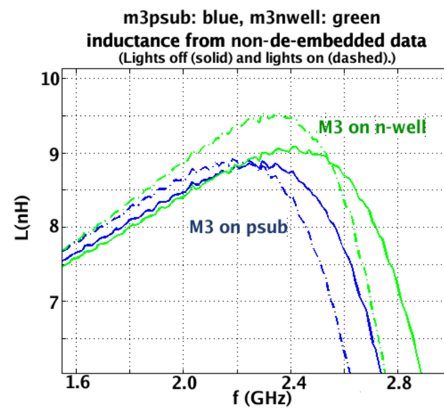


Figure 5.24: Zoom near the peak impedance point of Figure 5.22.

After de-embedding is done to negate the effects of the shunt impedance of the probe pads, the f_{sr} shifts are about 570 MHz for the p-substrate device and 1.14 GHz for the n-well device.

5.4.2 Interpretation and Modeling

To explain the resonant frequency shift in more detail, we adapted a more elaborate model for an on-chip inductor, proposed in [131] and also used by [140]. This model, shown in Fig. 5.25, takes into account the series inductance and resistance of the interconnect coil (L_s and R_s), capacitive coupling between loops and the center tap bridge (C_s), oxide capacitance (C_{ox}), and substrate capacitance and resistance (C_{sub} and R_{sub}).

These parameters can be calculated from the process parameters using the procedure given in our references. Some of the necessary process data for these

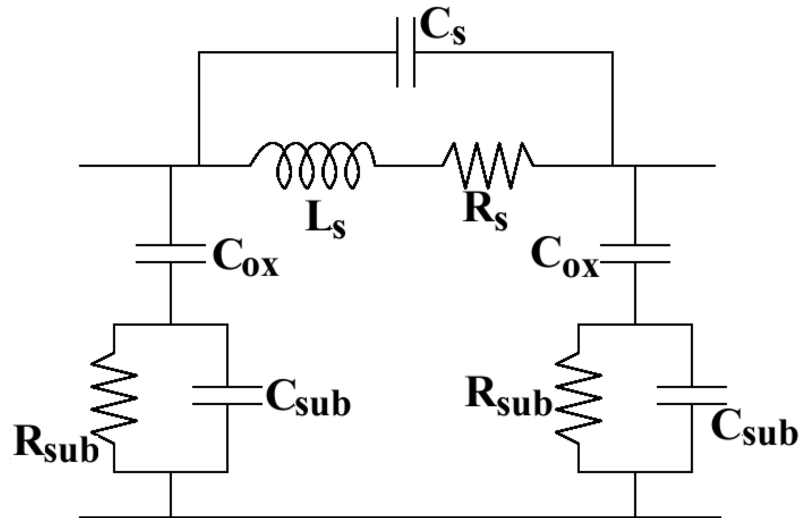


Figure 5.25: A more advanced lumped circuit model for on-chip inductors [131].

calculations, such as the inter-layer capacitances, were available through our fabrication facility MOSIS [101]. Not having access to proprietary information, we had to use our best guesses for some, such as the substrate and metal layer thicknesses. For use in our simulations, using the geometry of the M3 inductor featured above, we calculated L_s=6.0331 nH, C_s=28.5 fF without considering fringe capacitance and 42.3 with the fringe capacitance, C_{ox}=356.65 fF, R_{sub}=39.35 Ω and C_{sub}=2.23 fF. The serial resistor R_s depends on the frequency thanks to skin effect and proximity effects [140]. The expression we calculated for the skin effect dependency is $R_s(f) = 1.033 \times 10^{-4}\sqrt{f}/[1 - exp(-1.22 \times 10^{-5}\sqrt{f})]$.

Figures 5.26 and 5.27 display the resonant frequency shift when we use the s-parameter simulation function of Spectre to obtain the inductance vs. frequency curve of this model in "no light" and "light on" configurations. For the "light on" configuration, we assume the charge concentration in the substrate rises through photogeneration. Thus we treat the substrate conductivity as higher, thereby reducing substrate resistance. To obtain the shift pictured in the Figure, we divided R_{sub} by 1.5. The obtained f_{sr}-lowering effect follows that observed during the experiments in the previous section.

5.5 Concluding Remarks on On-Chip Inductors

With many innovations, such as ground planes, low-conductivity substrates (including non-silicon semiconductors), and more metal layers available farther up from the substrate in modern processes, the on-chip integration of inductive pas-

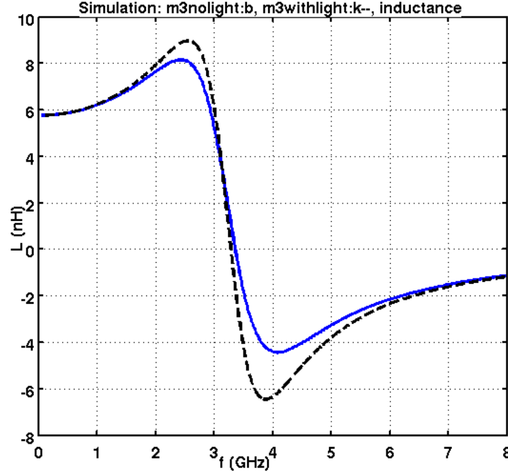


Figure 5.26: Spectre s-parameter simulation results for the lumped-circuit inductor model of Fig. 5.25 with the element values in the text. Solid line: "Light off". Dashed line: "Light on," with the substrate resistance reduced.

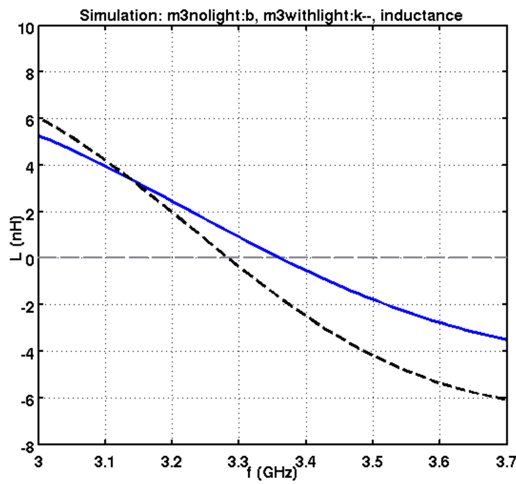


Figure 5.27: Zoom near the f_{sr} points of Figure 5.26.

sives for compact, system-on-a-chip communication and network applications will continue to develop. Certain modern processes even offer a metal layer set aside for this purpose, with a higher metal thickness to reduce serial resistance and increase the quality factor.

Beyond the model utilized above, there has been considerable work done in on-chip inductor modeling to facilitate the integration of inductors and transformers [129, 141] into the CAD flow. For instance, an enhancement to the model of Fig. 5.25 places a reactive branch parallel to the series resistance to model the serial element frequency dependency [139]. Segment-to-segment coupling through

the substrate can also be included [127, 142]. Model parameters are extracted from measurements and parameter-fitting, or calculated from the device geometry and parameters: Using partial element equivalent circuit methods [124] and a complex image method to include lossy substrate effects [19], we can obtain equivalent circuit parameters [5, 143]. Such element value calculations are also based on the classic methods developed in the 1940s and 1970s [15, 16, 125].

Finally, a method combining many of the features described above to create a segment-by-segment model of an on-chip inductor has been proposed and adapted for RF IC design tools [40, 144].

The inherent resonant nature of these devices, as demonstrated above, can be modulated and possibly put to use in circuits that would benefit from a tunable frequency response.

3-D integration, since it offers a higher geometrical degree of freedom, opens new possibilities for on-chip inductor design. We can propose a larger version of the coil-type inductors featured in Section 5.3.2, wrapped around one of the tiers as a "core." This would afford a wider coil area and a higher inductance in a small silicon footprint, but the higher parasitics and all-around proximity to a lossy substrate would create tradeoffs in quality factor—unless the design is implemented in an 3-D SOI process. In this latter case, we can also conceive a multi-tier stacked inductor with considerable gain in the inductance-to-surface area figure of merit.

5.6 A Study of Performance Improvement by Capacitive Load Reduction with 3D Integration

In Section 1.1, we observed that fewer lines needing to go off-chip to carry signals between different parts of an electronic system affords the designer benefits in avoiding extra load and speed and power loss. To demonstrate this point, we have measured the operation speeds of identical circuits integrated through off-chip bonds vs. on-chip bonds.

5.6.1 System Design

5.6.1.1 Measuring Bonding Pad Effects Using Ring Oscillators

The central structure in this system for measuring loading effects of bonding pads is a ring oscillator: A simple ring of an odd number of inverters feeding back into itself. Our system takes advantage of the fact that the operating frequency of a ring oscillator reflects the load on each of its stages, as we will review here.

A ring oscillator is a negative-feedback system. Once the inverters are powered, random noise in the input of one stage may cause the output of this stage to flip, which in turn switches the next stage, and so on; if the number of inverters in the chain is odd, all inverters keep changing states while being powered, causing the output signal tapped between any two stages to oscillate. The oscillation frequency of this signal depends on the total propagation delay of the circuit:

$$f_{osc} = \frac{1}{N(t_{PLH} + t_{PHL})}, \tag{5.4}$$

where N is the number of inverters. t_{PLH} and t_{PHL} are the high-to-low and low-to-high propagation delays respectively, measured between the 50% points of the input and the of the output. A five-stage ring oscillator is shown in Figure 5.28.

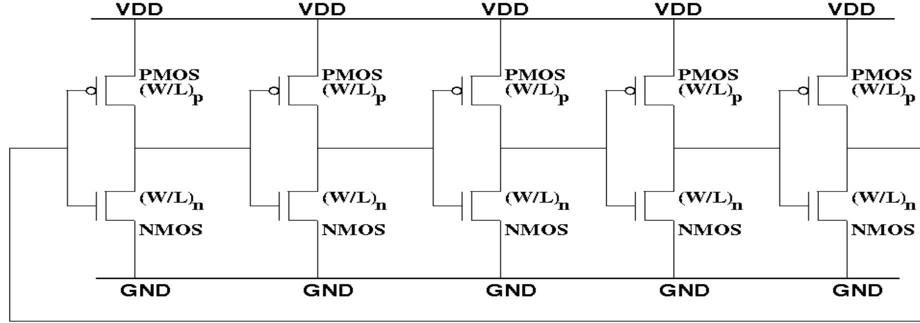


Figure 5.28: A five-stage ring oscillator.

Using the digital model for an inverter given in [12], shown in Figure 5.29, we calculate these propagation delays by considering the switching of the inverter as a load capacitor being charged through a large-signal equivalent switching resistance. In the figure, R_p and R_n donate these resistances. Note that the large-signal assumption implies the transistors are being taken to operate at either cut-off or saturation. For an inverter driving another inverter, a load capacitor of $C_{inn}+C_{inp}+C_{outn}+C_{outp}$, which is the net capacitance from the output to the AC ground, is being charged or discharged, through R_p to V_{DD} or through R_n to the ground respectively.

Therefore the low-to-high and high-to-low propagation delays are given by

$$t_{PLH} = 0.7R_p(2.5C_{oxn} + 2.5C_{oxp}), \quad t_{PHL} = 0.7R_n(2.5C_{oxn} + 2.5C_{oxp}), \tag{5.5}$$

and used in Eqn. 5.4 to find the operating frequency.

The key point for our demonstration is when there is an extra capacitive load between two of the stages, as shown in Figure 5.30. This will change the propagation

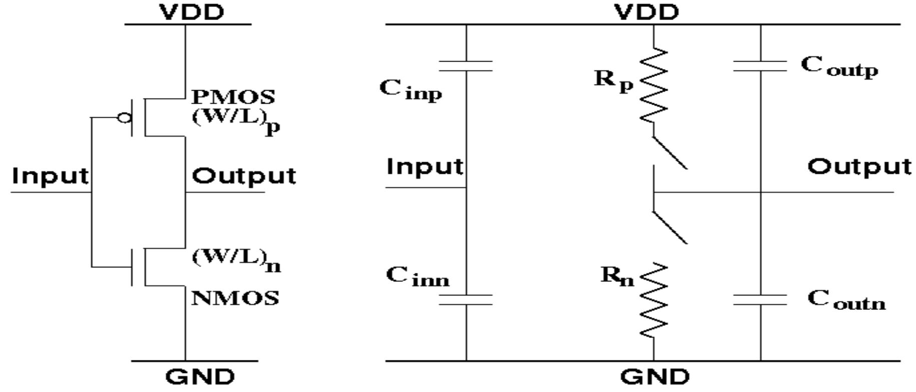


Figure 5.29: The CMOS inverter and its digital model. The input and output capacitances are proportional to the transistors' gate oxide capacitances: $C_{in} = 1.5C_{ox}$ and $C_{out} = C_{ox}$ approximately. The effective switching resistances are the average resistances of the relevant MOSFETs between the off and the on stages of the full swing: $R_{n,p} = V_{DD}/I_{Dn,p}$.

delay of the preceding stage:

$$t_{PLH,load} = 0.7R_p(2.5C_{oxn} + 2.5C_{oxp} + C_{load}). \tag{5.6}$$

Similarly for $t_{PHL,load}$. In this case, the total propagation delay for the N-stage ring oscillator becomes $(N-1)(t_{PHL} + t_{PLH}) + (t_{PHL,load} + t_{PLH,load})$.

5.6.1.2 Internal and External Ring Oscillators

We designed integrated circuits including two ring oscillator versions on a single die. Two versions of the chips were made: one using the AMI 1.6 μm (ABN) technology and other the AMI 0.6 μm (C5N) technology.

We designed the *internal ring oscillator* on each chip to have no extra load

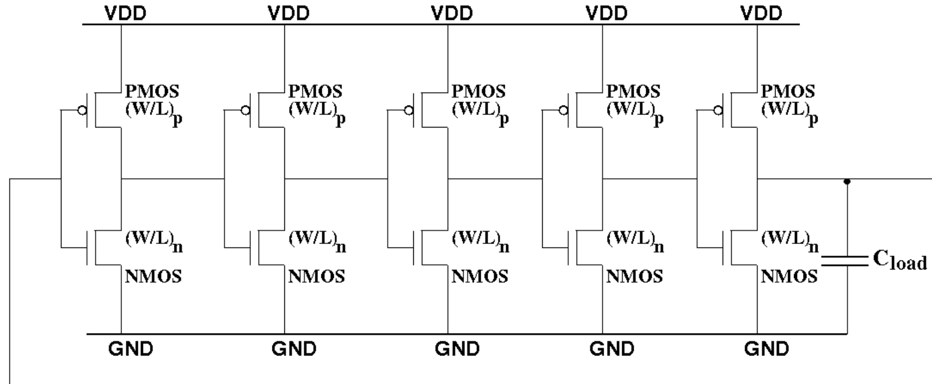


Figure 5.30: A five-stage ring oscillator with a capacitive load between two of the stages.

between stages: The output of each inverter is connected directly to the input of the next, with the "last" inverter driving the "first" to complete the circle. The C5N chip had a 31-stage internal ring oscillator, with the layout shown in Figure 5.31.

The output of each internal oscillator circuit went through an extra inverter, serving as a buffer, to a counter. In the C5N chip, this is a 6-bit counter comprised of D-flip-flops (which divides the output frequency by 64). We included the counter was to serve as an on-chip testing and diagnosis structure. As one advantage, it eliminated the need for the internal ring oscillator to drive outside loads directly; as another, it slowed down the expected operation frequency so that extremely fast measurement equipment would not be necessary. The layout of the oscillator and counter is shown on Figure 5.32.

We note that the buffer inverter acts as an extra capacitive load inserted between two stages, with $C_{load} = C_{inn} + C_{inp} = 1.5C_{oxn} + 1.5C_{oxp}$.

We laid out the *external ring oscillators* were as individual inverters with their

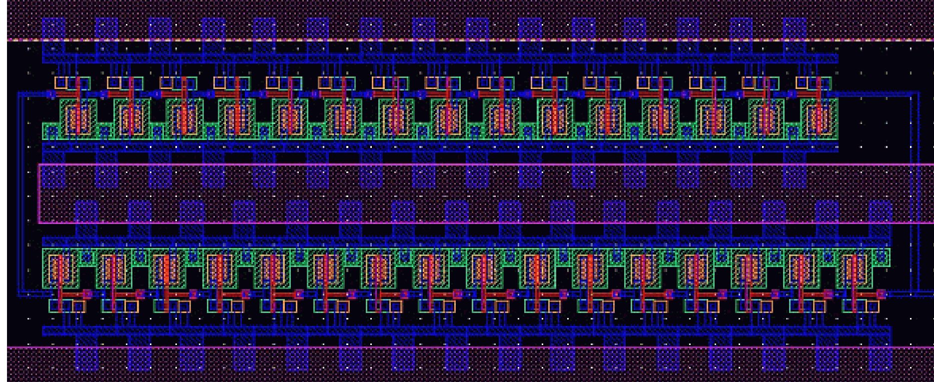


Figure 5.31: The 31-stage ring oscillator comprised of minimum-size transistors in the 0.6 μm technology chip. The output is taken from between the stages on the lower and upper right hand corners and goes through an additional inverter, which serves as a buffer (not shown).

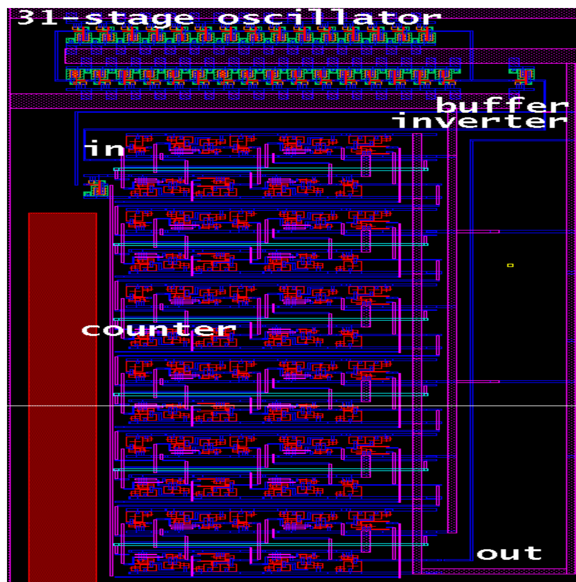


Figure 5.32: The layout for the 31-stage ring oscillator and 6-bit counter/divide-by-64 frequency divider.

input and outputs connected to separate bonding pads. We used standard bonding pads with ESD (electro-static discharge) protection, provided by MOSIS for these technologies [123]. The C5N chip had eleven inverters laid out for this purpose (three are shown in Figure 5.33). When these inverters are linked externally to form the ring oscillator, between each stage there is the extra load of two bonding pads,

bonding wires to package, and the out-of-the-chip connection elements. This was intended to simulate situations where two chips are integrated in a system and a signal has to go off chip to get from one part of the system to the other.

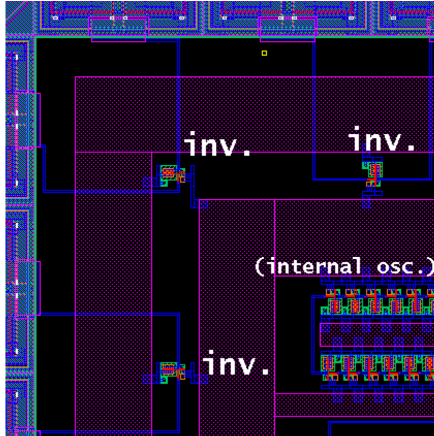


Figure 5.33: Three stages of the external ring oscillator. The bonding pad connections are visible on the top and left edges of the figure.

The full layout of the C5N chip is given in Figure 5.34. In addition to the internal oscillator, frequency divider and external oscillator, the chip had a pn junction diode and a large inverter as process check structures.

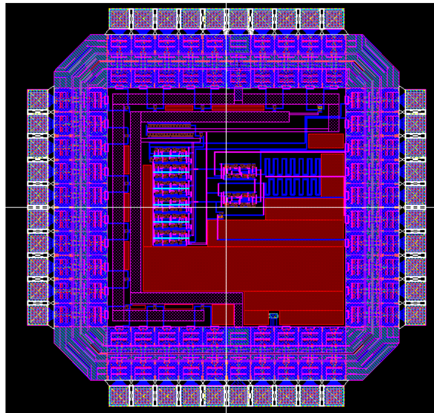


Figure 5.34: The full chip layout for the C5N chip.

5.6.2 Measurement Results

Comparing the performance of the 11-stage external ring oscillator with that of the 31-stage internal ring oscillator, we can conclude that the pads, pins and bread-

board connections caused nearly 800 times reduction in the operation frequency.

We first formed the 11-stage external ring oscillator on a breadboard and took measurements using a digital oscilloscope. Figure 5.35 shows the output waveform, with a frequency of 398.3 kHz.

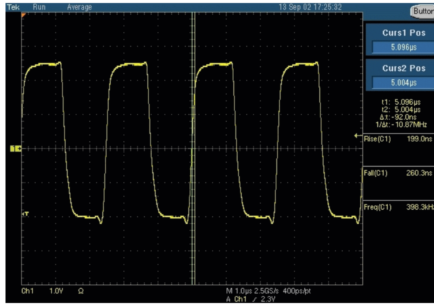


Figure 5.35: The output waveform of the 11-stage external ring oscillator.

The output of the divide-by-64 counter was then measured using the same scope. The output waveform, at 1.76 MHz, is shown in Figure 5.36.

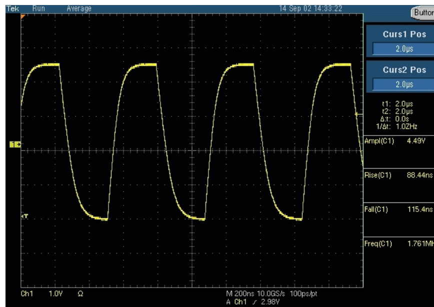


Figure 5.36: The output waveform of the divide-by-64 counter obtaining its input from the 31-stage internal ring oscillator.

This implies that the internal oscillator is running at $1.76 \times 64 = 112$ MHz.

For a comparison between the two structures, we look at what would happen if both had only 3 stages each. From Eqn. 5.4, ignoring the errors introduced by the buffer stage for the internal oscillator and the probe load for the external oscillator, we can find the equivalent frequencies. We can also calculate the single-stage delays for either structure. The results are given in Table 5.1, which shows that the speed improvement when the signal stays on-chip is 228ns/287ps=794-fold.

	Internal Oscillator	External Oscillator
Operating Freq.	112 MHz (31-stage)	398 KHz (11-stage)
3-stage equivalence	1.16 GHz	1.46 MHz
One-stage Delay	287 ps	228 ns

Table 5.1: The performance comparison of internal and external ring oscillators.

This effect is the direct result of the extra capacitive load on the oscillator inverters caused by the bonding pads (which the layout program Cadence extracts to be around 0.24 pF each, not including the ESD protection diode capacitances), bonding wires and the breadboard. To estimate this capacitance, we used the large-size inverter on the 1.6 μmABN chip, inserting varying additional load capacitances between its output and the ground and measuring the rise, fall and delay times. Extrapolating from the data points thus obtained, we can deduce an extra load capacitance of about 15 pF on the inverters of the external ring oscillator. This extra load exacerbates the trade-off requirement between speed and power in an unwelcome manner, due to the expanding current requirement to charge the extra capacitance to the same voltage level within the same rise time.

5.7 Summary

In this chapter, we performed an investigation of reactive elements on ICs. We first studied the properties and design of on-chip inductors and transformers, which are necessary components for the creation of system-on-a-chip type applications, such as "smart dust" wireless sensor networks and communication circuits. These are also good 3DIC implementation candidates, as we pointed out before.

We presented basic inductance concepts and described the physical features of typical on-chip inductors. We pointed out the electrical implications of the geometry and the presence of the lossy substrate on which these structures are built. We then provided an intuitive approach to the effects of different physical parameters, showing the typical inductance curve and its dependencies, concluding with some design guidelines. We continued by presenting measurement results from inductors and transformers we laid out and had fabricated, along with some remarks on their comparative performance.

Tuned loads and tunable LC structures are useful in many circuits for amplification, mixing, oscillation, matching and similar applications. We indicated the effect light shining on an on-chip inductor on its inductance, and explained this qualitatively as the rising substrate conductivity causing a lower resonant frequency. We demonstrated the same effect using a published on-chip inductor model.

Finally we focused on the effects of bonding pad parasitic capacitances, demonstrating that the higher load on signals having to go off-chip takes a heavy toll in operating frequency.

Chapter 6

Looking to the Future: Improving Electrical Engineering Education

6.1 Introduction, Motivation and Background

Integrated circuit technology is not the only area of electrical and electronics engineering to have undergone an exponential rise in complexity over the last decades. In some aspects, continuing miniaturization has been the driving force behind innovation, including the development of 3DICs—however the driving force behind *that* is an ever-increasing demand, both in amount and in scope, for electronic applications in society. Meeting this demand and creating further new demand by innovation, in turn, both depend on the ability of the engineering community to educate new engineers with a strong grasp of the state-of-the-art and an understanding of the directions this branch of human knowledge is poised to develop in. Accomplishing this end requires both provoking a wider interest in new generations for studying electrical engineering and presenting an ever-improving education to those who do choose to study this subject.

Another point worth note is that this mounting prevalence of electronics technology in society creates a need for the society to better understand the technologies in use, especially since “prevalence of ” can in many cases become “dependency on”. Thus a better introductory training for people who might not necessarily go on to conclude the full engineer’s training is beneficial.

In this chapter, we present our work on contributing to engineering education with an eye for furthering the goals outlined above. Our first contribution is the design and implementation of an introductory course to electrical and computer engineering intended for high school students. We next present a day-course designed for high school mathematics teachers; this course aims to assist them in emphasizing to their students the relevancy of their subject in electrical engineering research and development. Finally, we detail the design and implementation of a 400-level Capstone design course for electrical engineering students; this course is significant in its approach to attempt and bring together electrical engineering subdisciplines for a comprehensive system design project.

6.2 An Experiential Introduction to Electrical and Computer Engineering: A Summer Course for High-School Students

6.2.1 Introduction and Program Goals

With the current rate of advance in applied science, and the previously unmatched level of our dependence on its products, it is helpful that every student achieves a basic level of understanding and appreciation of technology. In the course of training students in basic skills, it has thus become important to teach the fundamentals of electronics and computer structures at the onset of the 21st century, just as it has been with chemistry, physics, biology and mathematics in the latter half of the 20th century. In addition, as with any field, Electrical and Computer Engineering is always in need of qualified students to become the next generation

of engineers. Early exposure to the concepts of engineering is an effective method of drawing in prospective students [146].

To help satisfy these emerging pedagogical needs, we have developed a new teaching methodology for educating high-school and beginning-college students in the areas of Electrical and Computer Engineering (ECE). The unique aspect of this approach is that it emphasizes the teaching of advanced engineering topics using concepts and relevant experience rather than high-level mathematics. By emphasizing hands-on laboratory work and qualitative description, many advanced topics, which are normally reserved for the sophomore and junior year ECE college curricula, can be presented to high-school and beginning-college students. While the treatment of these subjects for comprehensive professional training naturally requires the accompanying mathematical proficiency, we found that the descriptive approach, without detailed mathematics, is effective for giving students a conceptual anchor early-on in their technical training. This anchor should also provide an initial condition from which deeper mathematical treatments can be pursued and assimilated. Thus should help improve student retention in technology-related fields, which is a current concern along with new student recruitment [147, 148].

We applied this methodology under the auspices of the Maryland Governor's Institute of Technology, a living-learning program for selected high-school seniors sponsored by the State of Maryland. In the Summer of 2001 we designed and implemented the first ECE Technical Program for this Institute at the University of Maryland, College Park [149]. This living-learning program lasted five weeks, with daily lectures and laboratory sessions. We found that the students responded

favorably to the methodology. They successfully used the knowledge obtained in the lectures and laboratories to complete approximately 25 experiments, design a simple central processing unit (CPU), and fabricate high-fidelity audio amplifiers that they took home at the end of the program.

In the design process for this course, we identified design goals for this program from a pedagogical point of view. We summarize these goals here.

- Providing an enjoyable introduction to Electrical and Computer Engineering fundamentals to students, who might lack any prior background, at a high-school advanced placement course-equivalent conceptual level.

- With an eye out to improve retention in higher technical education, helping students build a conceptual foundation to a later potential analytical study of the subject matter. Also, giving potential ECE students the advantage of an early exposure to the subject matter as well as the material covered in particular in this course.

- Providing the students with a chance to build their hands-on laboratory experience, which is valuable no matter which particular technical field they might choose to branch into.

- Demonstrating the particular relevance and contributions of Electrical and Computer Engineering to modern society. Further, contributing to the students' future career choice by a solid, rather than abstract, presentation of what this branch of engineering entails.

- For the particular material covered in this course, developing a presentation method accessible to students with different learning styles, which is in increasingly sharper focus in modern pedagogy [150].

- Evaluating the effectiveness of our particular approach to elementary electronics and computer pedagogy, which is a conceptual and experiential, rather than a mathematical, presentation.

6.2.2 The Syllabus and the Textbook

The course included a laboratory component and a lecture component. Both components were supplemented by a textbook written by the course designers [151], which is covered in more detail ahead. The topics to cover were selected to introduce basic electronic components, their manner of use and applications, and some of the most fundamental concepts of electrical engineering, such as analog signals and digital electronics. Thus the course is an introductory summary to several sophomore- and junior-year courses [152]. Here is an overview of the course syllabus.

1. An overview of Electrical and Computer Engineering: branches and subjects.

2. Basic concepts in electrical physics and circuit theory: charges, electric fields, potential, capacitance, resistance, direct currents (DC), alternating currents (AC), Ohm's Law, Kirchhoff's Laws, and semiconductors.

3. Basic concepts in the mathematical representation of signals: sinusoidal signals, amplitude, frequency, trigonometric representation, frequency compo-

nents in signals and superposition.

4. PN-junction diodes: rectification, AC/DC conversion, light-emitting diodes (LEDs), and photodiodes.

5. Operational amplifiers: comparators, amplification, inverting-, non-inverting, and summing amplifiers.

6. Filters and filtering: active and passive high- and low-pass filters.

7. Bipolar junction transistors: current and voltage amplification.

8. A basic hi-fi audio amplifier: signal and power amplification, the concept of loading, power transistor usage, and heat sinks.

9. Digital logic and digital circuits: Boolean algebra, truth tables, AND and OR gates, combinational logic examples, sequential logic, and a simple adder.

10. Computer technology: standard computer architecture, programming languages, and a simple CPU design.

11. An optical arcade laser game design and implementation: phototransistors, revisiting comparators, and computer interfaces.

6.2.2.1 The Textbook

In order to present the material in a manner suitable to our experiential and conceptual pedagogical approach, we preferred to prepare and use a dedicated text for the course. This choice allowed us to avoid overwhelming the students with

the full-formal approach they would encounter if we had used a compilation of pre-existing course and laboratory texts. Too, we could tailor the text contents to our selected subject material and present the supplementary theoretical material for the lectures integrated with laboratory session guides.

In keeping with the goal of demonstrating the relevancy and contributions of electrical and computer engineering to our students in a concrete manner, for all three components of the course (lectures, textbook and lab work) we worked to present direct links to the real world for every syllabus item. To demonstrate some examples from the lectures and the textbook:

- Electric fields were introduced in the context of cathode-ray tubes.
- Sinusoidal signals, the concepts of frequency and amplitude and the frequency components present in a signal were presented not in an abstract, mathematical manner, but rather with direct physical demonstrations using musical instruments, voice, a microphone, a signal generator and an oscilloscope.
- While covering pn-junctions and rectification, the prevalent application of power supplies in household appliances were invoked.
- The "capstone" project of the analog portion of the course was an audio amplifier that the students built and took home at the end of the session. This project allowed them to immediately use the information presented with the concepts of operational amplifier (amplification), filtering (basic equalizer), and bipolar junction transistors (output driving stage).

- The concept of digital logic design was illustrated with a storybook approach to a vending machine design. Students later built an actual adder as a primitive calculator building block, and were introduced to the fundamental computer architecture in use today, complete with the CPU, memory, and bus structures.

Each subject was linked to least one experimental section. The Appendices of the textbook were intended for use as a laboratory handbook. In the experiment design, care was taken to cater to different learning styles by making sure that the experiments usually provided some immediate sensory feedback to the student, beyond the data they were directed to take. For instance, the relationship between voltage, resistance and current was illustrated in experiments where students could gauge the current level in a branch by the brightness of an LED, in addition to measuring and calculating the current.

6.2.3 Implementation

The selection process admitted twelve students from numerous Maryland high schools, of an achievement level approximately equivalent to that of an arriving freshman class to ECE at the University. Apart from a few brief exceptions for field trips and multi-department presentations, there were five-days-a-week lecture and lab sessions for six weeks of the summer.

For the lecture component, the lecture structure was the standard presentation by the professor, who accepted student questions. However, the content focus was on technological applications and descriptions of device and circuit operation and

applications, to be reinforced in the lab, rather than on rigorous mathematical derivation. The mathematical content of the course was limited to algebra and trigonometry, with no calculus included.

A dedicated laboratory was established in the summer of 2001 with funding set aside for this program. This laboratory contained standard EE lab setups including an oscilloscope, a DC power source, a signal generator, a multimeter and circuit prototyping boards. Except for the audio amplifier project, which was individual, students worked in pairs. At least one TA was always available during the lab sessions, which typically lasted for two to three hours. The experiments varied in length, with sometimes multiple experiments completed during the same session and sometimes a single experiment taking multiple sessions.

One tool the course designers used to help the students gain proficiency in lab equipment and procedures was providing them with lab report templates tailored for each experiment. Two example templates are included in Appendix C.

6.2.4 Program Outcomes

We used several different methodologies to assess program outcomes. These included an exam, student comments, instructors' observations, focus groups conducted by collaboration with an educational psychologist [153], and an end-course survey. Here we describe the course outcomes as they are related to the stated course goals.

- *Prevalence of ECE in society.* The students reported an increased awareness

of the application of ECE in daily life, as well as a deeper comprehension and appreciation of some common applications. This particular outcome was described by some students as a highlight of the course. In the survey, students stated their agreement with the claims that they "learned applications of electronics" and "increased their overall knowledge of technology."

- *Hands-on Laboratory Experience.* The laboratory component was nearly unanimously acclaimed as the favorite part of the program. The instructors observed that the students would, unprompted, routinely stay in the lab to work after hours on their own hi-fi audio amplifiers. The tangible end-result led the students to agree that "it was exciting to observe my audio system work" and "wiring and actually soldering [the system] was worth the time and energy." Another observation by the educators is that student response was most enthusiastic for experiments with results that the students could relate to directly and solidly, like the audio amplifier, the primitive CPU design and the optical arcade game.

  From the educators' point of view, early exposure to the standard pitfalls of measurement and prototyping equipment is another advantage.

- *Conceptual Presentation and the Course Level.* Students reported that the lectures were more fast-paced and challenging than standard high-school lectures. Most found the elementary mathematical content within their grasp and stated approval on the conceptual, rather than mathematical, focus.

- *Catering to Different Kolb Learning Styles.* At the course outset, the Kolb

Learning Style Inventory [150] was administered. The students were found to be divided almost evenly between Convergers (preferring to learn abstract concepts by concrete examples and methods, such as labwork) and Assimilators (preferring to learn abstract concepts by observing others' conduct experiments, and by interpolation and extension of possessed knowledge). By the end of the course both types of students appeared to benefit equivalently from the program, possibly thanks to the presence of both lab and lecture components [154]. As stated before, both types of learners seemed to grasp concepts that were presented with a practical component involving an immediate sensory feedback facet. Since our student group did not include any Divergers and Accommodators, which are the remaining two types of Kolb learning styles, we could not assess the effectiveness of our pedagogy on these two styles.

- *Impact on Students' Educational Choices.* At the end of the course, the students agreed or strongly agreed that they would "want to study and learn more about the subject matter." They also stated that the course was helpful in giving them a clearer idea about their future career choices and make more informed decisions about their post-secondary education prospectives.

- *Student Achievements.* Midway through the course, an 1.5 hour test covering the analog electronics portion of the syllabus was administered to monitor the students' gains. The results were remarkable: With example questions such as those presented in Figure 6.1, about half of the students scored above 90%. Considering that the course relied on the students' self-motivation, this

outcome is very favorable. The high level of success students achieved in completing some 25 experiments and projects within the laboratory component, including personal audio amplifiers built from scratch and compatible with household equipment, is a further indication that the students benefited broadly from the program.

- *General Student Commentary and Survey Responses.* In general, our students reported a strong agreement with the statements "the program was a worthwhile experience" and "[they would] recommend the program to others." They indicated that the subject matter was challenging appropriately for their level and related very favorable impressions.

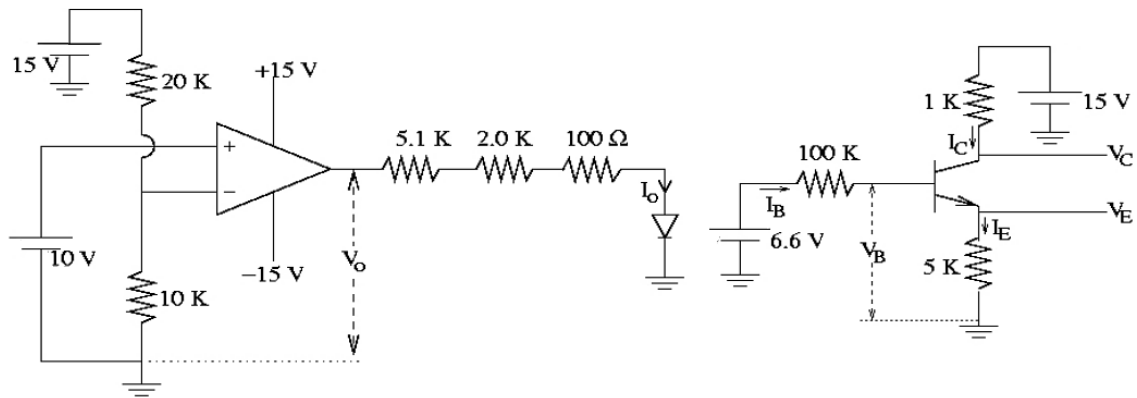


Figure 6.1: Example exam questions. These were the associated question statements: *(Left circuit)* For this op-amp circuit, what are I_o and V_o? Is the LED on, why or why not? *(Right circuit)* For the BJT circuit β=100 and V_{BE}=0.7. Draw the DC equivalent circuit. Calculate I_B, I_E, I_C, V_B, V_E and V_C.

6.2.5 Conclusion

Through a curriculum developed and implemented to present high school and starting college students with the electronics and computer architecture concepts with a focus on concepts and applications rather than their mathematical background, we were able to introduce our students to material ordinarily covered at sophomore and junior level courses, and observe a respectable amount of student benefit. Using a combined lecture/hands-on approach we could address the different learning styles encountered in our students. The students' gain was reflected in the favorable results of the one in-course formal test, as well as their enthusiastic response to and success in the lab modules. The students reported a rising appreciation of electrical and computer engineering in daily life, as well as a higher ability to make informed decisions about post-secondary education study choices. We expect our approach to help with retention problems in engineering education thanks to the course preparing the students with a stronger foundation on which to build the more rigorous, abstract approach of college-level technical courses.

6.3 Aiding the Mentors: A Short Course for High-School Maths Teachers

6.3.1 Introduction and Motivation

During the summer of 2004, a project supported by a grant provided by the General Electric Foundation Math Excellence Initiative and titled "TIME: Teachers

Integrating Mathematics and Engineering" was implemented at the University of Maryland. Part of the program was a two-week engineering training institute for middle and high school mathematics teachers [156, 157]. The program goal was to contribute to the solution of some prevailing problems in education in the United States: Insufficient interest in studying engineering among the newcomer students and a decline in the technological literacy at the level required by a technologically-oriented society [155]. It is expectable that a strong real-world relevancy component in mathematics education at the middle and high school would be beneficial, both from a technological literacy viewpoint and mathematical literacy viewpoint [158].

Our contribution to this program was in the design and implementation of a workshop for mathematics teachers [159]. We presented the program participants with an introduction to electrical concepts using algebra, gave a lecture/presentation introducing electrical and computer engineering in terms of its scope and how it uses mathematics as its language. We further provided the teachers with a workbook and a companion CD, intended for use in their own classrooms to expose their students to real-world engineering applications of mathematical concepts by means of worked-out examples and exercise questions.

6.3.2 Implementation

Our workshop consisted of three main components:

- A hands-on section for the presentation of the basic electrical concepts of voltage, current, resistance, power, and their relationship. After a basic the-

oretical instruction, the participants were provided with breadboards, power supplies, resistors and LEDs for experimentation. TAs were available to the participants to assist with their experiments.

- A presentation of how ECE uses mathematics as a language for its physics-based aspects as well as for representing and manipulating information. To illustrate this point, we picked three electronics applications and demonstrated their underlying mathematical background:

    1. A CD player: The "analog signal" concept, trigonometric functions combining to form analog signal representations of sound , digital representation of analog levels, digital-to-analog conversion, digital data storage.
    2. An AM radio receiver: Combinations of analog signals, filtering, modulation, demodulation, rectification.
    3. A Digital Camera: Image representation by pixels, grayscale levels, obtaining full-color pictures from red/green/blue channels.

- A worksheet suitable for the participants' use in their own classrooms, with example questions under the following headings

    1. Residential Power Calculations: Calculating cost from usage and reading a residential power bill. Example: "A family leaves their refrigerator plugged in and running while they go to vacation for three weeks. The refrigerator has a power rating of 1.8 kW. Assuming the refrigerator works for only 12 hours of each day, how much money could the family have

saved in electricity costs had they emptied all the perishables and plugged off the refrigerator? [Assume the cost of electricity is 5 cents per kWh.]"

2. Digital-to-Analog Conversion worksheet: The questions provide sequences of three-bit digital numbers, asking the participant to first plot this sequence as digital levels marching in time, then to do an "digital-to-analog conversion" by hand and comment on the differences of resulting signals. An example is shown in Figure 6.2.

3. Digital Camera Operation: Color image formation and grayscale pixel-array-to-image exercises. For the color image formation section, the CD provided to the participants included numerous images presented in full color and also split into their red, green and blue components, the suggested activities being along the lines of either trying to guess which regions would be brightest for each of the three channels, or looking at the channels and guessing what color would be prominent in which regions of the full-color picture. For the grayscale pixel-array-to-image exercises, two empty grids of pixels (16x16 and 32x32) were given, along with arrays (16x16 and 32x32) of numbers representing intensity values. The participants were asked to shade in the "pixel array" elements with shades of gray corresponding to the intensity values given in the number arrays to obtain were grayscale images, one of which is presented in Figure 6.3.

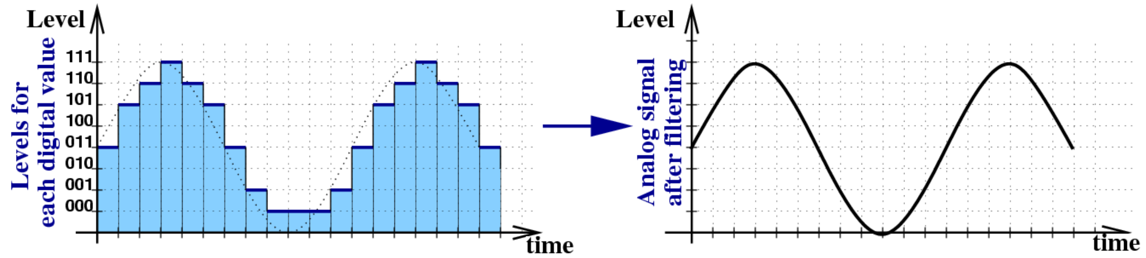


Figure 6.2: Example digital-to-analog conversion question. The sequence: **011 101 110 111 110 101 011 001 000 000 000 001 011 101 110 111 110 101 011**.

6.3.3 Conclusion

The main idea behind this program was to address the dual issues of engineering student recruitment and increasing the level of technological literacy in society by using the channel provided by middle and high school mathematics classes. The approach was designed to benefit the pedagogy of these classes themselves by affording them a way to showcase the "real-world relevancy" of the concepts which are their onus to present to their students. This particular implementation of our program was received very positively by the participants, who stated that they enjoyed and benefited from the presentation and the hands-on sessions. It is our hope that the material presented for classroom use has also been beneficial.

Further, this demonstrates that it is possible to use many fruits of recent research, not only in mathematics, but also in chemistry, physics, environmental science, biology and other related courses, to reinforce the subject immediacy and real-life relevancy. Some suitable areas such presentations could be solar power harvesting, novel battery designs, sensor networks and related miniaturization, computation and communication network concepts, biomedical technology and MEMS.

Figure 6.3: Example grayscale image converted into a 8-level 32x32 grayscale array. Top left: Original image. Top right: Empty grid and grayscale levels provided. Bottom left: Grayscale values given. Bottom right: Reconstructed image.

6.4 Multi-Track Capstone Design: A Senior-Year Course

6.4.1 Introduction, Motivation and Course Goals

As newly-graduated engineers enter the workforce, they find that prospective employers require knowledge and skills beyond the purely theoretical [160]. Graduates are expected to understand system-level design to a degree, and possess the skills of teamwork and creative thinking [161, 162]. Following the needs engineering education should meet, the Accreditation Board for Engineering and Technology (ABET) requires engineering programs to graduate students who can "design a system, component, or process to meet desired needs," "function on multi-disciplinary teams," and "identify, formulate and solve engineering problems." One of the more concrete criteria is that engineering curricula should culminate in "a major design experience based on the knowledge and skills acquired in earlier coursework and incorporating engineering standards and realistic constraints that include most of the following considerations: economic; environmental; sustainability; manufacturability; ethical; health and safety; social; and political" [42].

The Electrical and Computer Engineering department of the University of Maryland at College Park offers a number of senior-level single-semester Capstone Design courses following this requirement. In an attempt to address this criterion further, we developed and in the fall semester of 2003 implemented an Advanced Capstone Design class: "ENEE 498C, Topics in Electrical Engineering: Capstone II: Advanced Design" [163].

One important feature of this course is that it is designed to afford new gradu-

ates a "big picture" view of electrical and computer engineering. The ever-expanding depth and breadth of the subject matter necessarily limits our students to choosing a few of the existing subspecialties to focus on during their senior year. Providing them with a chance to retain an awareness of areas they drop out of formally studying after their junior year is thus beneficial. Therefore we designed the course to promote interaction between students with differing choices of subspecialties.

We summarize approach to the course design and implementation as follows:

- In order to emulate a multi-subspecialty system-level design experience, the original course concept envisions recruiting students who have already taken different Capstone design classes and creating course projects requiring combined experiences from numerous fields.

- To create a vertical as well as horizontal flow of information and experience within the student group, we planned to enroll students with different levels of training and internship/industry experience. Our pilot run had a few graduate students enrolled along width fifth and fourth year students.

- The course was shaped organically during the semester, and student input was actively requested and solicited at every stage. We give more details of this point below.

- In terms of the course structure, the lecture component was smaller than other Capstone classes, and included seminars by professors teaching Capstone classes in the department, aiming to introduce their subject areas to students who had not taken their particular course.

6.4.2 Implementation

During the pilot run of the course, we included seminars by professors with varying backgrounds and by the students themselves, a warm-up project for individuals or small groups, then the main course project for student teams. Student feedback and comments were taken throughout the semester and at the end.

During the first week, students were required to give short individual presentations on design projects they had carried out in their previous education and detail the technical skills they possessed and could contribute to a project group. Besides introducing each student's technical skills to their peers, these presentations also made all students aware of more design problems than they would have encountered by themselves previously.

The warm-up project was designing an integrated circuit using Cadence and submitting it for fabrication through the MOSIS fabrication facility. The know-how training on chip and printed circuit board came from two of the students themselves. The students picked which circuit to design, simulate and layout; the projects range in scope between gates to counters.

About three-fourths of the semester was taken by the main course project. The instructor provided several project outlines for the students to discuss and refine through in-class discussions, brainstorming sessions and group meetings. Through intense student involvement, the projects definitions were refined to focus on two final project ideas: a direction-finder and a rangefinder. Both projects integrated the skills of RF design, digital design, microcontroller-based design and PCB layout

experience. The instructor then gave, by way of specifications, *functional definitions* of these two projects, and suggestions on potential technologies and components to be used in the design. The students worked on tightening the specifications and moved on to the design and system integration. The students had access to a dedicated laboratory for their design work. At the end of the semester, they were required to give individual as well as group presentations on their projects.

6.4.3 Outcomes

We can summarize the couse outcomes from the students' point of view as follows.

- *Exposure to the Design Process.* As described above, the students shaped the system specifications and requirements as well as detailed descriptions by themselves during the project. Throughout the semester the instructors required the students to provide task lists, single-block design documents and system-scale dependency definitions. Thus the students, exposed to a more holistic design process, designed the projects as well as the products.

- *Teamwork and Project Management Skills.* The team projects required careful work division, intra- and inter-team communication, and project management/timeline skills. Student comments exhibit their awareness of the need to acquire these qualities and utilize what they already possessed.

- *Student Autonomy and Information/Skills Sharing* The students were required to come up with their own incremental deadlines as well as project scope and

details. They responded very enthusiastically and successfully to this open-ended course structure, creating a website for discussions, information-sharing and general ideas exchange, and turning in successful projects. The instructors observed the students were very open to learning from each other as well as aware of the role of task-division in their project teams.

- *Exposure to the Concept of System Engineering.* Students indicated a growing awareness of the higher-level system engineering activities the projects required. For systems where different circuits and subsystems are combined, an overall, organized understanding of compatibilities, dependencies and interactions is necessary. Student reports point to someone in the group taking on such a role and the others working with this person. One of the students had a striking suggestion about whether it would be feasible to offer an introductory-level system engineering course at the undergraduate level.

In the end, although the previous academic performance level varied widely in our group of students, their participation in and contribution to their respective group designs were notable. They indicated their appreciation for being exposed to a more realistic design project, took up novel responsibilities arising in such a project situation eagerly, and performed well.

6.5 Summary

In this chapter we presented our contributions to engineering education, at the pre-undergraduate level, senior undergraduate level and at one step removed by targeting pre-undergraduate educators.

Our first project was the design and implementation of an introductory course to electrical and computer engineering for high-school students. This course covered the basics of analog and digital electronics over a six-week period incorporating lecture and lab components. As the program achievements, the students gained an enhanced awareness of ECE applications and concepts and hands-on laboratory experience advanced enough for them to manufacture their own audio amplifiers, grasped the course material well enough to perform remarkably in a midterm test, and responded very favorably when questioned about their experience. The paper we presented about this program in the 2002 Frontiers in Education Conference was awarded the Benjamin Dasher Best Paper Award for this year.

Our next program was a short day-course developed for high school mathematics teachers, with the dual aims of assisting the teachers to enhance the real-world relevancy of their syllabus concepts and increasing technological awareness in their students through these applications. After a presentation of basic electrical concepts, we provided examples of how electrical engineering uses mathematics as a language and how mathematical concepts are used in the design process and operation of everyday-life ECE products. Once again, we received a positive response from the participants while demonstrating that it is thus possible to incorporate

recent developments in electrical engineering in pre-college courses, making both the course material and these developments of higher immediacy in the students' perception.

Finally, we presented the design and implementation of a senior year multi-track capstone design course. This course was unique in several aspects: For example, it was conceived for students who have already taken single-subject capstone design classes from different areas, so that they could bring their different experiences together in student groups. Furthermore, the course was shaped organically as it was being taught, student input being sought after all throughout, up to and including the scope and specifications of the final project. Thus the course succeeded in its aims of allowing students to learn about the design process as well as the design project, and honing their teamwork and project management skills. Our observations were that the students respond openly and eagerly to an open-ended course structure which required them to share their knowledge and skills as well as carry on their own planning. The benefits of such an experience are clear at the senior level in an area where graduating engineers with an idea of the "big picture" have an unquestionable edge.

Chapter 7

Summary and Future Research

7.1 Three-Dimensional Integration

In Chapter 2, we described the physical design, analysis and testing of a self-contained three-dimensional integrated circuit (3DIC). To our knowledge, this is the first demonstration of a self-powering SOI 3DIC.

For harvesting the necessary energy to power the functional element of the system, we designed photodiode arrays, optimized within the process and geometry constraints. Through analysis and simulations, we demonstrated that the expected photocurrent amount is sufficient to operate the functional element. We presented quantitative and qualitative analyses of the self-regulating behavior of our system. We then presented the measurement results from the circuit operation.

We have also designed and submitted for fabrication an expanded second-generation version of this system for the next run of this 3-D process. Section 2.3 describes the new features and design decisions for this version.

The chapter concluded with reviews of the states-of-the-art for 3-D integration and self-powering methods. 3-D integration is a promising technology at a stage of its development where its limitations arising from physical constraints and problems are well-defined. Further sections of this present work contribute to the development of techniques to address some of the physics-based issues, especially thermal and

noise-coupling related problems, in this new approach to VLSI.

7.2 Interconnect Network Modeling

In Chapter 3, we presented the development and results a method to analyze internal and external noise and signal coupling on interconnect networks on integrated circuits. The method is based on calculating the impulse response of the network, which is modeled by an RC network comprised of different unit cells which combine to make up the layout. If specific on-chip locations are identified as points of interest, the method can be used then to find out how vulnerable they are to noise or unintentional signal coupling from different parts of the chip. The storage and, once the impulse responses are calculated, computational requirements of our method are clearly advantageous over SPICE and similar full-circuit solvers.

The chapter introduced the method, the physical modeling approach, and demonstrate the implementation. The simulator is shown to match SPICE results over small networks and is then utilized to investigate the responses of an interconnect network with physical parameters based on a real fabrication process available through MOSIS.

Two extension projects following up on this work may be suggested. For the first project, given the layout of an integrated circuit, the researcher identifies critical points of interest, for instance the gate connection points of certain transistors, or the locations in the substrate corresponding to their body, where current injection might affect transistor operation by threshold-voltage shifting and similar problems.

Next, the impulse response to impulses originating at many points spread over the IC are calculated and stored. A random signal generator can then be used to emulate random noise coupling at these potential originating locations, which will yield a vulnerability map by rapidly calculating the outputs to these random inputs at all the target locations.

For the second project, a three-dimensional integrated circuit with at least two tiers should be modeled as a network, following the methodology given in the present work, but also including the substrate layers. The questions then to be addressed are related to the noise and crosstalk issues of 3DI presented in Section 2.4.3. After calculating the impulse responses in the substrate of one tier to impulses injected into the substrate of the other tier, the full response to typical substrate-noise waveforms originating in one substrate can be investigated in the other substrate, allowing the researcher to evaluate the advantages of 3DI in the integration of mixed signal systems with analog and digital subcircuits in separate substrates. The next thing to check for is unintentional coupling between the interconnect layers of one tier and the substrate of the next, as is suggested in the literature to be a potential problem. A designer could then implement a form of shielding, such as ground planes, between tiers and include this in the interconnect network model, repeating the simulation to determine the effectiveness of such an approach.

7.3 Thermal Modeling

Chapter 4 investigated heat generation and distribution on integrated circuits, with a focus on how the different materials present in the IC system affect and shape heat conduction and thermal distribution under different geometric conditions. We started by a simplified solver for the heat flux equation to get an intuitive sense of what effects non-uniform, non-isotropic thermal conductivity has on temperature distributions for a given geometry and heat source. Later, we expanded an advanced solver to implement the thermal conductivity of different material layers and applied this expansion to two integrated circuits of our own design.

To validate and calibrate this solver we designed and had fabricated integrated circuits which included individually addressable heaters and pn-junction diodes as thermal sensors. We performed measurements and used our data, along with the known physical properties of the IC process, to calibrate the package resistivity of and validate the solutions from our simulator. Thereby we determined that for a bulk process such as the one used here, the lateral metal interconnect network formed of the ordinary interconnect metal layers does not have an extensive impact on temperature distribution, as the substrate itself is the dominant thermal conductor.

Later we used our simulation capability to investigate several approaches to cooling and controlled heating in integrated circuits. We demonstrated that using a "cage" or "fence" formed of vertical vias around a heater element, we can both reduce the peak temperature and contain the thermal diffusion. We repeated this demonstration assuming an SOI process. We also demonstrated that for the SOI

process, ordinary planar interconnect network does affect temperature distribution.

To suggest the extension of this work in the context of three-dimensional integration technology, we refer back to the thermal problems and proposed solutions described in Section 2.4.3. Our simulator can be expanded further to simulate multiple stacked tiers, with each material layer of the individual tiers still implemented. A designer can then include vertical vias, connected to the ambient directly or through a known thermal resistance, and evaluate the effect of different placement and sizes of such vias in bringing down the peak temperature. If the 3DI process is based on an SOI process, the effect of ordinary interconnect networks will also be discernible and should be investigated.

This work is also useful in that it demonstrates controlled, localized or directional heating of selected regions in integrated circuits is possible. This is helpful in several applications, including electronics to operate in extreme temperatures. It is conceivable to merge the capabilities of this simulator and its predecessor, which also self-consistently solves for device and chip heating by simultaneously considering transistor behavior at different temperatures and transistor power dissipation as a heat source. One step further is a CAD program which gives the designer the ability to create a feedback system for controlled heating: The temperature distribution arising from an operating heater unit can be obtained by the thermal solver. The temperature at the location of a preset thermal sensor can then be input into a circuit simulator that solves for the circuit, including this sensor. The simulated output of this circuit can be set to control the heater operation, which takes the simulation back to the thermal side, until self-consistent operation is achieved.

7.4 Reactive Structures and Load Effects in ICs

Chapter 5 provided an introduction to on-chip inductor and transformer technologies and an investigation of parasitic capacitive load effects in integrated circuits. The former structures are necessary for compact communication circuits used in wireless network systems, which are suitable as 3-D integration applications thanks to the reasons outlined earlier in the dissertation. Furthermore, new geometric freedoms afforded by 3DIC invite the development of novel inductor and transformer structures on that platform, which should be studied based on the understanding of these devices started here.

We introduced the basic concepts and physical features of on-chip inductors, describing the electrical effects of the geometry with the help of a simplified model. We presented some design guidelines within the geometrical degrees of freedom afforded by the standard CMOS process. We showed the measurement results from our manufactured inductors and transformers and gave some remarks on their comparative performances.

We pointed out that shining light on an inductor structure laid out on a silicon substrate causes a shift in the self-resonant frequency. We explained the effect arising from the substrate conductivity change using a lumped-element type on-chip inductor model. As an extension to this work, we can propose the design of a semiconductor device underlying the inductor, conceived in order to change the substrate conductivity controllably through the application of an optical or electrical signal. Such a tunable LC structure would be useful in many different

circuit applications related to communications and RF electronics.

We concluded the chapter with the investigation of the effects of signals having to go off-chip vs. signals conducted through the equivalent of inter-tier connections in a stacked technology. This investigation verified the power and operating frequency related advantages of 3DIC as were sketched in the Introduction and later featured in Section 2.4.3.

7.5 Electrical and Computer Engineering Education

In Chapter 6 our work in improving engineering education at the undergraduate and pre-undergraduate level was presented.

We showcased three programs: A six-week summer course targeting high school students, a day-course for high school mathematics teachers, and a 400-level advanced Capstone design course for ECE majors. Through the summer course, we succeeded in giving our students an experiential introduction to the basic concepts of ECE, plus rather well-received laboratory experience and an increased awareness of ECE in daily life. The day course fulfilled its dual aims of raising the same awareness in high school teachers and through them, their students, and helping the teachers enhance their curricula by providing them with "real-world relevant" example applications of the concepts they have to teach. The advanced Capstone design course took the concept of a capstone design course one step further by allowing the students to work in design teams including people with different specialties, thus giving them a taste of system engineering. More, it made the students a part

of the design process by letting them shape the project specifications and timeline.

One of the fruits of the summer course was the textbook we wrote for it, *An Experiential Introduction to Electrical and Computer Engineering* by Goldsman and Dilli. This book can be revised and offered for wider publication. It is possible to model day-long outreach programs from colleges to local high schools based on the day course we presented here. Finally, the advanced Capstone design course ENEE498D as described here, with the student autonomy it affords, is suitable for entirely different procession and outcomes every time it is implemented. It is also possible to conceive of it as an introductory master's level course, or design an expansion in the form of a system engineering course.

Appendix A

Temperature Dependence of Diffusion Constants and Diffusion Lengths

Here we reassess the assumption in Section 4.3.2 that the diffusion constants $D_{p/n}$ and diffusion lengths $L_{p/n}$ are temperature-independent.

In the diode current expression, Eqn. 4.11, the term depending on the diffusion constant and length were considered temperature-independent. From a physical point of view, this assumption is not entirely correct. Since the diffusion lengths $L_{p/n} = \sqrt{D_{p/n}\tau_{p/n}}$, where $\tau_{p/n}$ are the recombination lifetimes, this relevant term can be rewritten as follows:

$$\frac{D_p}{N_A L_p} - \frac{D_n}{N_D L_n} = \frac{1}{N_A}\sqrt{\frac{D_p}{\tau_p}} - \frac{1}{N_D}\sqrt{\frac{D_n}{\tau_n}} = \sqrt{\frac{kT}{q}}\left(\frac{1}{N_A}\sqrt{\frac{\mu_p}{\tau_p}} - \frac{1}{N_D}\sqrt{\frac{\mu_n}{\tau_n}}\right) \quad . \tag{A.1}$$

The recombination lifetimes can be obtained from trap densities N_{trap}, the thermal velocity v_{ther} and trap cross-sections σ_o:

$$\tau_p = \frac{1}{N_{trap} v_{ther} \sigma_o} \quad , \tag{A.2}$$

similarly for τ_n. v_{ther} is given by [43]:

$$v_{ther} = \sqrt{\frac{3kT}{m_p^*}} \quad , \tag{A.3}$$

thus

$$\tau_p = \frac{\sqrt{m_p^*}}{N_t \sigma_o \sqrt{3kT}} \quad . \tag{A.4}$$

Reconsidering Eqn. A.1, the mobility terms' dependency on temperature has been extensively investigated in literature [117]. For $N_D = N_A = 10^{19}$ cm^{-3}, between 280 and 350 °K, both electron and hole mobilities can be approximated by

$$\mu_{n/p} = \gamma_{n/p} T^{-0.44} \tag{A.5}$$

where γ is a constant different for the n-type and p-type carriers. Plugging this, Eqn. A.3 and A.4 into Eqn. A.1 gives

$$\frac{D_p}{N_A L_p} - \frac{D_n}{N_D L_n} = \beta_1 (T)^{3/4} T^{-0.22} = \beta_2 T^{0.53} \tag{A.6}$$

with β_1 and β_2 constants which depend on the effective masses, trap densities, cross-sections, k and q.

Taking this dependency into account we can rewrite Eqn. 4.16 as

$$I_D = \alpha\beta T^{3.53} \exp(\frac{qV_d/n - E_g}{kT}) \quad . \tag{A.7}$$

Using this modified form for the nonlinear function whose root will give us the diode temperature, as described in Section 4.3.2, we determined that disregarding the diffusion coefficient and lifetime temperature dependencies results in overestimating sensor temperatures by about 0.5 ° K.

Appendix B

Temperature Dependency of MOSFET Current in Saturation

Standard threshold-based models for MOSFET devices give the saturation current of an NMOS, as [12]:

$$I_D = \mu_n \frac{W}{L} \frac{\epsilon_{ox}}{t_{ox}} (V_{GS} - V_{thn})^2 (1 + \lambda(V_{DS} - V_{DS,sat})) \tag{B.1}$$

Here, the temperature-dependent terms are the mobility and threshold voltage, μ_n and V_{thn}. We ignore the temperature variation of the channel-length modulation for simplicity. For the mobility, we use a power law [118] with a proportionality constant α:

$$\mu_n = \alpha (\frac{T}{300})^{-2.5} \tag{B.2}$$

The threshold voltage is comprised of a number of contributing factors: The contact potential difference between the Fermi-level potentials of the gate material ϕ_G and the body ϕ_{sub}, the potential change required for inverting the surface potential, which is defined as twice the magnitude of ϕ_{sub}, and for attracting enough electrons to the surface to fill the acceptor states. The latter, assuming Q'_b is the fixed negative charge per area in the depletion region between the inverted channel and the substrate, is Q'_b/C'_{ox}, where C'_{ox} is also given per area. Here the source is assumed shorted to the body, so there is no potential difference between the substrate and the source to affect the depletion region width. If we consider an extra interface charge Q'_{if}, already present due to imperfections in the oxide, this potential should

be modified accordingly. Then the threshold voltage when the source is shorted to the body is given by [12]:

$$V_{thn} = -\phi_G + \phi_{sub} - 2\phi_{sub} + \frac{Q'_b - Q'_{if}}{C'_{ox}} \tag{B.3}$$

We rewrite this, explicitly stating the temperature dependencies of its components:

$$\begin{aligned} V_{thn}(T) &= -\frac{kT}{q} ln(\frac{N_{D,poly}}{n_i(T)}) - \frac{kT}{q} ln(\frac{N_A}{n_i(T)}) + 2\frac{kT}{q} ln(\frac{N_A}{n_i(T)}) + \frac{Q'_b(T)}{C'_{ox}} - \frac{Q'_{if}}{C'_{ox}} \\ &= -\frac{kT}{q} ln(\frac{N_{D,poly}}{n_i(T)}) + \frac{kT}{q} ln(\frac{N_A}{n_i(T)}) + \frac{Q'_b(T)}{C'_{ox}} - \frac{Q'_{if}}{C'_{ox}} \end{aligned} \tag{B.4}$$

The dependency of n_i, the intrinsic carrier concentration, to temperature was covered in Section 4.3.2. Note that here complete ionization is assumed, which is actually not the case for low-temperature operation [30]. The surface charge depends on the depletion region width, which itself depends on the body potential:

$$Q'_b(T) = \sqrt{2qN_A\epsilon_{Si}| - 2\phi_{sub}(T)|} = \sqrt{4qN_A\epsilon_{Si}}\sqrt{\frac{kT}{q}}\sqrt{ln\frac{N_A}{n_i(T)}} \tag{B.5}$$

Inserting this in B.4 and separating the logarithms of fractions into sums of logarithms, we obtain

$$\begin{aligned} V_{thn}(T) &= \frac{kT}{q}(-ln(N_{D,poly}) + ln(n_i(T)) + ln(N_A) - ln(n_i(T))\,) \\ &+ \sqrt{\frac{4kqN_A\epsilon_{Si}}{q}}\sqrt{Tln(N_A) - Tln(n_i(T))} - \frac{Q'_{if}}{C'_{ox}} \\ &= \frac{kT}{q} ln(\frac{N_A}{N_{D,poly}}) \\ &+ \sqrt{\frac{4kqN_A\epsilon_{Si}}{q}}\sqrt{Tln(N_A) - Tln(c_{Si}T^{3/2}exp(\frac{-E_g}{2kT}))} - \frac{Q'_{if}}{C'_{ox}} \end{aligned} \tag{B.6}$$

by using the n_i dependency as mentioned above, with c_{Si} a constant set by silicon

density of states. Then

$$\begin{aligned} V_{thn}(T) &= \frac{kT}{q} ln(\frac{N_A}{N_{D,poly}}) \\ &+ \sqrt{\frac{4kqN_A\epsilon_{Si}}{q}} \sqrt{Tln(N_A) - Tln(c_{Si}) - 1.5Tln(T) - \frac{E_g}{2k}} - \frac{Q'_{if}}{C'_{ox}} \\ &= aT + b\sqrt{cT + 1.5Tln(T) + d} + e \ . \end{aligned} \tag{B.7}$$

Here we redefined the constants as a, b, c, d and e. Inserting B.7 into B.1 and defining c_1 as a further constant, we arrive at Eqn. 4.25:

$$I_D(T) = c_1 T^{-2.5} [V_{gs} - aT - b\sqrt{cT - 1.5T\ln(T) + d} - e\ ]^2 (1 + \lambda(V_{DS} - V_{DS,sat})) \ . \tag{B.8}$$

Appendix C

Example Lab Template Excerpts from “An Experiental Introduction to Electrical and Computer Engineering”

160 *APPENDIX A. BLANK LAB REPORTS*

A.4 3.3: The Voltage Drop Across A Diode

Fill in the table by going through steps 1 through 6.

V_{supply} (V)	V_D (V)	V_R (V)	I (A)
5			
2			
3			
6			
8			
10			
1			
0.7			
0.5			
0.3			

Plot the current vs. voltage across the diode on the following plot. Label each point on the axes clearly.

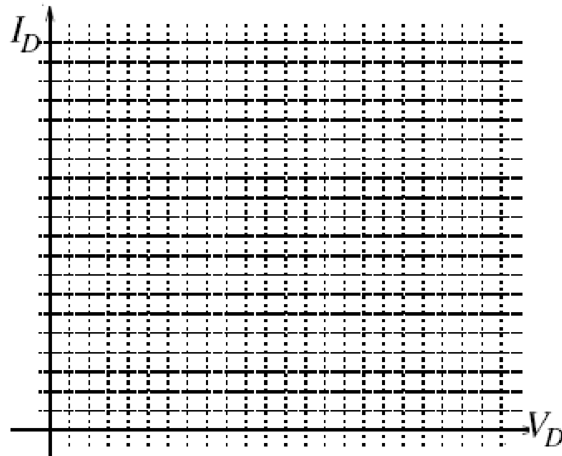


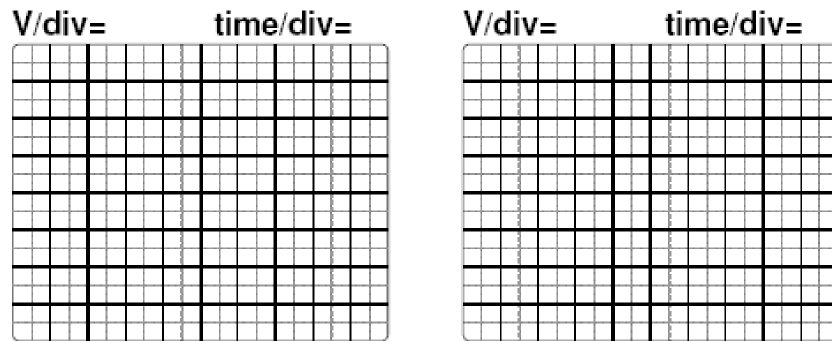
Figure C.1: Lab report template excerpt example for an experiment studying the voltage drop across a diode under different biases.

A.5. 3.4: THE HALF-WAVE RECTIFIER 161

A.5 3.4: The Half-Wave Rectifier

8. First sketch what you would expect the oscilloscope displays with the square and the triangle inputs to be. Then sketch the actual displays by changing the waveform of the signal generator. Do they look like what you expected?

Square Input



Triangle Input

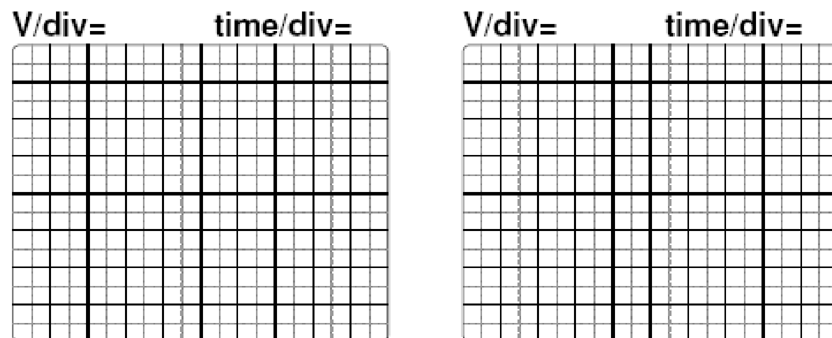


Figure C.2: Lab report template excerpt example for an experiment studying the behavior of a half-wave rectifier circuit.

Bibliography

[1] Akasaka, Y. "Three-dimensional IC trends," *Proceedings of the IEEE*, v. 74, no. 12, pp. 1703-1714, Dec. 1986.

[2] Neudeck, G. "Three-dimensional CMOS integration," *IEEE Circuits and Devices Magazine*, v. 6, no. 5, pp. 32-38, Sept. 1990.

[3] Al-Sarawi, S.F., Abbott, D. and Franzon, P.D. "A review of 3-D packaging technology," *IEEE Transactions on Components, Packaging, and Manufacturing Technology—Part B*, v. 21, no. 1, pp. 2-14, Feb. 1998.

[4] Samavedam, A., Mayaram, K. and Fiez, T., "A scalable substrate noise coupling model for design of mixed-signal IC's", *IEEE Journal of Solid-State Circuits*, vol. 35, no. 6, pp. 895-904, June 2000.

[5] Yun Bai, *Radio-Frequency Effects in Integrated Circuits*, M.Sc. Thesis, University of Maryland, College Park, 2003.

[6] Benthien, S., Lulé, T., Schneider, B., Wagner, M., Verhoeven, M. and Böhm, M. "Vertically integrated sensors for advanced imaging applications," *IEEE Journal of Solid-State Circuits*, v. 35, no. 7, pp. 939-945, July 2000.

[7] Nishikawa, K., Kamogawa, K., Piernas, B., Tokumitsu, M., Sugitani, S., Toyoda, I. and Araki, K. "Three-dimensional MMIC technology for low-cost millimeter-wave MMICs," *IEEE Journal of Solid-State Circuits*, v. 36, no. 9, pp. 1351-1359, Sept. 2001.

[8] Rahman, A. and Reif, R. "System-level performance evaluation of three-dimensional integrated circuits," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, v. 8, no. 6, pp. 671-678, Dec. 2000.

[9] Zhang, R., Roy, K., Koh, C.-K. and Janes, D.B., "Stochastic interconnect modeling, power trends, and performance characterization of 3-D circuits," *IEEE Trans. on Electron Devices*, v. 48, no. 4, pp. 638-652, Apr. 2001.

[10] Rahman, A. and Reif, R., "Thermal analysis of three-dimensional (3-D) integrated circuits (ICs)," in *IEEE 2001 International Interconnect Technology Conference*, pp. 157-159, June 2001.

[11] The MOSIS Service web site, " Fabrication Processes Available Through MOSIS," http://www.mosis.com/products/fab/vendors, last visited in June 2007.

[12] Baker, R. J., Li, H.W. and Boyce, D.E.. *CMOS Circuit Design, Layout and Simulation*, IEEE Press, New York, 1998.

[13] Meuris, P. and Schoenmaker, W., "Strategy for electromagnetic interconnect modeling," *IEEE Trans. on Computer-Aided Design of Integr. Circuits and Systems*, v. 20, no. 6, pp. 753-762, June 2001.

[14] Hajimiri, A. "Distributed integrated circuits: An alternative approach to high-frequency design", *IEEE Communications Magazine*, v. 40, no. 2, pp. 168-173, Feb. 2002.

[15] Hasegawa, H., Furukawa, M. and Yanai, H., "Properties of microstrip line on $Si-SiO_2$ system," *IEEE Trans. on Microwave Theory and Tech.*, v. MTT-19, no. 11, pp. 869-881, Nov. 1971.

[16] Wheeler, H.A., "Transmission-line properties of a strip on a dielectric sheet on a plane," *IEEE Trans. on Microwave Theory and Tech.*, v. MTT-25, no. 8, pp. 631-647, Aug. 1977.

[17] Wee, J.-K., Park, Y.-J., Min, H.-S., Cho, D.-H., Seung, M.-H. and Park, H.-S. "Modeling the substrate effect in interconnect line characteristics of high-speed VLSI circuits," *IEEE Trans. on Microwave Theory and Tech.*, v. 46, no. 10, pp. 1436-1443, Oct. 1998.

[18] Zheng, J., Tripathi, V.K. and Weisshaar, A. "Characterization and modeling of multiple coupled on-chip interconnects on silicon substrate," *IEEE Trans. on Microwave Theory and Tech.*, v. 49, no. 10, pp. 1733-1739, Oct. 2001.

[19] Weisshaar, A., Lan, H. and Luoh, A., "Accurate closed-form expressions for the frequency-dependent line parameters of on-chip interconnects on lossy silicon substrate", *IEEE Trans. on Adv. Packaging*, vol. 25, no. 2, pp.288-296, May 2002.

[20] Kleveland, B., Qi, X., Madden, L., Furusawa, T., Dutton, R.W., Horowitz, M.A. and Wong, S.S. "High-frequency characteristics of on-chip digital interconnects," *IEEE Journal of Solid-State Circuits*, v. 37, no. 6, pp. 716-725, June 2002.

[21] Jiao, D., Mazumder, M., Chakravarty, S., Dai, C., Kobrinsky, M.J., Harmes, M.C., List, S., "A novel technique for full-wave modeling of large-scale three-dimensional high-speed on/off-chip interconnect structures," in *2003 Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD)*, pp. 39-42, Sept. 2003.

[22] Wolbert, P.B., Wachutka, G.K.M., Krabbenberg, B.H. and Mouthaan, T.J.,"Nonisothermal device simulation using the 2-D numerical process/device simulator TRENDY and application to SOI-devices," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, v. 13, no. 3, pp. 293-302, Mar. 1994.

[23] Ozisik, M.N., *Heat Conduction*, Wiley & Sons, 1980.

[24] Link, G.M. and Vijaykrishnan, N., "Thermal trends in emerging technologies," in *Proceedings of the 7th International Symposium on Quality Electronic Design (ISQED'06)*, pp. 625-632, Mar. 2006.

[25] Yuan, L., Leventhal, S. and Qu, G., *Temperature-aware leakage minimization techniques for real-time systems,* University of Maryland Institute for Advanced Computer Studies (UMIACS), available at http://hdl.handle.net/1903/3032, Jan. 2006.

[26] Skadron, K., Abdelzaher, T. and Stan, M.R., "Control-theoretic techniques and thermal-RC modeling for accurate and localized dynamic thermal management," in *The 8th International Symposium on High-Performance Computer Architecture (HPCA'02)*, pp. 17-28, Feb. 2002.

[27] Naveh, A., Rotem, E., Mendelson, A., Gochman, S., Chabukswar, R., Krishnan, K., and Kumar, A., "Power and thermal management in the Intel Core Duo processor," *Intel Technology Journal,* v. 10, no. 2, pp. 109-122, May 2006.

[28] Boumaiza, S. and Ghannouchi, F.M., "Thermal memory effects modeling and compensation in RF power amplifiers and predistortion linearizers," *IEEE Trans. on Microwave Theory and Tech.,* v. 51, no. 12, pp. 2427-2433, Dec. 2003.

[29] Nenadovic, N., Mijalkovic, S., Nanver, L.K., Vandamme, L.K.J., d'Alessandro, V., Schellevis, H. and Slotboom, J.W., "Extraction and modeling of self-heating and mutual thermal coupling impedance of bipolar transistors," *IEEE Journal of Solid-State Circuits,* v. 39, no. 10, pp. 1764-1772, Oct. 2004.

[30] Akturk, A., Allnutt, J., Dilli, Z., Goldsman, N. and Peckerar, M., "Device modeling at cryogenic temperatures: Effects of incomplete ionization," *IEEE Trans. on Electron Devices*, v. 54, no. 11, pp. 2984-2990, Nov. 2007.

[31] Akturk, A., Goldsman, N., Dilli, Z. and Peckerar, M., "Effects of cryogenic temperatures on small-signal MOSFET capacitances," to be presented in *International Semiconductor Device Research Symposium (ISDRS) 2007*, College Park, MD, Dec. 2007.

[32] Akturk, A., *Thermal and Performance Modeling of Nanoscale MOSFETs, Carbon Nanotube Devices and Integrated Circuits*, Ph.D. Dissertation, University of Maryland, College Park, 2006.

[33] Hung, W.-L., Link, G.M., Xie, Y., Vijaykrishnan, N. and Irwin, M.J.,"Interconnect and thermal-aware floorplanning for 3-D microprocessors," in *Proceedings of the 7th International Symposium on Quality Electronic Design (ISQED'06)*, pp. 98-104, Mar. 2006.

[34] Lee, T. and Wong, S.S., "CMOS RF integrated circuits at 5 GHz and beyond," *Proceedings of the IEEE,* v. 88, no. 10, pp. 1560-1571, Oct. 2000.

[35] Burghartz, J.N. "Integrated multilayer RF passives in silicon technology," in *1998 IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems Digest of Papers*, pp. 141-147, Sept. 1998.

[36] Burghartz, J.N. and Rejaei, B. "On the design of RF spiral inductors on silicon," *IEEE Trans. on Electron Devices*, v. 50, no. 3, pp. 718-729, Mar. 2003.

[37] Gao, W. and Yu, Z. "Scalable compact circuit model and synthesis for RF CMOS spiral inductors," *IEEE Trans. on Microwave Theory and Tech.*, v.54, no. 3, pp. 1055-1064, Mar. 2006.

[38] IBM Corporation, *CMOS8RF (CMRF8SF) Design Manual*, Aug. 2006.

[39] Zolfaghari, A., Chan, A. and Razavi, B., "Stacked inductors and transformers in CMOS technology," *IEEE Journal of Solid-State Circuits*, v. 36, no. 4, pp. 621-628, Apr. 2001.

[40] Koutsoyannopoulos, Y.K. and Papananos, Y. "Systematic analysis and modeling of integrated inductors and transformers," *IEEE Trans. on Circuits and Systems–II: Analog and Digital Signal Processing*, v. 47, no. 8, pp.699-713, Aug. 2000.

[41] Zhang, H., Gao, H. and Li, G.-P., "Broad-band power amplifier with a novel tunable output matching network," *IEEE Trans. on Microwave Theory and Tech.*, v. 53, no. 11, pp. 3606-3611, Nov. 2005.

[42] ABET Engineering Accreditation Commission, *Criteria for Accrediting Engineering Programs, Effective for Evaluations during the 2004-2005 Accreditation Cycle.* Current (2006-2007) criteria available at http://www.abet.org.

[43] Muller, R.S. and Kamins, T.I. *Device Electronics for Integrated Circuits*, 2nd ed., John Wiley & Sons, USA, 1986.

[44] Burns, J.A., Aull, B.F., Chen, C.K., Chen, C.-L, Keast, C.L., Knecht, J.M., Suntharalingam, V., Warner, K., Wyatt, P.W. and Yost, D.-R.W., "A wafer-scale 3-D circuit integration technology," *IEEE Trans. on Electron Devices,* v. 53, no. 10, pp. 2507-2516, Oct. 2006.

[45] Dilli, Z., Goldsman, N., Peckerar, M. and Metze, G., "Realization of self-powered electronics by 3-D integration," in *2005 International Semiconductor Device Research Symposium*, pp. 324-325, Bethesda, MD, Dec. 7-9 2005.

[46] Dilli, Z., Goldsman, N., Peckerar, M, Akturk, A. and Metze, G., "Design and testing of a self-powered 3D integrated SOI CMOS system," to be published in *Microelectronics Engineering*. Digital object identifier: doi:10.1016/j.mee.2007.07.013

[47] MITLL Low-Power FDSOI CMOS Process: Application Notes, Revision 2004:1 (December 2004)

[48] Saleh, B.E.A. and Teich, M.C., *Fundamentals of Photonics*, Wiley & Sons, USA, 1991.

[49] "How Solar Energy Works," http://www.ucsusa.org/clean_energy/renewable_energy/page.cfm?pageID=79, visited on April 25, 2005.

[50] Li, H.H., "Refractive index of Silicon and Germanium and its wavelength and temperature derivatives," *J. of Physical and Chemical Data*, v. 9, no. 3, pp. 561-658, 1980.

[51] Dr. Craig Keast and Dr. Peter Wyatt, personal communication, 2005.

[52] Virtuoso Spectre Circuit Simulator, http://www.cadence.com/products/custom_ic/spectre/index.aspx, last visited March 2007.

[53] MITLL Low-Power FDSOI CMOS Process: Device Models, Revision 2005:1, Jan. 2005.

[54] MITLL Low-Power FDSOI CMOS Process: Design Guide, Revision 2005:3, Apr. 2005.

[55] MITLL Low-Power FDSOI CMOS Process: Design Guide, Revision 2006:5, Sept. 2006.

[56] MITLL Low-Power FDSOI CMOS Process: Application Notes, Revision 2006:1, June 2006.

[57] Philipose, M., Smith, J.R., Jiang, B., Mamishev, A. and Sindora-Rajan, K., "Battery-free wireless identification and sensing," *Pervasive Computing,* v. 4, no. 1, pp. 37-45, Jan.-March 2005.

[58] Paradiso, J.A. and Starner, T. "Energy scavenging for mobile and wireless electronics," *Pervasive Computing,* v. 4, no. 1, pp. 18-27, Jan.-Mar. 2005.

[59] Seo, H.-M., Moon, Y., Park, Y.-K., Kim, D., Kim, D.-S., Lee, Y.-S., Won, K.-H., Kim, S.-D. and Choi, P., "A low power fully CMOS integrated RF transceiver IC for wireless sensor networks," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, v. 15, no. 2, pp. 227-231, Feb. 2007.

[60] Wang, C. and Devos, F., "Design and implementation of electrical-supply-free VLSI circuits," *IEE Proceedings: Circuits, Devices, Systems*, v. 152, no. 3, pp. 272-278, June 2005.

[61] Xue, L., Liu, C.C., Kim, H.-S., Kim, S.(K.) and Tiwari, S. "Three-dimensional integration: Technology, use, and issues for mixed-signal applications," *IEEE Trans. on Electron Devices*, v. 30, no. 3, pp. 601-609, Mar. 2003.

[62] Lea, R.M., Jalowiecki, I.P., Boughton, D.K., Yamaguchi, J.S., Pepe, A.A., Ozguz, V.H. and Carson, J.C. "A 3-D stacked chip packaging solution for miniaturized massively parallel processing", *IEEE Trans. on Adv. Packaging*, v. 22, no. 3, pp. 424-432, Aug. 1999.

[63] Coello-Vera, A., Albinet, S. and Venet, N. "Hybrid 640 MBits 3D memory stack," in *1999 Int. Conf. on High Density Packaging and MCMs*, pp. 91-96, Jan. 1999.

[64] Caterer, M.D., Daubenspeck, T.H., Ference, T.G., Holmes, S.J. and Quinn, R.M. "Processing thick multilevel polyimide films for 3-D stacked memory" *IEEE Trans. on Adv. Packaging*, v. 22, no. 2, pp. 189-199, May 1999.

[65] Carson, J., "The emergence of stacked 3D silicon and its impact on microelectronics systems integration," in *1996 IEEE Int. Conf. on Innovative Systems in Silicon*, pp. 1-8, Oct. 1996.

[66] Gutmann, R.J., Lu, J.-Q., Devarajan, S., Zeng, A.Y. and Rose, K. “Wafer-level three-dimensional monolithic integration for heterogeneous silicon ICs,” in *2004 Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems*, pp. 45-48, Sept. 2004.

[67] Bajaj, J., “HgCdTe infrared detectors and focal plane arrays,” in *1998 Conf. on Optoelectronic and Microelectronic Materials and Devices*, pp. 23-31, Dec. 1998.

[68] Koyanagi, M.,Nakamura, T., Yamada, Y., Kikuchi, H.,Fukushima, T., Tanaka, T. and Kurino, H., “Three-dimensional integration technology based on wafer bonding with vertical buried interconnections,” *IEEE Trans. on Electron Devices,* v. 53, no.11, pp. 2799-2808, Nov. 2006.

[69] Knecht, J.M., Yost, D.R.W., Burns, J.A., Chen, C.K., Keast, C.L. and Warner, K., “3D via etch development for 3D circuit integration in FDSOI,” in *2005 IEEE International SOI Conference*, pp. 104-105, 3-6 Oct. 2005.

[70] Pinel, S., Tasselli, J., Marty, A. and Bailbe, J.-P. “A new ultra-thin 3D integration technique: Technological and thermal investigations,” *Design, Test, Integration, and Packaging of MEMS/MOEMS*, Paris, France, v. 4019, pp. 324-332, May 2000.

[71] Culurciello,E. and Andreou, A.G., “Capacitive coupling of data and power for 3-D silicon-on-insulator VLSI,” in *IEEE Int. Symp. on Circuits and Systems (ISCAS) 2005*, v. 4, pp. 4142-4145., May 23-26, 2005.

[72] Kleiner, M.B., Kühn, S.A., Ramm, P. and Weber, W., ”Performance improvement of the memory hierarchy of RISC-systems by application of 3-D technology,” *IEEE Trans. on Components, Packaging, and Manufacturing Technology—Part B*, v. 19, no. 4, pp. 709-718, Nov. 1996.

[73] Zhang, R., Roy, K. and Janes, D.B. “Architecture and performance of 3-dimensional SOI circuits,” in *1999 IEEE Int. SOI Conference*, pp. 44-45, Oct. 1999.

[74] Suntharalingam, V., Berger, R., Burns, J.A., Chen, C.K., Keast, C.L., Knecht, J.M., Lambert, R.D., Newcomb, K.L., O’Mara, D.M., Rathman, D.D., Shaver, D.C., Soares, A.M., Stevenson, C.N., Tyrrell, B.M., Warner, K., Wheeler, B.D., Yost, D-R.W. and Young, D.J. “Megapixel CMOS image sensor fabricated in three-dimensional integrated circuit technology,” in *2005 IEEE Int. Solid-State Circuits Conference*, pp. 356-357, Feb. 2005.

[75] Joyner, J.W. and Meindl, J.D. "Opportunities for reduced power dissipation using three-dimensional integration," in *IEEE 2002 Int. Interconnect Technology Conf.*, pp. 148-150, June 2002.

[76] Kim, S.K., Liu, C.C., Xue, L. and Tiwari, S. "Crosstalk reduction in mixed-signal 3-D integrated circuits with interdevice layer ground planes," *IEEE Trans. on Electron Devices,* v. 52, no. 7, pp. 1459-1467, July 2005.

[77] Alam, S.M., Troxel, D.E. and Thompson, C.V. "A comprehensive layout methodology and layout-specific circuit analyses for three-dimensional integrated circuits," in *Int. Symposium on Quality Electronic Design*, pp. 246-251, Mar. 2002.

[78] Jain, V.K. and Horiguchi, S. "VLSI considerations for TEST: A new hierarchical interconnection network for 3-D integration," *IEEE Trans. on Very Large Scale Integration (VLSI) Systems*, v. 6, no.3 , pp. 346-353, Sept. 1998.

[79] Alkalai, L. "Advanced packaging technologies used in NASA's 3D space flight computer", in *1999 International Conference on High Density Packaging and MCMs*, Denver, Colorado, v. 3830, pp. 85-90, 1999.

[80] Kanbach, H., Wilde, J., Krievel, F. and Meusel, E. "3d Si-on-Si stack package", in *1999 International Conference on High Density Packaging and MCMs*, Denver, Colorado, v. 3830, pp. 248-254, 1999.

[81] Strasser, M., Aigner, R., Lauterbach, C., Sturm, T.F., Franosch, W. and Wachutka, G., "Micromachined CMOS thermoelectric generators as on-chip power supply," in *12th Int. Conf. On Solid State Sensors, Actuators and Microsystems*, Boston, MA, pp. 45-48., June 2003.

[82] Guilar, N., Chen, A., Kleeburg, T. and Amirtharajah, R., "Integrated solar energy harvesting and storage," in *The Int. Symposium on Low Power Electronics and Design (ISLPED) 2006*, pp. 20-24, Tegernsee, Germany, Oct. 2006.

[83] Kaya, T., Koser, H. and Culurciello, E. "Low-voltage temperature sensor for micro-power harvesters in silicon-on-sapphire CMOS," *Electronics Letters*, v. 42, no. 9, pp. 526-528, Apr. 2006.

[84] Lal, A., Duggirala, R. and Li, H. "Pervasive power: A radioisotope-powered piezoelectric generator," *Pervasive Computing*, v. 4, no. 1, pp. 53-61, Jan.-Mar. 2005.

[85] McSpadden, J.O., Yoo, T. and Chang, K. "Theoretical and experimental investigation of a rectenna element for aicrowave power transmission," *IEEE Trans. on Microwave Theory and Tech.*, v. 40, no. 12, pp. 2359-2366, Dec. 1992.

[86] McSpadden, J.O., Fan, L. and Chang, K. "Design and experiments of a high-conversion-efficiency 5.8 GHz rectenna", *IEEE Trans. on Microwave Theory and Tech.*, v. 46, no. 12, pp. 2053-2060, Dec. 1998.

[87] Zbitou, J., Latrach, M. and Toutain, S. "Hybrid rectenna and monolithic integrated zero-bias microwave rectifier," *IEEE Trans. on Microwave Theory and Tech.*, v. 54, no. 1, pp. 147-152, Jan. 2006.

[88] Hagerty, J.A., Helmbrecht, F.B., McCalpin, W.H., Zane, R. and Popović, Z.B., "Recycling ambient microwave energy with broad-band rectenna arrays", *IEEE Trans. on Microwave Theory and Tech.*, v. 52, no. 3, pp. 1014-1023, Mar. 2004.

[89] Joe, J. and Chia, M.Y.W., "Voltage, efficiency calculation and measurement of low power rectenna rectifying circuit", in *IEEE Antenna and Propagation Society International Symposium*, v. 4, pp. 1854-1857, June 1998.

[90] Ryan, D.M., LaFollette, R.M. and Salmon, L., "Microscopic batteries for micro electromechanical systems (MEMS)," in *Proceedings of the 32nd Intersociety Energy Conversion Engineering Conference (IECEC-97)*, v. 1, pp. 77-82, July-Aug. 1997.

[91] Pushparaj, V.L., Shaijumon, M.M., Kumar, A., Murugesan, S., Ci, L., Vajtai, R., Linhardt, R.J., Nalamasu, O. and Ajayan, P.M., "Flexible energy storage devices based on nanocomposite paper," *Proceedings of the National Academy of Sciences,* v. 104, no. 34, pp. 13574-13577, Aug. 2007.

[92] Ngu, Y., Peckerar, M., Dilli, Z. and Goldsman, N. "High capacitance battery for powering distributed networks node devices," to be presented in *International Semiconductor Device Research Symposium (ISDRS) 2007*, College Park, MD, Dec. 2007.

[93] Shao, X.,Goldsman, N., Ramahi, O. and Guzdar, P., "A new method for simulation of on-chip interconnects and substrate currents with 3D alternating-direction-implicit (ADI) Maxwell solver," in *International Conf. on Simulation of Semiconductor Processes and Devices (SISPAD)'03*, pp. 315-317, Dec. 2003.

[94] Shao, X.,Goldsman, N., Ramahi, O. and Guzdar, P., "Modeling RF effects in integrated circuits with a new 3D alternating-direction-implicit Maxwell

equation solver," in *International Semiconductor Device Research Symposium (ISDRS) 2003*, pp. 532-533, Dec. 2003.

[95] Yang, B., Shao, X., Goldsman, N., Ramahi, O. and Guzdar, P.N., "Full Wave Modeling of Substrate Doping Effects and Nonideal Conductors in Integrated Circuit Interconnects," in *International Semiconductor Device Research Symposium (ISDRS) 2005*, pp. 368 - 369, Dec. 2005.

[96] Dilli, Z., Goldsman, N. and Akturk, A., "An impulse-response based methodology for modeling complex interconnect networks," in *International Semiconductor Device Research Symposium (ISDRS) 2005*, pp. 64-65, Dec. 2005.

[97] Dilli, Z., Goldsman, N., Akturk, A. and Metze, G., "A 3-D time-dependent Green's Function approach to modeling electromagnetic noise in on-chip interconnect networks," in *2006 Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD '06)*, pp. 240-243, Sept. 2006.

[98] Oppenheim, A.V., Willsky, A.S. and Young, I.T. *Signals and Systems,* Prentice Hall, USA, 1983.

[99] "SPICE Algorithm Overview," http://www.ecircuitcenter.com/SpiceTopics/Overview/Overview.htm, visited in Feb. 2006.

[100] Kumashiro, S., Rohrer, R.A. and Strojwas, A.J. "A new efficient method for the transient simulation of three-dimensional interconnect structures," in *IEEE Electron Devices Meeting*, pp. 193-196, Dec. 1990.

[101] "MOSIS Test Results for AMIS 0.50 Micron Runs (C5)", http://www.mosis.org/Technical/Testdata/ami-c5-prm.html, visited in June 2006.

[102] Zheng, J., Hahm, Y.-C., Tripathi, V.K. and Weisshaar, A., "CAD-oriented equivalent-circuit modeling of on-chip interconnects on lossy silicon substrate," *IEEE Trans. on Microwave Theory and Tech.*, v. 48, no. 9, pp. 1443-1451, Sept. 2000.

[103] Akturk, A., Goldsman, N. and Metze, G. "Coupled modeling of time-dependent full-chip heating and quantum non-isothermal device operation," in *2003 Int. Conf. on Simulation of Semiconductor Processes and Devices (SISPAD'03)*, pp. 311-314, Sept. 2003.

[104] Akturk, A., Parker, L., Goldsman, N. and Metze, G. "Mixed-mode simulation of non-isothermal quantum device operation and full-chip heating," in

International Semiconductor Device Research Symposium (ISDRS) 2003, pp. 508-509, Dec. 2003.

[105] Akturk, A., Goldsman, N., Parker, L. and Metze, G. "Mixed-mode temperature modeling of full-chip based on individual non-isothermal device operations," *Solid-State Electronics*, v. 49, no. 7, pp. 1127-1134, July 2005.

[106] Akturk, A., Goldsman, N. and Metze, G. "Self-consistent modeling of heating and MOSFET performance in 3-D integrated circuits," *IEEE Trans. on Electron Devices*, v. 52, no. 11, pp. 2395-2403, Nov. 2005.

[107] Dilli, Z., Akturk, A. and Goldsman, N. "Controlled localized heating on integrated circuits for cold-ambient temperature applications," to be presented in *International Semiconductor Device Research Symposium (ISDRS) 2007*, College Park, MD, Dec. 2007.

[108] Wallace, P.R. *Mathematical Analysis of Physical Problems,* Dover, USA, 1984.

[109] Gelsinger, P.,"Microprocessors for the new millenium: challenges, opportunities, and new frontiers," in *International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, pp. 22-25, 2001.

[110] Ronen, R., Mendelson, A., Lai, K., Lu, S-L., Pollack, F. and Shen, J.P., "Coming challenges in microarchitecture and architecture," *Proc. of the IEEE*, v. 89, no. 3, pp. 325-340, 2001.

[111] Im, S. and Banerjee, K., "Full chip thermal analysis of planar (2D) and vertically stacked integrated (3D) high performance ICs," in *International Electron Devices Meeting (IEDM) Technical Digest*, pp. 727-730, 2000.

[112] Akturk, A., Goldsman, N., Dilli, Z. and Peckerar, M. "Device performance and package induced self heating effect at cryogenic temperatures," in *International Conference on Simulation of Semiconductor Processes and Devices (SISPAD) 2006*, pp. 240-243, Monterey Bay, CA, Sept. 2006.

[113] Yang, Y., Gu, Z., Zhu, C., Dick, R. and Shang, L., "ISAC: Integrated space-and-time-adaptive chip-package thermal analysis," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, v. 26, no. 1, pp. 86-99, Jan. 2007.

[114] Swart, N.R. and Nathan, A.,"Coupled electrothermal modeling of microheaters using SPICE," *IEEE Trans. on Electron Devices,* v. 41, no. 6, pp. 920-924, June 1994.

[115] Anderson, G., Kim, J., Goldsman, N. and Dilli, Z., "Investigating pool boiling using a novel temperature distribution sensor," in *ECI International Conference on Boiling Heat Transfer*, May 2006.

[116] Elman, H., class notes for AMSC/CMSC 660: Scientific Computing I, University of Maryland, College Park, Fall 2002.

[117] Arora, N.D., Hauser, J.R. and Roulston D.J., "Electron and hole mobilities in silicon as a function of concentration and temperature," *IEEE Trans. on Electron Devices*, v. 30, no. 2, pp. 292-295, July 1983.

[118] Selberherr, S., *Analysis and Simulation of Semiconductor Devices,* Springer-Verlag/Wien, 1984.

[119] Badaroglu, M., van der Plas, G., Wambacq, P., Balasubramanian, L., Tiri, K., Verbauwhede, I., Donnay, S., Gielen, G. and de Man., H. "Digital circuit capacitance and switching analysis for ground bounce in ICs with a high ohmic substrate", *IEEE Journal of Solid-State Circuits*, vol. 39, no. 7, pp.1119-1130, July 2004.

[120] Xu, M., Su, D., Shaeffer, D.K., Lee, T.H. and Wooley, B.A. "Measuring and modeling the effect of substrate noise on the LNA for a CMOS GPS receiver," *IEEE Journal of Solid-State Circuits*, v. 36, no. 3, pp. 473-485, Mar. 2001.

[121] Nagata, M., Nagai, J., Hijikata, K., Morie, T. and Iwata, A. "Physical design guides for substrate noise reduction in CMOS digital circuits," *IEEE Journal of Solid-State Circuits*, v. 36, no. 3, pp. 539-549, Mar. 2001.

[122] Su, D.K., Loinaz, M.J., Masui, S. and Wooley, B.A. "Experimental results and modeling techniques for substrate noise in mixed-signal integrated circuits," *IEEE Journal of Solid-State Circuits*, v. 28, no.4, pp. 420-430, Apr. 1993.

[123] MOSIS Integrated Circuit Fabrication Service, http://www.mosis.org, last visited Oct. 2007.

[124] Ruehli, A.E., "Inductance calculations in a complex integrated circuit environment", *IBM Journal of Research and Development,* vol. 16, no. 5, pp. 470-481, Sept. 1972.

[125] Grover, F.W., *Inductance Calculations*, 1946, re-publication, Dover Publications, New York, 2004.

[126] Bai, Y., Dilli, Z., Goldsman, N. and Metze, G., "Frequency-dependent modeling of on-chip inductors on lossy substrates," in *International Semiconductor Device Research Symposium (ISDRS) 2003*, pp.292-293, Dec. 2003.

[127] Tang, C.-C., Wu, C.-H. and Liu, S.-I., "Miniature 3-D inductors in standard CMOS process", *IEEE Journal of Solid-State Circuits*, v. 37, no.4, pp. 471-479, Apr. 2002.

[128] Yue, C.P. and Wong, S.S., "On-chip spiral inductors with patterned ground shields for Si-based RF IC's", *IEEE Journal of Solid-State Circuits*, v. 33, no. 5, pp. 743-752, May 1998.

[129] Ng, K.T., Rejaei, B. and Burghartz, J.N., "Substrate effects in monolithic RF transformers on silicon," *IEEE Trans. on Microwave Theory and Tech.,* v. 50, no. 1, pp. 377-383, Jan. 2002.

[130] Ng., K.T., Rejaei, B., de Kruijff, T.R., Soyuer, M. and Burghartz, J.N., "Analysis of generic spiral-coil RF transformers on silicon," in *2000 IEEE Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems Digest of Papers*, pp. 103-107, 2000.

[131] Yue, C.P. and Wong, S.S., "Physical modeling of spiral inductors on silicon," *IEEE Trans. on Electron Devices,* v. 47, no. 3, pp. 560-568, Mar. 2000.

[132] Pozar, D.M., *Microwave Engineering*, 2nd ed., John Wiley & Sons, USA, 1998.

[133] O, K., "Estimation methods for quality factors of inductors fabricated in silicon integrated circuit process technologies," *IEEE Journal of Solid-State Circuits*, v. 33, no. 8, pp. 1249-1252, Aug. 1998.

[134] "Introduction to Bipolar Device GHz Measurement Techniques," Cascade Microtech Application Note.

[135] Fraser, A., Gleason, R. and Strid, E.W., "GHz on-silicon-wafer probing calibration methods," in *IEEE 1988 Bipolar Circuits & Technology Meeting*, paper no. 7.6, pp.154-156, 1988.

[136] Lee, T.H., *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 1998.

[137] Nguyen, N.M. and Meyer, R.G., "Si IC-compatible inductors and LC passive filters," *IEEE Journal of Solid-State Circuits*, v. 25, no.4, pp. 1028-1031, Aug. 1990.

[138] Li, S.-H., Jang, S.-L., Chuang, Y.-S. and Li, C.-F., "A new LC-tank voltage controlled oscillator," in *IEEE Asia-Pacific Conference on Circuits and Systems*, pp. 425-427, Dec. 2004.

[139] Yin, W.-Y., Pan, S.J., Li, W. and Gan, Y.B., "Experimental characterization of coupling effects between two on-chip neighboring square inductors", *IEEE Trans. on Electromagnetic Compatibility*, v. 45, no.3, pp. 557-561, Aug. 2003.

[140] Zhan, Y., Harjani, R. and Sapatnekar, S.S., "On the selection of on-chip inductors for the optimal VCO design," in *IEEE 2004 Custom Integrated Circuits Conference*, pp. 277-280, Oct. 2004.

[141] Mayevskiy, Y., Watson, A., Francis, K., Hwang, K. and Weisshaar, A., "A new compact model for monolithic transformers in silicon-based RFICs," *IEEE Microwave and Wireless Component Letters*, v. 15, no. 6, pp.419-421, June 2005.

[142] Gil, J. and Shin, H., "A simple wide-band on-chip inductor model for silicon-based RF ICs," *IEEE Trans. on Microwave Theory and Tech.*, v. 51, no. 9, pp. 2023-2028, Sept. 2003.

[143] Melendy, D. and Weisshaar, A., "A new scalable model for spiral inductors on lossy silicon substrate," in *2003 IEEE MTT-S International Microwave Symposium Digest*, vol. 2, pp. 1007 - 1010, June 2003.

[144] Bantas, S., Koutsoyannopoulos, Y. and Liapis, A., "An inductance modeling flow seamlessly integrated in the RF IC design chain," in *Design, Automation and Test in Europe Conference and Exhibition*, v. 3, pp.16-20, Feb. 2004.

[145] Ramo, S., Whinnery, J.R. and van Duzer, T., *Fields and Waves in Communication Electronics*, 2nd ed., John Wiley & Sons, USA, 1984.

[146] Hirsch,L.S., Kimmel, H., Rockland, R. amd Bloom, J. "Implementing pre-engineering curricula in high school science and mathematics," in *Frontiers in Education (FIE) 2005*, pp. S2F-21-26, Oct. 2005.

[147] Okyere, E., Rothkopf, E., Corti, F. and Beninati, C., "Innovative approaches to student retention in the technologies," in *Frontiers in Education (FIE) 1991*, pp. 351-354, Sept. 1991.

[148] Sash, R., Detloff, H., Chen, B., Gendgenett, N. and Duranm D., "Work in progress: Retention of freshmen computer and electronics engineering students," in *Frontiers in Education (FIE) 2006*, pp. 19-20, Oct. 2006.

[149] Dilli, Z., Goldsman, N., Schmidt, J.A., Harper, L. and Marcus, S.L., "A new pedagogy in electrical and computer engineering: An experiential and conceptual approach," in *Frontiers in Education (FIE) 2002*, v. 1, pp. T2C-3-7, Nov. 2002.

[150] Kolb, D. *Kolb Learning Styles Inventory*, McBer & Company, Boston, MA, 1985.

[151] Goldsman, N. and Dilli, Z., *An Experiential Introduction to Electrical and Computer Engineering*, University of Maryland, College Park, 2001.

[152] University of Maryland, College Park, *Undergraduate Catalog,* 2001.

[153] Dr. Janet Schmidt, personal communication, 2001.

[154] Larkin-Hein, T. and Budny, D.D., "Research on learning style: Applications in the physics and engineering classrooms," *IEEE Trans. on Education* ,v. 44, no. 3, pp.276-281, Aug. 2001.

[155] Parcover, J.A. and McCuen, R.H., "Discovery approach to teaching engineering design," *Journal of Professional Issues in Engineering Education and Practice*, vol. 121, no. 4, pp. 236-241, Oct. 1995.

[156] Srikantaiah, D., "The TIME Project's on-going training for PGCPS, Region II math and science teachers," *K-16 Partnership Development Center: Collaborations*, U. of MD, College of Education, no. 4, pp. 3-7, Winter 2005.

[157] Greene, R. M., "The Dean and faculty of the A. James Clark School of engineering capture the attention of 200 students at Arundel High School," *K-16 Partnership Development Center: Collaborations*, U. of MD, College of Education, no. 4, pp.1-6, Winter 2005.

[158] Paulos, J.A., *Innumeracy: Mathematical Illiteracy and Its Consequences*, Vintage, reprint ed., 1990.

[159] Dilli, Z. and Goldsman, N., "Relating electrical and computer engineering to the high school classroom," to be presented in *International Semiconductor Device Research Symposium (ISDRS) 2007,* College Park, MD, Dec. 2007.

[160] Freid, C., "Are required courses meeting industry's demands?" *IEEE Potentials*, v. 70, no. 3, pp. 39-40, Aug./Sept. 2000.

[161] Christiano, S.J.E. and Ramirez, M.R., "Satisficing: Student needs and industry expectations, in *Frontiers in Education (FIE) 1994*, pp.491-495, Nov. 1994.

[162] Ray, J. and Farris, J., "Integration of design and manufacturing processes in first-year engineering curriculums," in *Frontiers in Education (FIE) 2000*, v. 2, pp. S2G7-S2G11, Oct. 2000.

[163] Dilli, Z. and Goldsman, N., "Development and implementation of a multispecialty advanced capstone design course," in *Information Technology Based Higher Education and Training (ITHET) '04*, pp. 425-428, June 2004.