Characterization and Modeling of TSV Based 3-D Integrated Circuits

by

Ioannis Savidis

Submitted in Partial Fulfillment
of the
Requirements for the Degree
Doctor of Philosophy

Supervised by
Professor Eby G. Friedman

Department of Electrical and Computer Engineering
Arts, Sciences and Engineering
Edmund J. Hajim School of Engineering and Applied Sciences

University of Rochester
Rochester, New York
2013

Dedication

To my family. My achievements are but an indication of your love and support.

Biographical Sketch

Ioannis Savidis was born in Rochester, New York in December 1982. He received the B.S.E. degree in electrical and computer engineering and biomedical engineering from Duke University, Durham, NC, in 2005. He received the M.Sc. degree in electrical and computer engineering from the University of Rochester, Rochester, NY, in 2007. He pursued his research in high performance integrated circuits for the Ph.D. degree in electrical and computer engineering under the guidance of Professor Eby G. Friedman.

He interned at Freescale Semiconductor Corporation, Austin, TX during the summers of 2006 and 2007, where he worked on the electrical characterization and modeling of interplane 3-D vias. During the summers of 2008, 2009, 2010, and 2011 he interned at

the IBM T. J. Watson Research Center, Yorktown Heights, NY, where he worked on electrical characterization and modeling, reticle design, and DC and high frequency electrical measurements of test vehicles implementing various interplane 3-D via topologies. His research interests include analysis, modeling, and design methodologies for high performance digital and mixed-signal integrated circuits, emerging integrated circuit technologies, and interconnect related issues, with an emphasis on electrical and thermal modeling and characterization, signal and power integrity, and power and clock delivery for 3-D IC technologies.

The following publications were a result of work conducted during his doctoral study:

PUBLICATIONS

Book Chapter

[1] I. Savidis and E. G. Friedman, "Physical Design Trends for Interconnects," *On-Chip Communication Architectures System on Chip Interconnect*, S. Pasricha and N. Dutt, Morgan Kaufmann Publishers, Elsevier, Chapter 11, pp. 403-437, 2008.

Journal Papers

[1] I. Savidis and E. G. Friedman, "Experimental Analysis of Thermal Coupling in TSV-Based 3-D Integrated Circuits," *IEEE Transactions on Electron Devices*, (in submission).

[2] I. Savidis, S. Kose, and E. G. Friedman, "Power Noise in TSV-Based 3-D Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 48, No. 2, pp. 587-597, February 2013.

[3] B. Ciftcioglu, R. Berman, S. Wang, J. Hu, I. Savidis, M. Jain, D. Moore, M. Huang, E. G. Friedman, G. Wicks, and H. Wu, "3-D Integrated Heterogeneous Intra-Chip

Free-Space Optical Interconnect," *Optics Express*, Vol. 20, No. 4, pp. 4331-4345, February 2012.

[4] V. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Networks in 3-D Integrated Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 19, No. 12, pp. 2256-2266, December 2011.

[5] J. Wang, I. Savidis, and E. G. Friedman, "Thermal Analysis of Oxide-Confined VCSEL Arrarys," *Microelectronics Journal*, Vol. 42, No. 5, pp. 820-825, May 2011.

[6] B. Ciftcioglu, R. Berman, J. Zhang, Z. Darling, S. Wang, J. Hu, J. Xue, A. Garg, M. Jain, I. Savidis, D. Moore, M. Huang, E. G. Friedman, G. Wicks, and H. Wu, "A 3-D Integrated Intra-Chip Free-Space Optical Interconnect for Many-Core Chips," *IEEE Photonics Technology Letters*, Vol. 23, No. 3, pp. 164-166, February 2011.

[7] I. Savidis and E. G. Friedman, "Closed-Form Expressions of 3-D Via Resistance, Inductance, and Capacitance," *IEEE Transactions on Electron Devices*, Vol. 56, No. 9, pp. 1873-1881, September 2009.

Conference Papers

[1] I. Savidis and E. G. Friedman, "Test Circuits for 3-D Systems Integration," *Proceedings of the Government Microcircuit Applications & Critical Technology Conference (GOMACTech)*, pp. 181-184, March 2012.

[2] H. Wu, B. Ciftcioglu, R. Berman, S. Wang, J. Hu, I. Savidis, M. Jain, D. Moore, M. Huang, E. Friedman, and G. Wicks, "Chip-Scale Demonstration of 3-D Integrated Intra-Chip Free-Space Optical Interconnect (Invited Paper)," *Photonics West: Proceedings of SPIE Optoelectronic Integrated Circuits XIV*, Vol. 8265, pp. 82650C-1 82650C-10, February 2012.

[3] I. Savidis, V. Pavlidis, and E. G. Friedman, "Clock Distribution Models of 3-D Integrated Systems," *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2225-2228 May 2011.

[4] I. Savidis, S. Kose, and E. G. Friedman, "Power Grid Noise in TSV-Based 3-D Integrated Systems," *Proceedings of the Government Microcircuit Applications & Critical Technology Conference (GOMACTech)*, pp. 129-132, March 2011.

[5] B. Ciftcioglu, R. Berman, J. Zhang, Z. Darling, A. Garg, J. Hu, M. Jain, P. Liu, I. Savidis, S. Wang, J. Xue, E. G Friedman, M. Huang, D. Moore, G. Wicks, and H. Wu, "Initial Results of Prototyping a 3-D Integrated Intra-Chip Free-Space Optical Interconnect," *Proceedings of the Workshop on the Interaction between Nanophotonic Devices and Systems (WINDS 2010)*, December 2010.

[6] J. Xue, A. Garg, B. Ciftcioglu, J. Hu, S. Wang, I. Savidis, M. Jain, R. Berman, P. Liu, M. Huang, H. Wu, E. Friedman, G. Wicks, and D. Moore, "An Intra-Chip Free-Space Optical Interconnect," *Proceedings of the 37th Annual International Symposium on Computer Architecture (ISCA)*, pp. 94-105, June 2010.

[7] J. Xue, A. Garg, B. Ciftcioglu, S. Wang, I. Savidis, J. Hu, M. Jain, M. Huang, H. Wu, E. G. Friedman, G. W. Wicks, and D. Moore, "An Intra-Chip Free-Space Optical Interconnect," *Proceedings of the 3rd Workshop on Chip Multiprocessor Memory Systems and Interconnects (CMP-MSI '09) held in conjunction with the 36th International Symposium on Computer Architecture*, June 2009.

[8] I. Savidis, E. G. Friedman, V. F. Pavlidis, and G. De Micheli, "Clock and Power Distribution Networks for 3-D Integrated Circuits," *Proceedings of the Workshop on 3D Integration, Design, Automation & Test in Europe Conference*, March 2009.

[9] V. F. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Architectures for 3-D SOI Integrated Circuits," *Proceedings of the IEEE International Silicon-on-Insulator Conference*, pp. 111-112, October 2008.

[10] V. F. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Networks for 3-D Integrated Circuits," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 651-654, September 2008.

[11] I. Savidis and E. G. Friedman, "Electrical Modeling and Characterization of 3-D Vias," *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 784-787, May 2008.

Workshop Poster Presentations

[1] I. Vaisband, S. Kose, I. Savidis, and E. G. Friedman, "On-Chip Power Delivery," *CEIS University Technology Showcase*, Rochester, New York, April 6, 2011.

[2] B. Ciftcioglu, R. Berman, J. Zhang, Z. Darling, S. Wang, J. Hu, J. Xue, A. Garg, M. Jain, I. Savidis, D. Moore, M. Huang, E. G Friedman, G. Wicks, and H. Wu, "3-D Integrated Intra-Chip Free-Space Optical Interconnect," *IEEE International Solid-State Circuits Conference (ISSCC) Student Research Forum*, February 2010.

[3] I. Savidis and E. G. Friedman, "Clock Distribution Topologies for 3-D Integrated Circuits," *CEIS University Technology Showcase*, Rochester, New York, February 12, 2009.

[4] V. F. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Networks for 3-D Integrated Circuits," *Massachusetts Institute of Technology Lincoln Laboratorys 3-D Integrated Circuit Multi-Project Wafer Review*, Boston, Massachusetts, September 15, 2008.

Acknowledgments

As I began writing this section, it became quite evident to me that it is very difficult to express gratitude to all the individuals who have affected my life throughout the years. Various family members, teachers, professors, and close friends have each contributed in molding me into the person that I currently am. Without your support, patience, and encouragement, I would not be where I am today

I would like to begin by sincerely thanking my advisor and mentor, Professor Eby G. Friedman, for his continuous support, encouragement, and direction during the course of my Ph.D. studies. His knowledge, experience, insight, and enthusiasm towards his work has provided me with a model that will guide me through my own research pursuits. The knowledge I have gained goes beyond the basic process of research. Lessons learned on running a laboratory efficiently and productively, properly funding projects, academic writing style, verbal expression, and other nuggets of knowledge that are essential for productive research laboratories are all attributed to my adviser's tutelage. Professor Friedman, it was a pleasure completing my thesis work under your supervision.

I would like to thank Professor Engin Ipek, Professor Hui Wu, Professor Michael Gage, Professor Diane Dalecki, and Professor John Lambropoulos for serving on my proposal and defense committees. Your insightful comments, critical reviews, and helpful advice enhanced both my research and thesis dissertation.

I have had the pleasure of working in an environment with a diverse group of individuals coming from different professional and cultural backgrounds. Members of the High

Performance Integrated Circuit Laboratory made my Ph.D. experience much more pleasurable and enjoyable. I would like to thank previous and current members of the laboratory: Dr. Guoqing Chen, Dr. Mikhail Popovich, Prof. Vasilis F. Pavlidis, Prof. Emre Salman, Dr. Jonathan Rosenfeld, Dr. Renatas Jakushokas, Prof. Selcuk Kose, Ravi Patel, Inna Vaisband, Shen Ge, Boris Vaisband, and Alex Shapiro. The long discussions on a variety of topics have expanded the scope of my professional and personal knowledge. I would also like to thank RuthAnn Williams for her endless effort on my behalf, allowing me to focus my time on my research. The staff in the Department of Electrical and Computer Engineering has always assisted and supported me throughout my time at the University of Rochester, and I thank you all for your efforts. I would also like to thank Dr. Berkehan Ciftcioglu and Prof. Aaron Carpenter for discussions that went beyond technical jargon, and focused on pressing matters such as the best footballer in the world, or whether Duke was going to win another basketball championship. For anyone that I have not mentioned, it is not that I have forgotten you, but rather, there are not enough pages in this manuscript to highlight all the times we have spent together, all the moments we have lived through, all the support we have provided each other. You too hold a special place in my heart.

I also want to express my warmest gratitude to Dr. Robert Jones, Prof. Ankur Jain, Dr. Ritwik Chatterjee, Dr. Syed Alam, Dr. Douglas Reber, and Mr. Brad Smith from Freescale Semiconductor who guided me through two enjoyable summer internships in the beautiful city of Austin, Texas. In addition, I would like to thank Dr. John Knickerbocker, Dr. Steven Wright, Dr. Bing Dang, Dr. Edward Sprogis, Dr. Bucknell Webb, Dr. Chirag Patel, Mr. Paul Andry, Mr. Chris Jahnes, and Mrs. Cornelia Tsang from IBM Research for your guidance during three insightful summer internships in Yorktown Heights, New York. I have formed bonds with you through these internships that have and will continue to remain strong. I look forward to potentially working with all of you again.

Finally, I would like to thank my family, who have provided me with joy and happiness while I was working through long hours in the laboratory. I would especially like to thank my father Konstantinos Savidis, grandmother Stella Savidis, and aunt Melanie and uncle

Michael Savidis for their support and their help watching over my son while I spent time in the laboratory pursuing my Ph.D studies. My brothers Panagioti and Nickolaos, my wife's parents Pedro and Julia Hurtado, and my mother Fotini have provided me with a listening ear on numerous occasions when I just needed another perspective. I would also like to thank my cousins, Yianni, Walter, and Adoni Savidis, for their support during my studies. Most importantly, I owe a special thanks to my wife Ana Lucia and my son Ioannis A. for their love, support, and endless encouragement. You have always brought sunshine to my life.

Abstract

Through silicon via (TSV) based three-dimensional integrated circuits have rapidly progressed over the past decade. TSV based three-dimensional integration has the potential to significantly boost the performance and capabilities of state-of-art integrated circuits, while supporting the integration of disparate heterogeneous technologies. The research presented here provides insight into some of the most pressing issues currently being addressed by the research community, and provides guidelines for designing these evolving heterogeneous 3-D systems.

The organization of the dissertation begins with an introduction to TSV electrical models, power delivery, and thermal behavior in 3-D ICs. Characterization and physical design methodologies for 3-D integrated circuits are discussed. Electrical modeling of TSVs is presented, culminating in the development of closed-form expressions for the TSV resistance, capacitance, and inductance.

Synchronization and power delivery are critical design considerations in 3-D ICs. Models of three distinct clock distribution networks are provided, and a comparison of the power and delay of each topology is presented. Three power delivery topologies are discussed, with experimental evidence describing the effects of the TSV density on the noise profile

of 3-D power delivery networks. A comparison of the peak and average noise for each topology with and without board level decoupling capacitors is provided, and suggestions for enhancing the design of 3-D power delivery networks are offered.

Thermal properties in 3-D integrated circuits are also discussed. The placement of two highly active and aligned circuit blocks has a significant effect on the thermal profile of 3-D ICs. A test circuit exploring thermal coupling between device planes is presented. The experimental results provide insight into heat flow within 3-D ICs.

Three-dimensional integration is an evolving technology that will prolong the semiconductor roadmap for several generations. This dissertation provides insight into the 3-D IC design process, with the goal of strengthening the design capabilities for 3-D integrated circuits and systems.

Contributors and Funding Sources

The research presented in this dissertation was supervised by a committee consisting of Professors Eby G. Friedman (advisor), Engin Ipek, and Hui Wu of the Department of Electrical and Computer Engineering, and Professor John Lambropoulos of the Mechanical Engineering Department. The committee was chaired by Professor Diane Dalecki of the Department of Biomedical Engineering.

The author, I. Savidis, characterized the electrical impedance of the through silicon via, derived closed-form expressions characterizing the electrical impedance of the through silicon via, developed equivalent models of the clock distribution topologies for SPICE simulation, designed test circuits examining power distribution topologies and thermal coupling in 3-D integrated circuits, experimentally examined the power network and thermal coupling test circuits, developed equivalent models of the power distribution networks for SPICE simulation, and assisted in the development of models for the thermal analysis of bundled vertical cavity surface emitting lasers. The introduction, Chapter 1, and the background chapters, Chapters 2 through 4, are based on academic literature presented by other researchers. A chapter by chapter description of contributions from colleagues is listed below.

Chapter 5: I. Savidis is the primary author of this chapter. The development and evaluation of the research were performed in collaboration with co-author E. G. Friedman. The research effort was published in the *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS)* and the *IEEE Transactions on Electron Devices*.

Chapter 6: I. Savidis contributed the equivalent electrical models of the clock distribution topologies used for SPICE simulation. The models were used to compare results with experimentally collected data on clock skew, slew, and power consumption of three different clock topologies. He also assisted V. Pavlidis in the layout and experimental testing of the clock distribution test circuit. The development and evaluation of the research were performed in collaboration with co-author E. G. Friedman. Results from this study were published in four conferences (*Proceedings of the IEEE Custom Integrated Circuits Conference*, *Proceedings of the IEEE International Silicon-on-Insulator Conference*, *Proceedings of the Workshop on 3D Integration, Design, Automation & Test in Europe Conference*, and *Proceedings of the IEEE International Symposium on Circuits and Systems*) and a journal (*IEEE Transactions on Very Large Scale Integration (VLSI) Systems*).

Chapter 7: I. Savidis is the primary author of this chapter. He designed, participated in the layout, and experimentally evaluated the test circuit examining noise propagation in 3-D integrated power distribution networks. He also developed the equivalent electrical models of the power network used to determine the resonant characteristics, noise characteristics, and impedance of the power distribution network through SPICE simulations. The development and evaluation of the research were performed in collaboration with co-author E. G. Friedman. Savidis was assisted by S. Kose in the circuit layout, and by B.

Ciftcioglu, J. Hu, and Professor H. Wu on the experimental testing of the test circuit. Results from this study were published in two conferences (*Proceedings of the Workshop on 3D Integration, Design, Automation & Test in Europe Conference* and *Proceedings of the Government Microcircuit Applications & Critical Technology Conference (GOMACTech)*) and a journal (*IEEE Journal of Solid-State Circuits*).

Chapter 8: I. Savidis is the primary author of this chapter investigating decoupling capacitor placement in 3-D ICs. The development and evaluation of the research were performed in collaboration with co-author E. G. Friedman.

Chapter 9: I. Savidis was a member of a team of twelve Ph.D. students (J. Xue, A. Garg, R. Parihar, B. Ciftcioglu, J. Hu, S. Wang, I. Savidis, M. Jain, R. Berman, P. Liu, J. Zhang, Z. Darling) and five professors (E. G. Friedman, H. Wu, M. Huang, G. Wicks, and D. Moore) that developed a 3-D integrated free space optical interconnect system for direct core-to-core communication. His primary contributions were in the development of models representing the vertical cavity surface emitting lasers (VCSEL) for thermal simulation, as described in Chapter 10. This work was published in four conferences (*Proceedings of the 3rd Workshop on Chip Multiprocessor Memory Systems and Interconnects (CMP-MSI '09) held in conjunction with the 36th International Symposium on Computer Architecture*, *Proceedings of the 37th Annual International Symposium on Computer Architecture (ISCA)*, *Proceedings of the Workshop on the Interaction between Nanophotonic Devices and Systems (WINDS 2010)*, and *Photonics West: Proceedings of SPIE Optoelectronic Integrated Circuits XIV*), three journals (*IEEE Photonics Technology Letters*, *Optics Express*, and *Microelectronics Journal*) and has been submitted to the *IEEE Transactions on Very Large*

Scale Integration (VLSI) Systems.

Chapter 10: I. Savidis contributed to the development of the electrical model and material properties for the vertical cavity surface emitting lasers used to determine the thermal characteristics of bundled VCSELs on the thermal profile of the die. The development and evaluation of the research were performed in collaboration with co-author E. G. Friedman. Simulations were completed in COMSOL multiphysics by J. Wang, who also contributed to the development of the electrical model, and is the primary author of a journal paper published in the *Microelectronics Journal*.

Chapter 11: I. Savidis is the primary author of this chapter investigating intra- and inter-plane thermal coupling in 3-D stacked ICs. The development and evaluation of the research were performed in collaboration with co-author E. G. Friedman. Professor K. Hirschman of the Rochester Institute of Technology provided equipment and guidance during the experimental testing of the circuit. The research effort was published in the *Proceedings of the Government Microcircuit Applications & Critical Technology Conference (GOMACTech)* and has been submitted to the *IEEE Transactions on Electron Devices*.

Chapters 12 and 13 are, respectively, the concluding and future work chapters. I. Savidis is the primary author of both of these chapters.

Graduate study was supported by a Dean's Fellowship from the University of Rochester and in part by grants from the National Science Foundation, New York State Office of Science, Technology and Academic Research to the Center for Advanced Technology in Electronic Imaging Systems, Intel Corporation, Freescale Semiconductor Corporation, Samsung Electronics, Cisco Systems, and Qualcomm Corporation.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The rapid progression of ever more complex technological systems has brought an era of ubiquitous computational devices that touch all aspects of our lives. These technologies are found in the simplest of devices that we now take for granted such as TV remote controls, microwave ovens, refrigerators, electronic toothbrushes, and many more, to complex supercomputers and next generation devices. Two of which, the smartphone and tablet, have come to pass only in the past decade and whose global impact is just beginning to take hold. The technological revolution of the past half century is destined to have significant long term consequences, some of which have already past, some that are beginning, and many more that are bound to come.

These remarkable devices can be traced back to two significant innovations of the past century; the development of the point contact bipolar junction transistor by John Bardeen, Walter Brattain, and William Shockley at Bell Laboratories in 1947 [1–4], and the introduction of the integrated circuit (IC) by Jack Kilby at Texas Instruments and Robert Noyce at Fairchild Seminconductor, respectively, in 1958 and 1959 [5–8]. A microphotograph

of both the ground breaking bipolar junction transistor and the pioneering integrated circuit are shown, respectively, in Figures 1.1 and 1.2. Bardeen, Brattain, and Shockley won the Nobel prize in physics in 1956 for sowing the seeds to what would soon become the explosive semiconductor era [3]. The invention of the transistor revolutionized the semiconductor industry as the transistor replaced the much larger and often times electrically volatile vacuum tube that was first introduced at the turn of the 20th century [9]. These two revolutionary breakthroughs mark the beginning of the semiconductor era, and initiated the microelectronics revolution of the past half century. One of the most remarkable aspects of this era is the near doubling of transistor count every one to two years. This phenomena of near doubling transistor count has even been given ”law” status, being referred to as Moore’s law after the source, Gordon E. Moore, published a paper describing this technological trend in the mid 1960’s [10–13].

From these early beginnings, the IC has experienced an exponential growth in the number of transistors. With this growth came increased system complexity, where systems incorporated greater functionality into the hardware design. The next natural evolution of the IC was to incorporate more functional variety, for example, analog and digital circuitry, which led to systems-on-chip (SoC) integrated circuits. Insight on the ever progressing integrated circuit, from the early years in Kilby’s and Noyce’s labs, to current state-of-the-art complex multiprocessors and systems-on-chip (SoC) that are designed and sold by a multitude of companies such as Intel, AMD, IBM, Qualcomm, Freescale, Samsung, and the like, is provided in Subsection 1.1.

Although few accolades are bestowed upon packaging technologies, without the proper

housing to both protect the IC and provide an ever increasing demand for bidirectional communication to the outside world, these rapid advancements in IC technology and the corresponding progression in consumer electronics are certain to have been limited in scope. Similar to the evolution of integrated circuit technology, packaging technologies have evolved to accommodate the increasing power and input/output (I/O) demand of the IC. The evolution in packaging can also be traced back to the 1960's, where ceramic flat packs were first used for military applications and NASA [14]. The progression in packaging technology evolved with demands on the IC, such as increased power density and number of input/output (I/O) signals. Packaging technology has evolved to the point that, in the past decade, three-dimensional (3-D) packaging has passed from a research concept to a production level process.

It is from the evolution of both IC and packaging technologies that three-dimensional integrated circuits have emerged as a natural progression for system level design. As integrated circuits became more complex, 3-D integration continued this trend towards massive systems level integration with gigascale complexity. 3-D integration provides enhanced inter-connectivity, greater device integration density, a reduction in the number and length of the long global wires, and the potential to combine disparate heterogeneous technologies [15–17].

1.1 The evolution of the integrated circuit

The story of the integrated circuit begins in the summer of 1958 at Texas Instruments (TI). Jack Kilby was a newly hired Texas Instruments employee tasked with building

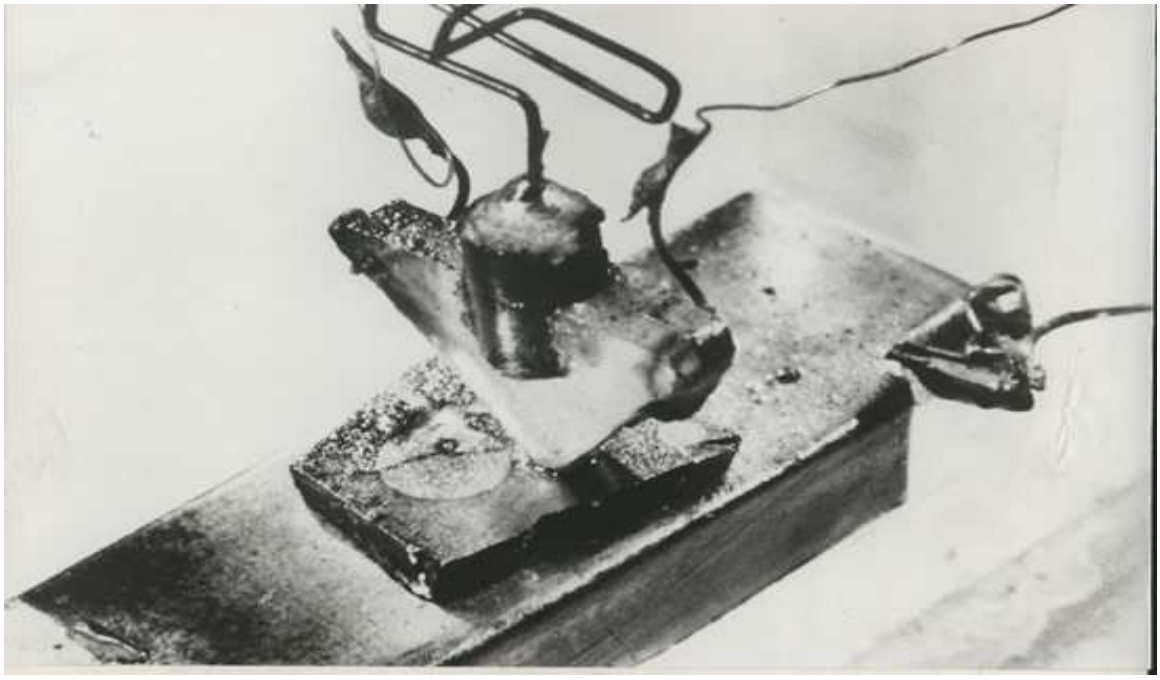


Figure 1.1: The first point contact transistor developed by Bardeen, Brattain, and Shockley at Bell Laboratories in 1947 [3, 4].

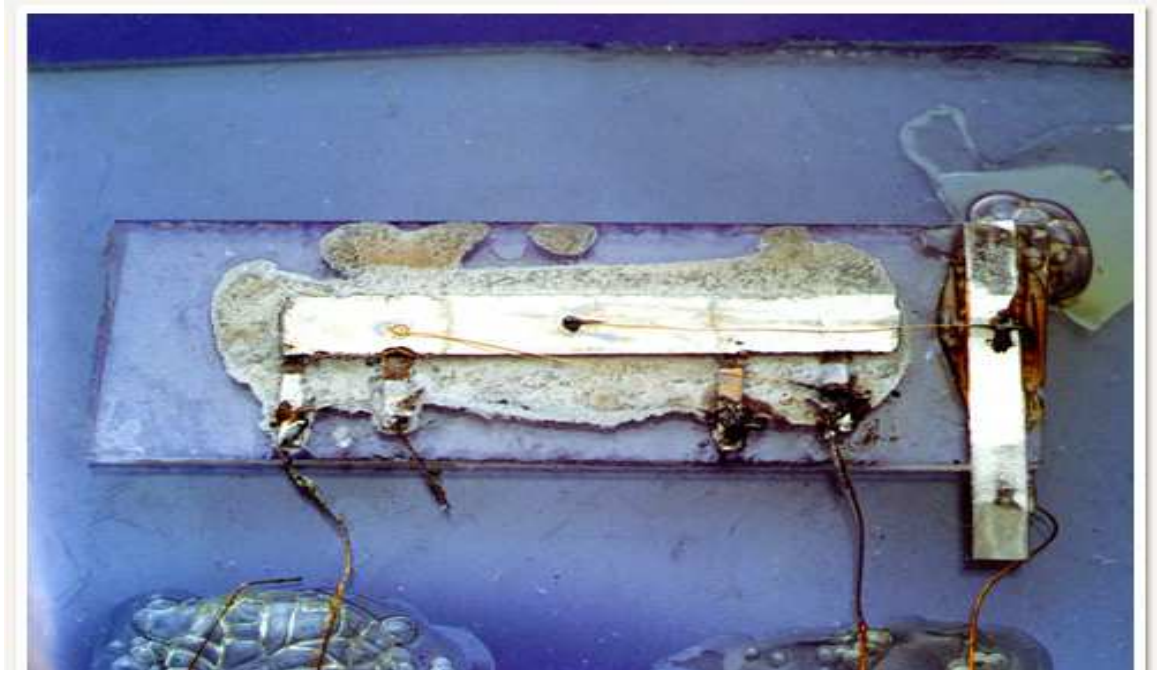


Figure 1.2: The first integrated circuit developed by Jack Kilby at Texas Instruments in 1958 [3].

smaller electrical circuits. Unhappy with TI's current miniaturization project, Kilby saw an opportunity to develop his own solution to the miniaturization problem. During that summer, Kilby worked alone in the lab and formulated a groundbreaking idea: he would produce all electrical components and the circuit out of the same piece of semiconductor material. This ingenuity produced the first monolithic integrated circuit. By fabricating all parts from the same piece of material and adding metal connections as a layer on top, Kilby removed the need for individual discrete components. Wires and components were no longer assembled manually, and circuits were made smaller with a manufacturing process that could now be effectively structured. After Kilby presented his new idea to his superiors in the late summer of 1958, he was permitted to build a test version of his circuit. By September 1958, Kilby had his first functional integrated circuit. The integrated circuit earned Jack Kilby the Nobel Prize in Physics in 2000.

Although the idea was groundbreaking, Jack Kilby's integrated circuit was pretty crude and had some problems. Robert Noyce at Fairchild Semiconductor solved several practical problems of Kilby's integrated circuit, specifically, interconnecting all of the components onto the same substrate. Noyce added metal as a final layer and etched a portion to form the wires needed to connect the individual components. The integrated circuit was now more suitable for mass production. Noyce's breakthrough came about six months after Kilby produced the first IC. Robert Noyce, with the knowledge acquired from his early pioneering work on the development of the IC, co-founded Intel, today one of the largest and most important integrated circuit manufacturers.

Early commercial development of the integrated circuit was difficult and costly. These

early ICs were only sought after by the military (for the US Air Force Minuteman Program) and NASA (for the Apollo Guidance Computer) as the benefits of miniaturization justified paying the high costs [18–20]. When these IC's were first introduced, functionality was limited to the most simplistic of tasks, and integrated a few resistors, capacitors, and transistors to form basic amplifiers and filters. The 1960's marked the beginning of an era from small scale integration (SSI), to medium and large scale integration (MSI and LSI). Fairchild and TI had an early lead in the integrated circuit business, producing the first commercial devices. Microphotographs of these early ICs are provided in Figure 1.3 [3]. Different terms were coined during these early years to describe the different circuit technologies. Names such as direct coupled transistor logic (DCTL), resistor transistor logic (RTL), and diode transistor logic (DTL) emerged from the various semiconductor companies, each claiming a technological advantage over the other. By the mid 1960's, Motorola also entered as a competitive IC manufacturer, and by the end of the 1960's, a wide variety of semiconductor companies had emerged, with one of the most significant being Intel. Intel was founded in Mountain View, California on July 18, 1968 by Robert Noyce and Gordon Moore, primarily as a dynamic random access memory (DRAM) company, and would soon become the leading semiconductor company in the world. A brief history of early Intel based microprocessors is presented below to provide an indication of the technological drivers semiconductor companies were pursuing. Although the historical data is for Intel based microprocessor devices, this information is indicative of industry-wide trends in the high end semiconductor business that also includes Intel rivals, IBM and AMD.

Beginning with the early 1970's, Intel switched from making DRAM circuits to producing high performance microprocessors [21]. One of the first marketed single chip microprocessors was the Intel 4004. A microphotograph of the 4004 is shown in Figure 1.3d [22–25]. The 4004 was a simple 4-bit PMOS processor, with a multiplexed address and data bus due to limited I/O pins. The 4004 was housed in a 16-pin ceramic dual inline package (DIP). Although the Intel 4004 worked fine for calculators and similar SSI applications, the circuit was not suited for microcomputers due to limitations of the architecture. The 4004 had only a three-level deep stack, lacked interrupt support, and used a complicated method to access the RAM. Some of these shortcomings were fixed in the Intel 4040 with the introduction of a seven-level deep stack and a simplified RAM access protocol, but the circuit still did not include interrupt support. Additional information including the power consumption, operating frequency, processor area, and other relevant parameters are listed in Table 1.1 [22].

The next major family of Intel processors was the 8000 series. The Intel 8008 was released five months after the Intel 4004, and was the first 8-bit microprocessor. The 8008 contained 3,500 transistors and was fabricated in a 10 μm enhancement load PMOS process. A microphotograph of the 8008 is shown in Figure 1.3e [24, 26, 27]. Two years later, more specifically, in 1974, Intel released the 8080, a 6,000 transistor processer fabricated in a 6 μm enhancement load NMOS process. There were significant improvements to the 8080 processor as compared to the 8008 [26, 28]. The maximum memory size on the 8080 was increased from 16 KB to 64 KB, the number of I/O ports was increased to 256, and many new instructions and direct addressing modes were added. The 8080 included a Stack

Pointer (SP) register to specify the position of the external stack in CPU memory. This register permitted the stack to grow as large as the size of the memory, not limiting the 8080 to a seven-level internal stack like the 8008. The third CPU in the 8000-family of processors was the 8085. Process enhancements included a reduction of the transistor channel length to 3 μm, which increased the clock speed to a frequency of 8 MHz. In addition, the 8085 included a single 5 volt power supply (rather than the +5, -5, and +12 volt supplies in the previous generations), a serial I/O port, and the clock oscillator and system controller were integrated on-chip. By the late 1970's, Intel introduced the 8086, the first member of the x86 family of processors. The 8086 had 29,000 transistors, fabricated in a 1.5 μm process, and boasted a core speed of 10 MHz [29, 30]. Intel produced this first 16-bit processor by dividing the memory into 8-bit odd and even banks that were simultaneously read to produce a 16-bit data stream every clock cycle. The 8086 also incorporated a set of powerful string instructions that supported block level operations, operations on the smallest unit of data permitted in memory, such as moving a block of data, comparing data blocks, and setting data blocks to certain values, amongst many functions.

The next family of Intel microprocessors were based on the x86 architecture and included the 80186, 80286, 80386, and 80486. These processors were developed in the 1980's. Both the 80186 and 80286 were released in 1982 and are 16-bit processors. The 80286 was a 134,000 transistor processor that operated up to 25 MHz, had a 16 MB addressable memory, and was fabricated in a 1.5 μm process [31]. The 80286 processor was also the first to include a protected mode, where a 16 MB block of memory was addressable by the CPU, protecting this memory block and other system resources from user

programs. A microphotograph of the 80386 is shown in Figure 1.3f [24, 32, 33]. The 80386 and 80486 were the first 32-bit Intel processors and were released, respectively, in 1985 and 1989 [32, 34]. Numerous improvements to the microarchitecture of the 80486 had been incorporated. The 486 included a 8 KB unified level 1 cache that was increased to 16 KB, much faster bus transfers (1 CPU cycle as opposed to two or more cycles for the 80386), clock doubling and tripling technology, reduced execution time of instructions, an integrated floating-point unit (prior processors included separate CPU and FPU ICs that communicated through an interchip bus), and power and system management features. Intel was simultaneously producing integrated circuits with 1.2 million transistors with a feature size of 1 μm. The remaining families of Intel based microprocessors are all x86 based except for one, the XScale 32-bit RISC microprocessor based on the ARM architecture introduced in 2000. These families include the Pentium based processors (Pentium I, II, III, IV, Pro) developed during the 1990's [35–41], the core and core2 processors of the early and mid 2000's, and the current Nehalem microarchitecture (Intel i3, i5, and i7) based processors [42, 43]. A microphotograph of a Pentium II microprocessor is shown in Figure 1.3g [24, 37, 44]. Information on the operating frequency, power consumption, number of cores, processor area, and other relevant parameters for each processor is also listed in Table 1.1.

The critical point of this discussion is that cutting edge microprocessors evolved over time. A common thread throughout the evolution of the integrated circuit is that the feature size of the transistor had been reduced, increasing the number and type of functions integrated on-chip. This process has continued since the late 1960's, and continues to this

day. The design objectives may have changed over time, *i.e.*, higher speed for each successive generation of Pentium based processors, or more recently, portability with cores that consume far less power but at lower speeds. Another trend that has occurred in the high end microprocessor market is the switch to multiple core based systems, exemplified by both the Core2 and Nehalem processors, where the Intel i7 processor (Nehalem based) is a six core device [42, 43]. Intel is not the only semiconductor company pursuing multi-core systems. An image of an 18 core IBM Blue Gene Q integrated circuit is shown in Figure 1.3h [45–47]. This trend towards incorporating a larger number of individual cores appears to be continuing. As a final note, over the past decade, it is not just the increasing number of cores, but the type and number of functions that have been included on-chip that have also drastically increased, opening the door to intellectual property (IP) based integration, often described as systems-on-chip (SoC) integration. Not only did the number and type of functions increase, but the vast variety of functions incorporated on present day SoC based integrated circuits is astonishing. It is this concept of systems level integration that 3-D integrated circuits can support, which appears to be the next natural evolution of the integrated circuit [48–50].

1.2 Packaging technology evolves too

From the beginning of the 1960's, when Fairchild and TI were producing the early ICs, packaging technologies evolved to accommodate the increasing power and input/output (I/O) demands of each successive IC generation. There are two primary aspects to this packaging evolution, the first being a physical transition from wirebonding both power and I/O pads onto the IC to the pads of the package (see Figure 1.4a [51, 52]), to thermosonic

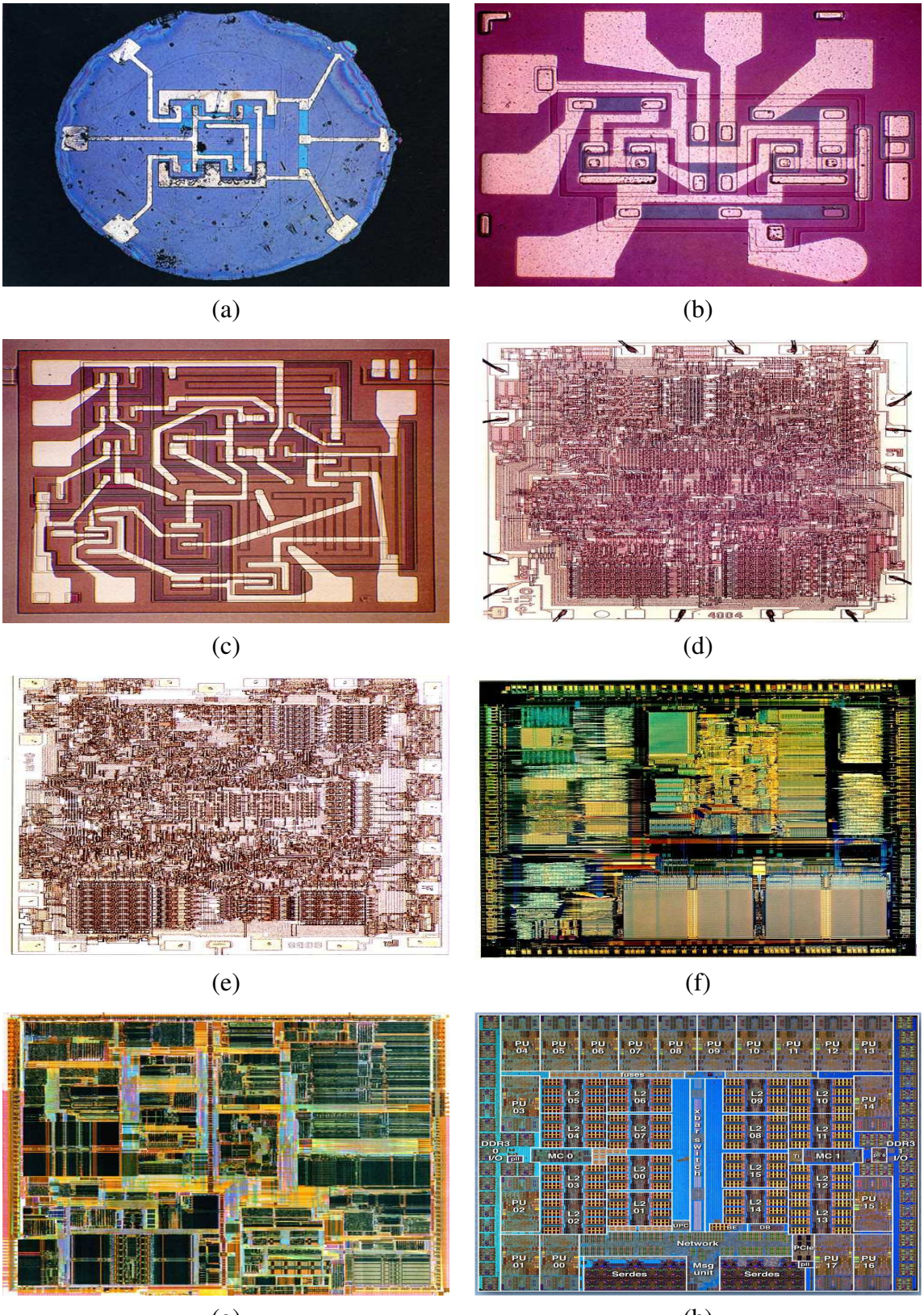
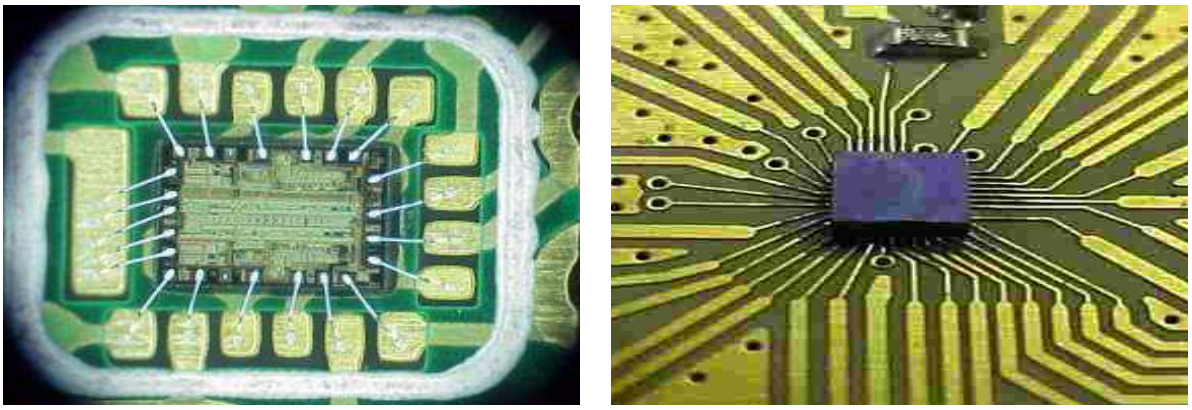


Figure 1.3: The evolution of the integrated circuit from early Fairchild integrated circuits (a)-(c), to single core Intel processors such as the (d) 4004, (e) 8008, (f) 80386, and (g) Pentium II, and finally to multicore processors such as the (h) IBM 18 core Blue Gene Q IC.

Table 1.1: Intel microprocessor evolution [22, 26, 28–32, 34–43]

Processor	Year	Transistors	Process technology (μm)	Process	Power consumed (watts)	Supply voltage (Volts)	Package	Bus width (bits)	Clock frequency (MHz)	Number of cores
4004	1971	2300	10	PMOS	0.6	15	16-pin plastic DIP	4	0.108	1
4040	1974	3000	10	PMOS	0.9	15	24-pin ceramic DIP	4	0.74	1
8008	1972	3500	10	PMOS	0.8	5, −9	18-pin ceramic DIP	8	0.2, 0.5, 0.8	1
8080	1974	6000	6	NMOS	1.5	±5, 12	40-pin ceramic DIP	8	2, 3.125	1
8085	1976	6500	3	HMOS	1	5	40-pin DIP, 44-pin PLCC	8	3, 6	1
8086	1978	29,000	3	HMOS	1.7	5	40-pin ceramic DIP	16	4.77, 8, 10	1
8088	1979	29,000	3	HMOS	1.7	5	40-pin ceramic DIP	8	4.77, 8	1
80186	1982	29,000	2	CMOS	3	5	68-pin ceramic PGA, LCC, PLCC, TQFP	16	6-20	1
80286	1982	134,000	1.5	CMOS	3.3	5	68-pin ceramic PGA, CLCC, PLCC	16	6, 10, 12.5	1
i386 DX	1985	275,000	1, 1.5	CMOS	1.5-2.8	5	132-pin ceramic PGA	32	16, 20, 25, 33	1
i486 DX	1989	1.2 million	0.8, 1	CMOS	3.5	5	168-pin ceramic PGA	32	25, 33, 50	1
Pentium P5	1993	3.1 million	0.25-0.8	CMOS	15	5	273-pin ceramic PGA	32	60-300	1
Pentium Pro P6	1995	5.5 million	0.25-0.6	CMOS	40	3.3	387-pin ceramic SPGA	32	166, 180, 200	1
Pentium II	1997	7.5 million	0.25	CMOS	32	1.6, 3.3	242-pin mini-cartridge	32	233-450	1
Pentium III Katmai	1999	9.5 million	0.25	CMOS	24	2.0, 3.3	242-pin Slot 1 SECC2, 370-pin FC-PGA	32	450-600	1
Pentium 4 Willamete	2000	42 million	0.18	CMOS	90	1.75	478-pin FC-PGA	32	1400-1600	1
Pentium M Banias	2003	77 million	0.13	CMOS	22-24.5	0.96-1.48	478-pin FC-PGA	32	1200-1500	1
Pentium M Dothan	2004	140 million	0.09	CMOS	21-27	0.98-1.34	479-pin FC-BGA, FC-PGA	32	1500-2200	1
Pentium 4 Prescott	2004	125 million	0.09	CMOS	82-115	1.25-1.4	478-pin FC-PGA	32	2400-3800	1
Core 2 Duo Conroe	2006	291 million	0.065	CMOS	82-115	0.85-1.5	775-land FC-LGA	64	2400-3000	2
Core 2 Duo Wolfdale	2008	410 million	0.045	CMOS	65 (100)	0.85-1.36	775-land FC-LGA	64	2600-3300	2
Core 2 Quad Yorkfield	2008	820 million	0.045	CMOS	95 (132)	0.85-1.36	775-land FC-LGA	64	2800-3200	4
Core i7-975 Bloomfield	2009	731 million	0.045 (HKMG)	CMOS	160 (230)	0.8-1.3	1366-land FC-LGA	64	3000-3700	4
Core i3-560 Clarkdale	2010	559 million (CPU/GPU)	0.032	CMOS	129 (170)	0.65-1.4	1156-land FC-LGA	64	2900-3330	2
Core i5-760 Lynnfield	2010	774 million	0.045	CMOS	160 (190)	0.65-1.4	1156-land FC-LGA	64	2900-3500	4
Core i7-980 Gulftown	2011	1170 million	0.032 (HKMG)	CMOS	160 (230)	0.8-1.3	1366-land FC-LGA	64	3000-3700	6



(a) (b)

Figure 1.4: Evolution of bonding technology from (a) wirebonds [52] to (b) flip-chip bonding [56].

flip-chip bonding (see Figure 1.4b [53–56]) these same pads. This evolution increased the number of pads available to the IC about ten fold, from approximately 800 to 8,000, and limited the parasitic inductance of the long wirebonds at the expense of increased packaging complexity. The increase in the complexity of the package is the second aspect of the evolution in package technology. It was no longer appropriate to develop IC packages without performing electromagnetic (EM) simulations of the components and interconnects required for both power and signaling within the IC package. In addition to electromagnetic simulation, electronic design automation (EDA) tools were developed by major vendors such as Mentor, Cadence, and Synopsys to address this increase in board and package level complexity.

The board level package contains many electrical components that include ICs, resistors, diodes, and capacitors. These components are interconnected at the board level. The IC package provides mechanical support and environmental protection and connection to the board. There are four primary functions that the IC package provides: 1) heat removal,

2) power distribution, 3) signal distribution, and 4) physical and electrical protection. The following paragraphs describe the evolution of IC packaging technology in support of advancements in integrated circuit technology, as the IC revolution produced significant improvements in the four primary functions of the package.

The vast majority of integrated circuits are packaged in opaque ceramic or plastic insulation. The only connection to the outside world is through metal pins (traces) through the insulation. Traces originating at the die, through the package, and into the printed circuit board have different electrical properties as compared to the on-chip impedances. Circuit board traces require special design techniques and need greater electrical power than signals confined to the IC. Integrated circuits produced in the 1960's and 1970's were packaged in ceramic flat packs. These packaged ICs were reliable and small, and were used by the military for many years. Commercial circuits quickly moved to the dual in-line package (DIP), first in ceramic and later in plastic. By the 1980's, pin counts exceeded the practical limit for DIP packaging, leading to pin grid array (PGA), quad flat pack (QFP), and leadless chip carrier (LCC) packages. Surface mounted packages appeared in the early 1980's and became popular by the end of the 1980's. These packages used a finer lead pitch with leads formed as either a gull-wing or J-lead [57–59]. The small outline integrated circuit (SOIC) is an example of a surface mounted IC. The package for the small outline IC occupied an area about 30% to 50% less than an equivalent DIP, and was typically 70% thinner.

After the DIP, two new families of packages emerged: the previously discussed SOIC and plastic leaded chip carrier (PLCC) packages. By the late 1990's, the plastic quad flat

pack (PQFP) and the thin small outline packages (TSOP) were most prevalent for high pin count devices, although PGA packages are still used today for high end microprocessors. Intel and AMD have been shifting from PGA based packages for high end microprocessors to land grid array (LGA) packages. The evolution of packaging technology is shown in Figure 1.5 as it has progressed from the DIP of the 1970's to present 3-D packaging technologies [60].

The ball grid array (BGA) evolved from the PGA as the number of I/O per die increased. BGA packages have existed since the 1970's, but the BGA truly evolved in the 1990's with the development of the flip-chip ball grid array package (FCBGA), which supports a much higher pin count than other types of packages. Early BGA based 3-D integrated systems exploited the enhanced electrical characteristics of the solder balls as compared to wire bonded 3-D ICs. This capability was particularly useful for delivering power in these early 3-D systems as the solder balls lessen the IR and $L \cdot di/dt$ drops in the noise sensitive devices within the 3-D stack. These techniques integrate greater functionality onto a single package and reduce the power consumed in signaling between dies, as printed circuit board (PCB) connections are replaced by the better electrical characteristics of wire bonds or solder balls. In an FCBGA package, the die is flipped and mounted upside down, and connects to the package through (controlled collapse chip connection) C4 solder balls. This approach eliminates the long wirebonded interconnect necessary with other package types, thereby reducing the overall impedance of each I/O. The FCBGA package also supports an array of input-output signals distributed over the entire die rather than confined to the periphery of the die.

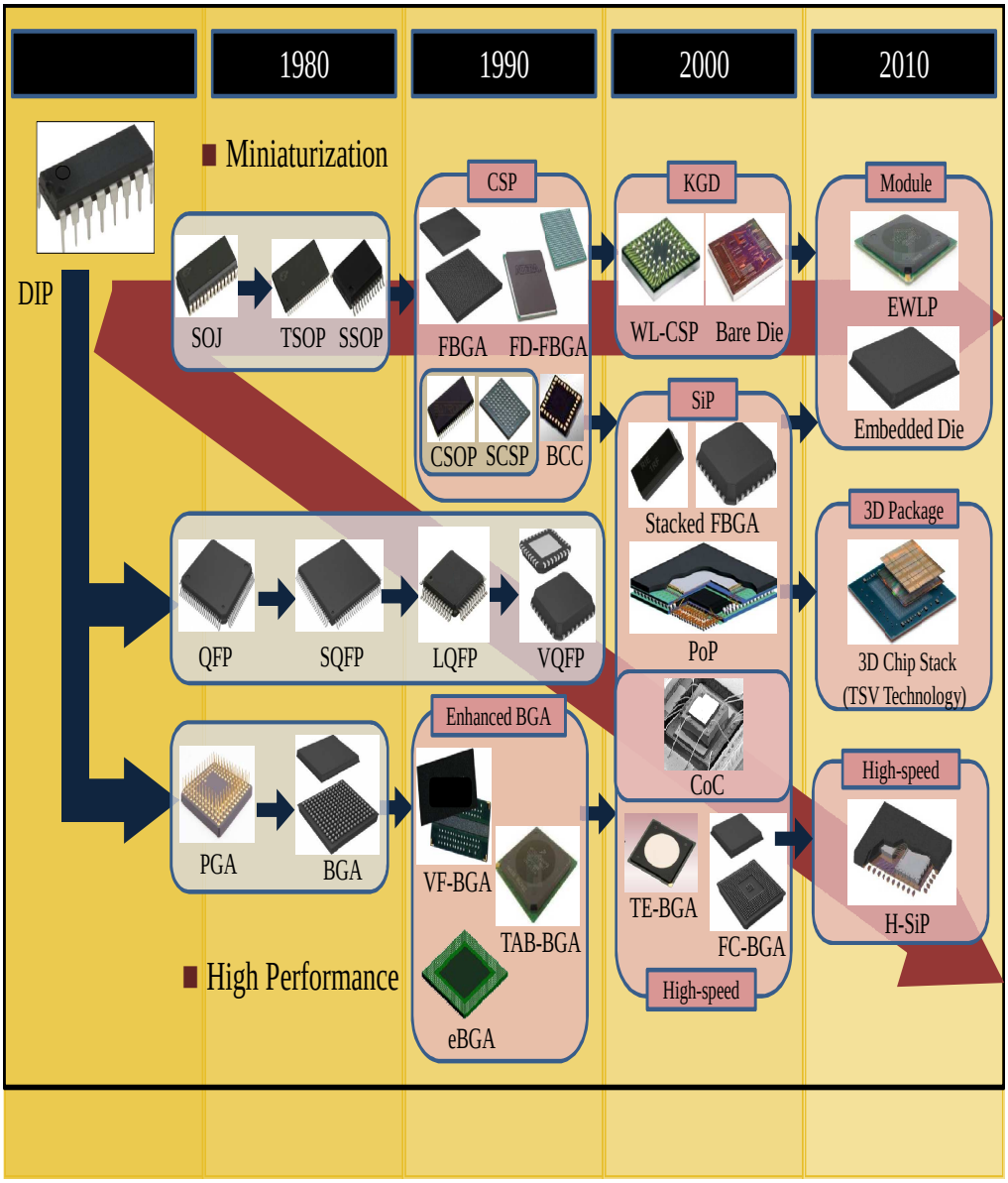


Figure 1.5: The evolution of packaging technology [60].

There are also IC packages that are not as common, but necessary for more exotic applications such as optical ICs. ICs that require proximity communication are just one example of a circuit that requires a non-standard package. Another exotic package is the chip on board (COB) package that attaches the raw die directly to a PCB, the die wires are bonded directly to the PCB traces, and the die and die wires are covered with an insulator that is partially or completely transparent to permit optical input and output of optoelectronic devices and electrically programmable read-only roms (EPROM). There are many more varieties of exotic ICs requiring exotic packaging, similar to the proximity communication IC [61–66].

The evolution of these different packaging technologies is shown in Figure 1.6. Beginning with the 1970's, different package technologies emerged as ICs requirements shifted. The two primary technology drivers for packages, shown on the axes of Figure 1.6, are the number of I/Os corresponding to the increase in circuit performance, and the need for small size as decreased form factor relates to lower cost [67]. Each package technology that emerged balanced these two issues differently. Over the past decade, there has been a shift in package technology, with a greater focus on systems-level integration.

Packaging technologies went through another evolutionary period in the past decade to accommodate greater system complexity. Multiple dies are now stacked in one package, producing a system-in-package (SiP). Another variation stacks multiple packages to produce a system-on-package (SoP) or package-on-package (PoP). A microphotograph of a PoP is shown in Figure 1.7 [68]. These are early forms of three-dimensional integrated circuits, and these configurations fall into a larger category that combines multiple dies

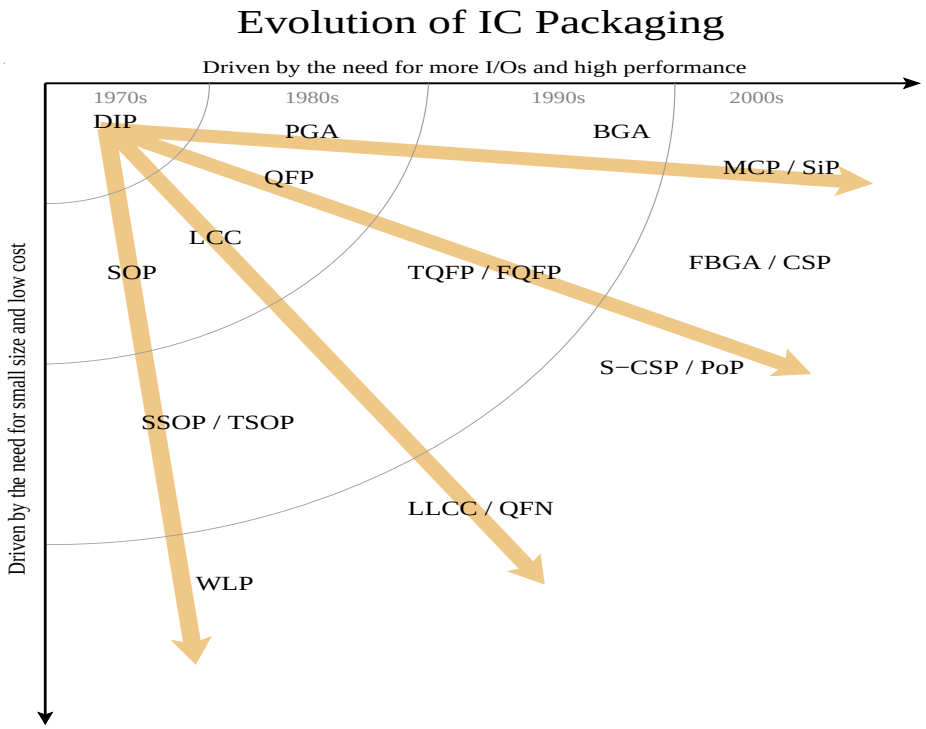


Figure 1.6: Evolution of the package from the 1970's to today [67]. Two different driving factors are compared: 1) the need for small size and low cost on the vertical axis, and 2) the need for more I/Os and high performance on the horizontal axis.

onto a single substrate. This category is referred to as a multi-chip module (MCM), as compared to a single chip module (SCP). MCMs combine two-dimensional as well as three-dimensional integrated circuits, and therefore the boundary between a large MCM and a small printed circuit board has become somewhat unclear. Another form of package that advances three-dimensional integration is wafer level packaging (WLP). With WLP, the integrated circuit is packaged at the wafer level, and the size of the resulting package is similar to the size of an IC. Wafer level packaging streamlines the manufacturing process by enabling wafer fabrication, packaging, test, and burn-in at the wafer level. Rather than initially dicing the wafer, and individually packaging each die in a plastic package, wafer level packaging attaches the top and bottom outer packaging layers and solder bumps onto the integrated circuit while still in the wafer. Wafer dicing is performed once packaging is completed. Wafer-on-wafer based 3-D integration is a natural next step, with the introduction of large vias through the silicon substrate to interconnect the different die. This step leads to a critical realization: Stacks of packaged die interconnected using wire-bonding, flip-chip bumps, ball grid arrays, wafer level packaging, and other techniques have already been used to create products, with through silicon via (TSV) based three-dimensional integration as the next potential evolutionary step.

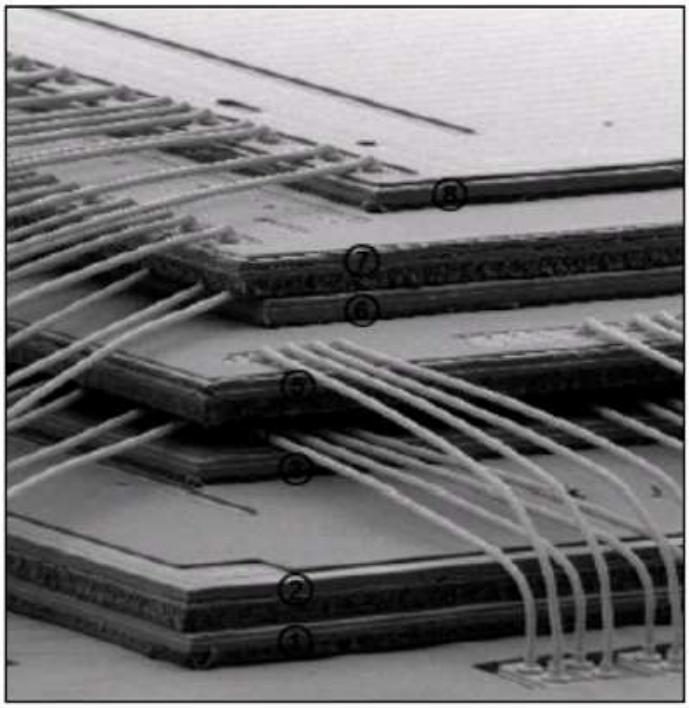


Figure 1.7: A microphotograph of a package-on-package three-dimensional integrated circuit [68].

1.3 3-D integration: Merging IC and packaging technologies

The era of rapid technological scaling has brought revolutionary advancements in systems level integration. Integrated circuits that once performed simple computations have evolved to advanced ubiquitous microprocessors. The natural evolution of the microprocessor to incorporate both analog and digital circuitry has led to system-on-chip (SoC) integrated technologies. As integrated circuits became more complex, three-dimensional integration has emerged as a new technology platform that continues this trend towards massive systems level integration.

Three-dimensional integration is a novel technology of growing importance that has the potential to offer significant performance and functional benefits as compared to 2-D ICs. 3-D integration provides enhanced inter-connectivity, high device integration density, a reduction in the number and length of the long global wires, and the potential to combine disparate heterogeneous technologies [15–17]. Three-dimensional integration has emerged as a class of solution that combines technology scaling with packaging techniques. Successful fabrication of vertically integrated transistors dates back to the early 1980's [69]. The structures included the so-called JMOS structures, a term that describes the joint use of a single gate for both devices [70]. One such example is a monolithic 3-D CMOS inverter where the PMOS and NMOS transistors share the same gate, reducing the total area of an inverter [69, 71, 72]. Research on three-dimensional integration remained an area of limited scientific interest during the 1980's and 1990's. By the early 2000's, this situation

began to change due to the increasing importance of the interconnect and the demand for greater functionality on a single substrate. Over the past ten years, three-dimensional integration has evolved into a design paradigm that encompasses multiple abstraction levels, including the package, die, and wafer. Different manufacturing processes and interconnect schemes have been proposed for each of these abstraction levels [16]. The central features of three-dimensional systems are briefly reviewed in the following subsections.

1.3.1 Three-dimensional integration as an enabling technology

3-D integrated circuit (3-D IC) technology is the next generation in chip stacking processes. Bare die are vertically stacked and interconnected using fine pitched through substrate vias. 3-D IC technology is currently the most practical solution to interconnect different dies requiring tens of thousands of interconnects with distances of less than 100 μm. Although still an expensive process, the advantages of integration density results in the smallest volume and form factor, the highest performance, and the greatest interconnect density among dies within a packaged stack.

A significant benefit of three-dimensional integration is the decrease in the length of the longest interconnects across an integrated circuit. Similar trends caused by the reduction in wire length are described by various interconnect prediction models adapted for 3-D ICs [73]. The decrease in length is a promising solution for increasing circuit speed while reducing the power dissipated by an integrated circuit.

One of the most critical attributes of 3-D ICs is the ability of these systems to integrate disparate technologies. This defining feature of 3-D ICs offers unique opportunities for

highly heterogeneous and multi-functional systems. Potential examples of heterogeneous 3-D systems-on-chip abound. One example is a real-time image processing system where the image sensors capture the light on the topmost plane, analog circuits on the plane below convert the signal to digital data, and the remaining two planes of digital logic process the information from the upper planes [74, 75]. A 3-D SoC version features considerably improved performance as compared to a planar implementation of the same system [76, 77]. A schematic illustration of another example with sensors and an antenna on the topmost plane and additional planes for RF, analog, digital, and communications circuitry below is shown in Figure 1.8 [17]. Applications in the military, medical, and wireless communication domains as well as low cost consumer products exist for vertically integrated systems. These benefits of 3-D integrated systems are discussed in the next sub-section.

1.3.2 Benefits of 3-D integration

There are substantial benefits to stacking ICs and vertically interconnecting these planes. The capability to vertically bond multiple dies or wafers to manufacture an interconnected stacked IC is, however, not mature, with much research currently in this area. The following few subsections focus on the benefits of 3-D integrated circuits with regard to heterogeneous systems integration, form factor, power reduction, and cost.

a. High degree of integration

A strong need to miniaturize consumer electronics is continually underway. The portability and form factor of electronic devices such as the smartphone and tablet have become

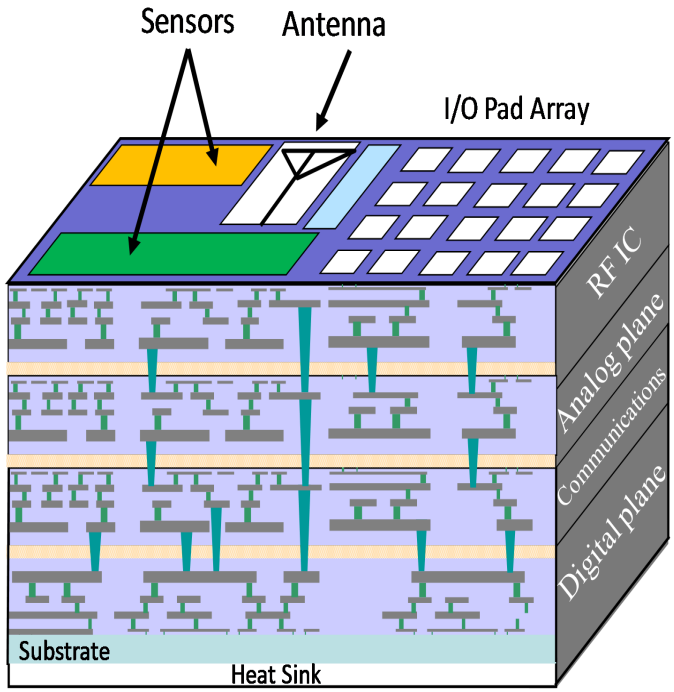


Figure 1.8: A 3-D integrated system with sensors and antennas on the topmost device plane, and RF, analog, digital, and communication circuitry below [17].

primary product criteria for consumers. In addition, consumers expect increasing functionality from their portable devices. These requirements contribute to economic pressure on integrated circuit manufactures to produce ever smaller integrated circuits in packages containing multiple ICs. 3-D integration provides a means for these vendors to produce highly integrated systems. By stacking thinned chips, 3-D IC technology densely packs multiple dies into a small footprint, providing enhanced functionality in a thin package [17, 78–80].

b. Heterogeneous integration

Modern integrated circuits include a number of heterogeneous functions. These functions can include data processing, sensing, memory, data transmission, and many other capabilities requiring multiple fabrication processes. Incorporating these functions into a single die requires different fabrication processes to be integrated, a complex and costly process. Three-dimensional integration offers the potential to merge once incompatible fabrication processes into a stacked 3-D system. 3-D manufacturers can place gallium arsenide devices above silicon transistors. Dies from different process technologies (*e.g.*, memory, CMOS, RF, analog, MEMS/NEMS), foundry lines from completely different vendors, or even emerging technologies such as integrated optics or graphene can be stacked to form a 3-D system. In addition, TSV-based 3-D integration offers additional benefits to the alternative, board-level and system-in-package 3-D processes, as TSV-based 3-D integration provides both an increase in interconnection density, and a reduction in the package footprint.

c. Reduced form factor

Squeezing more functionality out of smaller integrated circuits not only reduces cost, but paves the way for increased and novel mobile consumer electronics, where the size of the dies and boards is critical. Aside from a smaller die area through standard scaling principles, as described by Moore and Dennard [10–12, 81], 3-D integration provides an alternate approach to reducing the form factor. Where scaling theory proposes a reduction in area by reducing the transistor channel length (and other physical and material characteristics), 3-D integration permits the decomposition of an integrated circuit into multiple dies, each a subportion of the original circuit, which are vertically stacked. The limitation in reducing footprint area has become an issue of circuit block placement, thermal effects, and manufacturing constraints. Despite these limitations, the ability to reduce the IC form factor with 3-D integration paves the way for integrating systems into smaller packages, ideal for mobile consumer and military devices.

d. Enhanced power consumption

Power consumption remains an important design criterion. Integrated circuits for portable devices in consumer, military, and other applications require efficient use of power to maximize battery lifetime. The power consumed by an integrated circuit, in many applications and devices, has become more critical than either speed or area. Without effective power management techniques and strategies, on-chip temperatures pose a risk to circuit integrity. Manufactures are often required to install cooling devices such as fans to avoid catastrophic failure due to increased on-die temperatures at the expense of greater power

consumption. One of the primary sources of on-chip power consumption is the signal buffers that drive long on-chip wires. Fast switching signals that travel across a die over long distances require low power circuits. 3-D integration provides an additional solution to minimizing circuit level power consumption by reducing the number and length of the global signal interconnect. By partitioning functional blocks into multiple, vertically stacked die rather than placing these blocks on opposite sides of a planar 2-D die, the distance between these blocks is greatly decreased. Increased locality reduces interconnect length and power consumption.

e. Cost reduction

Since the beginning of the semiconductor revolution, increases in systems integration have been directly linked to the ability of semiconductor companies to scale the size of the active devices to smaller dimensions. Transistor scaling has become increasingly difficult and, more importantly, expensive [78]. The development of process technologies, the construction of fabrication facilities, and the equipment necessary to produce integrated circuits have become prohibitively expensive to the point which only a few of the most successful companies can afford these large capital expenditures [78]. In addition, die yield, a critical metric semiconductor companies use to quantify the per cent of functional die, degrades with each successive transistor generation as cross-die and cross-wafer process variations fundamentally affect these smaller devices. Another cause of reduced yield is the increased physical die size to accommodate greater functionality. Three-dimensional

integration is a potential solution to the greater expense in systems integration and reductions in yield. With 3-D circuits, the die size does not have to be increased to provide greater functionality, but rather can be reduced into smaller dies which are vertically stacked. By reducing the size of a 2-D die, the number of circuit defects per die is reduced. 3-D integration also supports the use of multiple technologies, allowing circuits to incorporate older and cheaper technologies for the non-critical components within a system.

In this research proposal, emerging 3-D technologies and design methodologies are described, and solutions for specific critical issues are proposed. An outline of the research proposal is presented in the next section.

1.4 Thesis outline

The primary focus of this research effort is the development of circuit-level techniques in support of integrating multiple dies or wafers into a vertically stacked three-dimensional system. Although there are several techniques to achieve vertically stacked systems integration, this research dissertation primarily focuses on through silicon via (TSV) and through oxide via (TOV) based technologies. As both TSVs and TOVs only differ with regard to the substrate material, silicon versus oxide, the two terms are used interchangeably throughout this dissertation.

The following three chapters provide background information on general design issues in 3-D integrated systems. An introduction to electrical characterization of the TSV is discussed in Chapter 2. An overview of power generation and distribution for stacked dies is offered in Chapter 3. Background material on thermal effects in 3-D integrated circuits

is provided in Chapter 4.

Background material on the electrical models and characterization of TSVs is presented in Chapter 2. Electrical characterization of single, multiple, and bundled TSVs is described. This work includes models based on full wave simulations, as well as closed-form expressions for the TSV resistance, inductance, capacitance, and substrate conductance. Experimental work on the resistance, inductance, and capacitance of the TSVs is also presented. The primary objectives of this chapter are two fold: 1) to describe an equivalent electrical model of a TSV for inclusion in SPICE based simulations that consider the physical parameters of a TSV, and 2) to reduce the computational time required to electrically characterize a TSV both as a single via as well as a group (or bundle) of TSVs. These models and expressions for the 3-D via impedance provide an efficient method to characterize the performance of signal paths containing through silicon vias.

A critical challenge in 3-D integrated circuits is the design of robust power distribution networks that provide sufficient current to every transistor within a vertically stacked system. In planar ICs where flip-chip packaging is adopted as the packaging technique, an array of power and ground pads is allocated throughout the surface of the integrated circuit. Increasing current densities and faster current transients, however, complicate the power delivery design process. Three-dimensional integration provides additional metal layers to distribute power that are unavailable in two-dimensional circuits. 3-D integrated systems however are immature, and much work is required to develop efficient power distribution topologies. Power delivery in 3-D integrated systems presents new challenges for delivering sufficient current to each of the device planes. Stacking device planes in the

vertical direction also leads to higher power densities. The effect of the increased power density within a 3-D power network is significant, as specialized design techniques are required to ensure that each device plane is operational while not exceeding the target output impedance. In addition, interplane power network noise suppression techniques, such as multi-plane decoupling capacitance, are required. Background material on 3-D power generation and delivery in 3-D systems is provided in Chapter 3.

Highly active circuit blocks produce localized hotspots that are detrimental to the operation of an IC. In 2-D ICs, heat removal is achieved primarily through passive techniques such as the proper placement of a heat sink. Direct thermally conductive paths exist from the active circuits, through the on-chip metalization, the package, and the heat sink. Two significant physical problems, with regard to heat generation and management, arise when considering 3-D ICs. First, since 3-D ICs permit a significantly greater packing density, the heat per unit area is much higher as compared to a 2-D IC. Second, as there are multiple stacked device planes, the distance to the heat sink differs depending upon the location of the die within the 3-D stack. These different distances imply that each device plane experiences a different thermal resistance from that plane to the heat sink. As thermal issues are of significant importance in 3-D ICs, background material on heat generation and propagation, modeling, and thermal management is provided in Chapter 4.

The next seven chapters, Chapters 5 through 11, describe work on different circuit-level aspects of 3-D integrated systems. TSV models and closed-form expressions developed for the resistance, inductance, and capacitance of both a single TSV as well as multiple TSVs

are described in Chapter 5. Models of three different clock distribution networks, comparing the root-to-leaf delay to experimental results, are provided in Chapter 6. The next chapter, Chapter 7, focuses on power delivery in 3-D stacked dies, where the propagation of noise within the power and ground networks is experimentally compared among three different power distribution topologies. The design of a test circuit examining different decoupling capacitor placement schemes for 3-D integrated circuits is discussed in Chapter 8. The systems level integration of a hybrid 3-D integrated free-space optical system for multi-core intrachip communication is described in Chapter 9. Research on characterizing thermal coupling between single and bundled vertical cavity surface emitting lasers (VCSELs) in a 3-D stack with digital CMOS, RF, and analog device planes is described in Chapter 10. Finally, inter- and intra-plane thermal coupling within a stacked 3-D IC is discussed in Chapter11. A more detailed breakdown of each chapter is provided below.

The chapter on the electrical properties of a through silicon via is composed of two distinct sections. The first section describes the electrical characterization of a TSV. Both capacitive and inductive coupling between multiple 3-D vias is described as a function of the separation distance and plane location. The effects of placing a shield via between two signal vias is investigated as a means to limit capacitive coupling. The location of the return path is examined to determine a preferred placement of the 3-D vias to reduce the overall loop inductance. A basic understanding of the electrical properties of these TSVs is necessary for both modeling the clock distribution network described in Chapter 6 and for analyzing the 3-D power delivery networks described in Chapter 7. The second component

of Chapter 5 focuses on closed-form expressions describing the TSV resistance, capacitance, and inductance. The closed-form expressions consider the 3-D via length, diameter, dielectric thickness, and spacing to ground. Numerical simulations of the TSV are used to quantify the electromagnetic behavior of the resistance, capacitance, and inductance for comparison with the closed-form expressions.

Models of 3-D clock distribution networks are presented in Chapter 6. Symmetric interconnect structures, such as H-trees and X-trees, are often used to distribute a clock signal in 3-D circuits. By exploiting the symmetry of the clock distribution network, the clock signal arrives simultaneously at each leaf of the tree. Maintaining this symmetry within a 3-D IC is however a difficult task. The problem is further complicated as both interplane and intraplane interconnects with different impedance characteristics are present. Models of three different 3-D clock network architectures and a description of each topology are reviewed in Chapter 6. The clock topologies have been designed and fabricated. A comparison between the analytic root-to-leaf delay and the experimental delay is also presented.

Chapter 7 is dedicated to 3-D power delivery. A quantitative analysis of the noise measured on each plane of a three plane 3-D integrated stack is provided. An analysis of the effects of the through silicon vias on *IR* voltage drops and *L·di/dt* noise is discussed, as the impedance of the power distribution network is affected by the TSV density. A comparison of two different via densities for identical power distribution networks is also described in this chapter, and implications of the 3-D via density on the power network design process is discussed. The effective placement of decoupling capacitors can potentially reduce *IR* and *L·di/dt* noise within the power network, while enhancing performance by stabilizing

the voltage fluctuations on the power rails. The effect of board level decoupling capacitors in 3-D circuits is examined. Methods for placing decoupling capacitors at the interface between planes to minimize the effects of inter-plane noise coupling are also suggested.

Once a basic understanding of topology dependent noise propagation within the power and ground networks has been provided (see Chapter 7), methods to alleviate the affect of this noise are discussed in Chapter 8. A particular focus is on the use of decoupling capacitors to minimize the effect of noise in 3-D circuits. Modern 2-D ICs already use decoupling capacitors to minimize IR and $L\frac{di}{dt}$ noise, however, best practices for 3-D IC decoupling capacitor placement have yet to be developed. Due to the proximity of device planes within a 3-D stack, where the distance between transistors can be less than 40 μm [82, 83], novel placement configurations, such as sharing decoupling capacitors between planes, are possible. A test circuit currently in fabrication by Tezzaron Semiconductor to evaluate several decoupling capacitor placement scenarios within a 3-D IC is also described in Chapter 8.

An optical interconnect system for intra-chip communications based on free-space optics has been developed by integrating heterogeneous technologies into a 3-D system. This hybrid optical system is described in Chapter 9. The optical communication is provided by dedicated lasers and germanium photodetectors, which avoids packet switching while offering ultra-low latency and scalable bandwidth. In addition, a test circuit that includes a microprocessor, transmitter and receiver circuits, as well as preconditioning circuits for data transmission among the cores via the free-space optical system has been designed and is currently being fabricated by Tezzaron Semiconductor. The analog transceiver circuits

are isolated from the microprocessor on separate device planes. The transceivers are designed to achieve a 10 Gb/s data rate. The basic operation of the system serializes the 64 bit data output from the 333 MHz microprocessor. The data are transmitted from the vertical cavity surface emitting lasers (VCSELs) and free space optical system, receiving the optical signal through the photodetectors, de-serializing the data, and analyzing the transmitted data packets on a second microprocessor.

VCSEL structures are typically afflicted with severe thermal effects. An understanding of the thermal behavior of VCSEL arrays is crucial since heat limits device performance, optical output power, threshold current, modulation speed, heat sink efficiency, and produces hot spots in 3-D systems that are already under high thermal stress. Continuing the research described in Chapter 9, a discussion of the heat generated by VCSEL arrays and thermal coupling to the active and passive devices in the planes below are discussed in Chapter 10.

A major challenge in TSV-based 3-D integrated circuits is the management of large thermal gradients across different device planes. A three-dimensional test circuit examining thermal propagation within a 3-D integrated stack has been designed, fabricated, and tested, as described in Chapter 11. Intra- and inter-plane thermal coupling is investigated through single point heat generation using resistive thermal heaters and temperature monitoring through calibrated four-point resistive measurements tested at low current to avoid joule heating. Each thermal source is paired with a thermal sensor on an adjacent metal level, and these pairs are distributed throughout each plane of the 3-D stack. Test conditions including the location of the active circuit relative to the heat sink and the alignment of two

active circuits are examined to characterize thermal propagation in 3-D integrated circuits. Design suggestions are also provided in Chapter 11 to better manage hot spot formation while reducing deleterious effects on neighboring circuit blocks.

A brief summary of the primary contributions of this dissertation is provided in Chapter 12. The primary goals of this research effort are to explore multiple aspects of the design of 3-D integrated circuits including clock and power distribution networks, electrical characterization and modeling, thermal propagation in 3-D ICs, and heterogeneous 3-D integration. Intuition behind critical design issues for 3-D ICs is provided through both experiments and models of these complex circuits. Ultimately, the material described in this dissertation provides insight that will help continue the evolution of highly integrated, multifunctional, multiplane microelectronic circuits.

Future research directions are provided in Chapter 13. Merging heterogeneous technologies in the third dimension by utilizing novel interface circuits is described. Two critical aspects of multi-plane integration is effective synchronization and power generation and distribution among disparate technologies. 3-D synchronization must consider the different frequencies of operation, clock network loads, clock distribution topologies, and skew/slew requirements of each heterogeneous device plane. 3-D power delivery must consider multi-voltage domains, different power noise requirements between heterogeneous device planes, and mismatched parasitic impedances within the power distribution networks, while ensuring thermal and current density limits are not exceeded. A universal interface is needed to

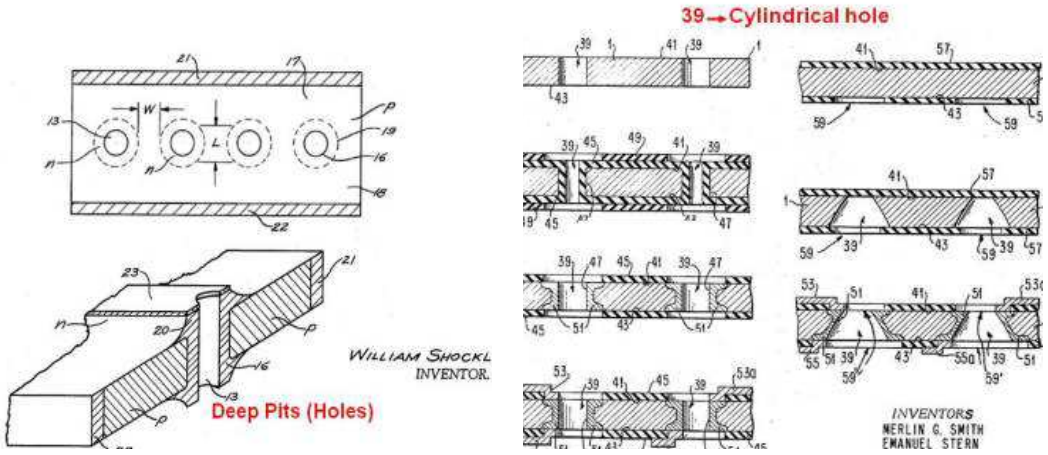
ensure effective clock delivery, power delivery, and signaling across the disparate technologies forming the stacked 3-D system. To achieve this universal interface, techniques developed for 2-D circuits and packaging, such as voltage shifters and electrostatic discharge (ESD) protection circuitry, should be adapted for 3-D stacked systems. The proposed techniques will determine the optimum number and placement of the voltage regulators and the size and placement of the decoupling capacitors that satisfy specific noise and area constraints of a 3-D system. Circuits and methodologies for efficiently delivering clock, I/O signals, and power at the interface between disparate technologies are also needed.

Chapter 2

Electrical Characterization and Modeling of Through Silicon Vias

One of the fundamental breakthroughs in the development of 3-D integrated circuits is the through silicon via (TSV). The early concept of the through silicon via was introduced in the late 1950's and 1960's, first by William Shockley, followed closely by work from Merlin Smith and Emanuel Stern, as these inventors were the first to patent a method of etching a cylindrical hole through silicon [84, 85]. Sketches of these early etched holes in silicon are depicted in Figure 2.1.

The TSV joins two separately processed wafers or dies to form a 3-D IC. Each additional device plane that is stacked requires another set of TSVs to connect signals, different power domains, and various clock signals to properly operate the newly connected device plane. As the TSV is the critical component for interconnecting the separate device planes, accurate electrical characterization of the impedance of the TSV is necessary to effectively model the inter-plane signal delay and power characteristics, to better understand the impedance of the power network, and to analyze the effects of inter- and intra-plane clock skew and slew. The primary objective is to develop an equivalent electrical π-model



(a) (b)

Figure 2.1: Sketch of early through silicon via from patents filed by (a) William Shockley [84], and (b) Merlin Smith and Emanuel Stern of IBM [85].

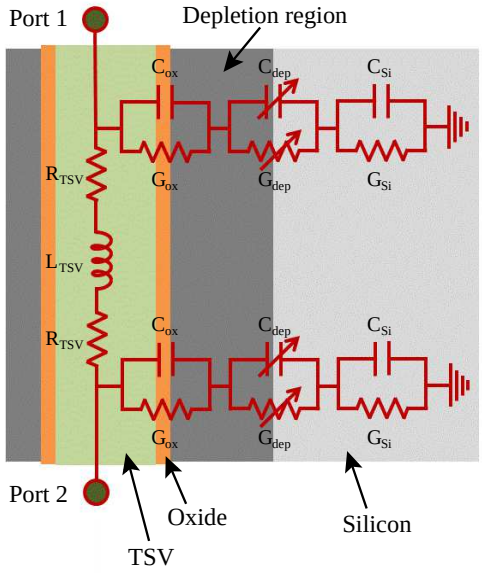


Figure 2.2: Equivalent π-model of the through silicon via.

of the TSV, as shown in Figure 2.2. These models are often used for simulation in SPICE.

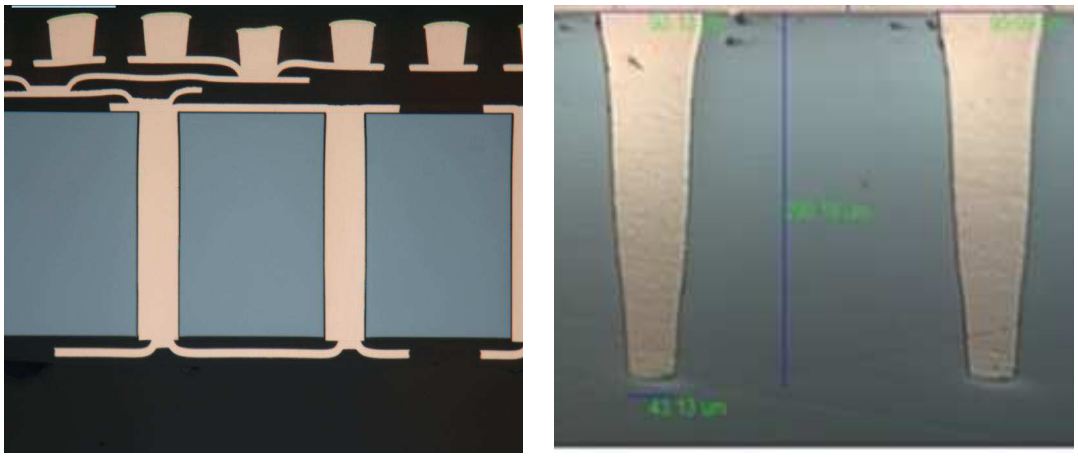
The research described in this chapter provides insight into the characterization and

compact modeling of the electrical impedance of the TSV. An introduction to through silicon via technology is provided in Section 2.1 to better understand the physical characteristics of the TSV. A description of work on numerically simulating the TSV is provided in Section 2.2. Once the electrical characteristics of the TSV are determined for a range of lengths, diameters, dielectric liner thicknesses, and other physical parameters, equivalent closed-form expressions have been developed to efficiently and accurately determine the impedance of the TSV without the need for additional numerical simulation. Existing work on compact modeling of the electrical properties of the TSV is described in Section 2.3. Some concluding remarks are provided in Section 2.4.

2.1 Through silicon via technology

Beginning with the mid 1990's, there has been an exponential growth in research to develop 3-D integrated circuits. A particular focus in the early years of 3-D integration was the development of process flows that incorporate TSV formation and wafer/die bonding into traditional fabrication processes. Some industrial leaders in this area are IBM [86, 87], LETI [88], Intel [87], and Texas Instruments [87], while public funding for fabrication process flows has produced important results at MIT Lincoln Laboratory [82, 89] and the U.S. National Laboratories [87]. Many more companies, universities, and national laboratories are developing TSV-based technology [87, 90], continuing the work on enhancing the fabrication of these stacked systems.

A multitude of process flows have been developed to generate stacked ICs. Process flows may vary depending on the depth of the silicon, material of the TSV, aspect ratio of



(a) (b)

Figure 2.3: Microphotograph of a (a) straight [99], and (b) tapered TSV [100].

the TSV, dielectric liner thickness that insulates the TSV from the silicon, and various other physical parameters. A description of these physical parameters characterizing a TSV is provided in this section.

The TSV is typically a cylindrical structure composed of either copper, aluminium, tungsten, polysilicon (for TSV fabricated during the front-end of the line processes [91]), or solder [88, 92–94]. Most current processes favor tungsten based TSVs over other materials as the conductivity is relatively high (approximately $\frac{1}{3}$ of copper), without damaging the transistors due to copper leaching into the silicon [95]. Nevertheless, a large body of work has focused on the production of copper through silicon vias [88, 96, 97]. Although the TSV is often designed as a cylinder, the shape often includes a narrowing or tapering from the surface where the transistors are patterned to the bottom of the silicon. The degree to which a TSV is tapered depends upon the fabrication process and typically ranges from close to 0° to 15° [98]. A microphotograph of both a straight and tapered TSV is shown in Figure 2.3.

There are a wide range of TSV lengths and diameters that can be produced depending upon the fabrication process. Most processes produce TSVs with aspect ratios (length to width ratio) of 10 to 15 to 1 [101–104]. There are, however, processes that produce aspect ratios as low as 2 to 3 [101] and as high as 25 [105]. The thickness of the material determines the length. For example, a silicon substrate thinned to 150 μm requires TSVs that are longer than 150 μm. The thickness of the substrate is therefore a primary consideration that determines the diameter of the TSV. This example of a 150 μm thick silicon substrate can be used to illustrate the correlation between substrate thickness and TSV diameter. If a predefined aspect ratio of 10 is assumed for a 150 μm substrate, a minimum TSV diameter of 15 μm is required. Although the thickness of the substrate appears to limit the diameter of a TSV, TSVs have been produced with various aspect ratios to accommodate the thickness of a particular substrate [88, 101, 105–107].

The dimensions of the through silicon via also affect the type of fabrication process for the transistors. The buried oxide of a silicon-on-insulator (SOI) process behaves like a natural etch stop layer during the wafer thinning process necessary to stack device planes [108]. Typical substrate thicknesses after wafer thinning for an SOI process are in the range of 10 to 20 μm [108], which suggests that TSV diameters of 1 to 3 μm can be fabricated. TSVs with diameters ranging from 1 to 60 μm and lengths of 10 [82] to 300 μm have been fabricated [96, 103].

Additional physical parameters of the TSV worth noting is the dielectric liner and seed layers. Both the liner (typically SiON, SiN, TiN, or SiO_2 [91, 109, 110]) and seed

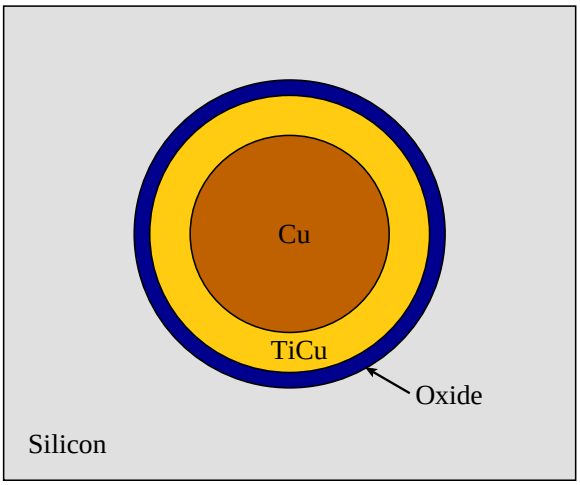


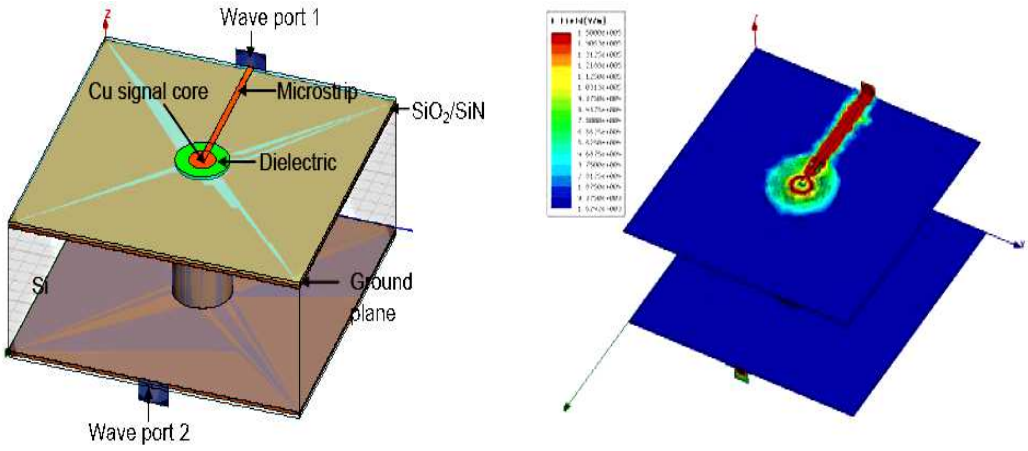
Figure 2.4: Top view of a TSV in silicon depicting the oxide layer, TiCu seed layer, and copper TSV.

layer (TiCu, TaN, or TEOS [111]) are deposited through plasma-enhanced chemical-vapor-deposition (PECVD) at temperatures of about 200°C [91], atomic layer deposition for low temperatures of 100°C [91], or electrochemical deposition (ECD) [104]. The oxide liner is typically 50 to 500 nm thick and is used to isolate the TSV from the silicon substrate [104, 109, 110, 112]. The thickness of the oxide is increased to 1 to 3 μm for TSV diameters larger than 20 μm. The seed layer is typically 500 to 600 nm thick and is used as a bonding layer for the TSV conductor material. The material for the seed layer depends on the TSV fill material. As an example, for a copper TSV process, a titanium copper (TiCu) seed layer is used with thicknesses of 60 to 80 nm of titanium and 400 to 500 nm of copper [104]. A top view of the oxide layer, TiCu seed layer, and copper TSV is shown in Figure 2.4. The combined thickness of the liner and seed layer is approximately 600 nm to 1 μm [104, 109]

2.2 Electrical characterization through numerical simulation

Characterization of the electrical properties of the through silicon via is often completed through numerical simulation. The basic requirement of a numerical simulator is the physical structure of the item under study, including the surrounding environment and corresponding material properties. There are several full wave and 2.5D electromagnetic simulation tools available including Quick 3D [113] and High-Frequency Structural Simulator (HFSS) [114–116] from Ansys, EIP from IBM [117], FastHensry [118] and FastCap from MIT [119], EMPro 3D from Agilent [120,121], Sonnet Suites from Sonnet [122,123], and 3D EM Field Simulation from CST [124], to name a few.

Current numerical simulators support both the graphical input of shapes and structures, material properties, and shape manipulations through a graphical user interface (GUI) or coded input that includes the coordinates of the shapes (rectangles, circles, spheres, prisms, etc.), shape manipulation commands, and the material properties of each shape. There are benefits and drawbacks for each technique. The use of the graphical interface is easy and intuitive, so one can quickly learn to form and define shapes and structures for simulation. The drawback is that any changes in the physical structure of the TSV, in particular the length, requires a completely new set of shapes and structures to match the change in length. Alternatively, there is a steep learning curve to properly code the structures, shapes, material properties, and shape manipulations for a given simulator. Any change in a physical structure is however easily propagated to other shapes through a simple search



(a) (b)

Figure 2.5: Image of (a) TSV structure in HFSS for high frequency simulation, and (b) resulting electric field distribution [125].

and replace. A physical structure described in HFSS through the graphical interface is depicted in Figure 2.5(a), and the resulting electric field distribution is shown in Figure 2.5(b). These field solvers can produce electric field distributions, magnetic field distributions, current and voltage distributions, and compute the capacitance, conductance, inductance, and resistance matrices by solving Maxwell's equations through the method of moments and finite-element method techniques. Maxwell's equations are listed in Table 2.1 in both differential and integral form as satisfying this set of equations for the physical geometries and material properties describes the electrical properties of the structure.

Table 2.1: Maxwell equations solved by electromagnetic field solvers to determine the electrical impedance of a TSV. The relationship of the field vectors is described by $\vec{D} = \epsilon\vec{E}$ and $\vec{B} = \mu\vec{H}$.

Law	Differential form	Integral form
Gauss's Law	$\nabla \cdot \vec{D} = \rho$	$\oint_S \vec{D} \cdot d\vec{A} = \int_V \rho d\vec{V}$
Gauss's Law (magnetism)	$\nabla \cdot \vec{B} = 0$	$\oint_S \vec{B} \cdot d\vec{A} = 0$
Faraday's Law	$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}$	$\oint_C \vec{E} \cdot d\vec{l} = -\frac{d}{dt} \int_S \vec{B} \cdot d\vec{A}$
Gauss's Law (magnetism)	$\nabla \times \vec{H} = \vec{J} + \frac{\partial \vec{D}}{\partial t}$	$\oint_C \vec{B} \cdot d\vec{l} = \int_S \vec{J} \cdot d\vec{A} + \frac{d}{dt} \int_S \vec{D} \cdot d\vec{A}$

The resistance, capacitance, inductance, and conductance of a TSV varies significantly depending upon the size of the via. Numerical simulations have produced DC resistances ranging from 200 to 250 mΩ for large aspect ratio TSVs [88, 126], to as small as 10 to 30 mΩ [86, 112, 127, 128] for small aspect ratios. The effect of skin depth on the AC resistance has also been explored through numerical simulations, and a near doubling of the DC resistance is possible at a frequency of 10 GHz [128].

The self-capacitance of a TSV, the capacitance assuming a ground at infinity, also varies depending on the size of the TSV and the applied voltage. The applied voltage affects the surrounding silicon material by operating like a MOS capacitor in accumulation, depletion, or inversion. In addition, the thickness of the oxide liner also significantly affects the capacitance of the TSV [112, 129, 130]. The self-capacitance ranges from as low as 1 to 5 fF for small TSVs with 1 to 2 μm diameters and lengths of 10 to 15 μm to as high as 200 to 250 fF for large TSVs with 30 to 40 μm diameters and lengths of 100 to 150 μm [112, 129].

The self-inductance of a TSV ranges from 2 to 5 pH for small TSVs to as large as 40 to 50 pH for larger sized TSVs (the same TSV dimensions as in the capacitance discussion) [112, 129]. The mutual inductance follows a similar trend with values ranging from 1 to 2 pH for small TSVs and 10 to 25 pH for larger sized TSVs. The mutual inductance does depend on the space between TSVs and falls off exponentially to a non-zero asymptotic value. The loop inductance, therefore, can range from as low as approximately 10 to 15 pH to values exceeding 60 pH [129]. An expression to calculate the loop inductance is

$$L_{loop} = L_{11} + L_{22} - 2 \cdot L_{21}, \tag{2.1}$$

where L_{11} and L_{22} are, respectively, the self-inductance of the forward and return current paths, and L_{21} is the mutual inductance between the two TSVs.

Characterization of the TSV through numerical simulation can be computationally expensive depending upon the physical structure of the environment, and time consuming as any change in physical parameters such as the TSV diameter and length and/or the number of TSVs requires a complete readjustment of the numerical analysis process. The results of numerical simulations are therefore often used to evaluate the accuracy of compact models of the electrical impedance of a TSV. Most of the compact models discussed in Section 2.3 are compared to numerical simulation, and a few models are also compared to experimental results. Compact models that have been verified with numerical simulation provide a computationally efficient and accurate method to determine the electrical impedances of a TSV.

2.3 Compact models

Compact models of the through silicon via provide a quick and accurate method to calculate the TSV impedance. A large body of work exists that focuses on the electrical modeling of the through silicon via [112, 129], including bundled TSVs [131] and high frequency effects [110, 125, 127, 132]. In addition, a subset of research on compact TSV models focus on including the tapered sidewall of the TSV [98, 133].

General closed-form expressions for the resistance, inductance, capacitance, and conductance of the TSV structure consider the diameter, length, dielectric liner thickness, distance to ground, distance to neighboring TSVs, and tapering angle [98, 127, 134]. A discussion of the closed-form expressions of the individual electrical impedances is provided below.

2.3.1 TSV resistance

A brief discussion of the physical principles governing the calculation of the TSV resistance is provided here. Properties that effect the resistance such as surface scattering, boundary scattering, and skin effect are also described.

The resistance relates the potential difference at the two ends of a wire to the total current flowing through a wire. The fundamental electrostatic law governing resistance is

$$R_{via} \equiv \frac{\Phi_{12}}{I} = \frac{-\int_L \vec{E} \cdot d\vec{l}}{\int_A \sigma \vec{E} \cdot d\vec{l}}, \tag{2.2}$$

where Φ_{12} is the electrostatic potential between two points, I is the current through the wire, σ is the conductivity of the material, and E is the electric field.

The DC resistance of a TSV is determined for a non-magnetic cylindrical wire conductor. A majority of closed-form expressions for the resistance of a TSV assumes a cylindrical wire [110, 112, 129, 131, 135, 136]. The resistance of a TSV is therefore a function of the length ℓ_v, cross-sectional area πr_v^2, and conductivity of the material σ.

$$R_{via} = \frac{\ell_v}{\sigma \pi r_v^2}. \tag{2.3}$$

a. High frequency effects

One of the primary high frequency effects that affects the resistance of the TSV is the skin effect. At DC, the current density through the TSV is uniformly distributed across the entire cross-sectional area. As the frequency of the signal propagating through the TSV increases, the current density becomes non-uniform, and drops exponentially with distance from the surface (sidewall of TSV). This phenomena is explicitly the result of a time-varing magnetic field H induced by a time-varying current I, with the assumption that the current return path is sufficiently far away that any influence can be neglected [137]. The skin effect leads to current crowding at the surface, reducing the effective cross-sectional area of the TSV. The skin effect is defined as the depth below the surface of the conductor at which the current density drops to $\frac{1}{e}$ (approximately 0.37) of the current density at the surface of the conductor. The skin depth is

$$\delta = \frac{1}{\sqrt{\pi f \mu_o \sigma}}, \tag{2.4}$$

where f is the frequency, σ is the material conductivity, and μ_o is the permeability of free space.

As a result of the skin effect, which reduces the effective cross-sectional area of a conductor, the resistance of a rectangular wire with width w and thickness t can be written as

$$R \approx \frac{\ell}{\sigma 2\delta[t + (w - 2\delta)]}, \tag{2.5}$$

and the resistance of a cylindrical wire as

$$R \approx \frac{\ell}{\sigma\pi[r^2-(r-\delta)^2]} \approx \frac{\ell}{\sigma\pi(2r-\delta)\delta}. \tag{2.6}$$

The -2δ term in (2.5) removes the overlap in area between the TSV thickness t and width w at the corners.

b. Effect of barrier thickness on resistivity

As previously mentioned, the TSV is isolated from the silicon by a dielectric liner that prevents direct contact to the silicon. In the case of copper based TSVs, the dielectric liner also includes a metal barrier with a higher resistivity that prevents the diffusion of copper into the silicon. The barrier thickness is dependent on both the deposition process and the barrier material [138]. When examining a cross-sectional view of the TSV, the area occupied by the metallic current carrying portion is the conducting material deposited after the liner and barrier is formed. A thicker liner/barrier, therefore, leaves a smaller diameter to fill with the metallic conductor. Since the resistivity of the dielectric liner is much higher than the metal, the result is an increase in resistance. There are two methods to account for this characteristic. One method reduces the effective diameter of the TSV depending on the thickness of the liner/barrier. For the second method, the resistivity is modified to include the effect of the liner/barrier. Banerjee in [138] applied the second technique and modified the resistivity, as described by

$$\rho_b = \frac{\rho_o}{1-\frac{A_b}{A_{metal}\cdot(\frac{p_w}{2})^2}}. \tag{2.7}$$

The effective resistivity due to the barrier ρ_b is inversely proportional to the area of the barrier A_b and proportional to the area of the metal conductor A_{metal}. Banerjee also accounted for the pitch between wires p_w, but this term can be ignored for TSVs. The bulk resistivity of the metal ρ_o is temperature dependent and must therefore be adjusted based on the temperature of operation.

c. Temperature effect

The conductivity σ of a material is dependent on the carrier concentration q and the mobility of the charge carriers μ, and is

$$\sigma = q\mu. \tag{2.8}$$

Charge carriers are created by the ionization of atoms comprising the lattice of a conductor. At a specific temperature, essentially all of the atoms of a conductor are ionized, and the supply of electrons is constant with temperature. The increase in resistance due to increasing temperature is caused by a higher rate of collision with the lattice, impurities, and grain boundaries of the conductor material. With this increase in collision rate, the mobility of electrons is reduced.

For conductors, the loss of mobility is due to ionic scattering. The relationship between temperature and resistance includes this effect on the material resistivity. The temperature dependent resistivity is described by [139]

$$\rho(T) = \frac{1}{\sigma(T)} = \rho_o(T_o)[1 + \alpha(T - T_o)], \tag{2.9}$$

where $\rho(T)$ is the wire resistivity, $\rho_o(T_o)$ is the resistivity at a reference temperature T_o (often taken to be 20°C), and α is the temperature coefficient of resistance. The temperature coefficient of resistance is $\frac{d\rho}{dT}$, the slope of ρ plotted against the temperature. The α term for a specific material changes with the reference temperature, and therefore, the relationship between resistivity and temperature only holds for a range of temperatures around this reference temperature [140].

2.3.2 TSV inductance

The inductance of a conductor is a measure of the magnetic field lines that are internal to and surround the wire, characterizing the ability of a conductor to couple magnetic flux or store magnetic energy. The physical layout of a conductor, current density, current direction, and signal frequency affects the inductance of a wire. The inductance of a material is defined as

$$L \equiv \frac{\oint \vec{B} \cdot d\vec{A}}{I}. \tag{2.10}$$

A critical characteristic that is necessary to determine the inductance of a wire is the current return path. Inductance is based on the 'loop' formed as current flows into and out of an area. The current return path is not easy to identify in modern integrated circuits as it is not a function of proximity to the source structure such as the capacitance. The current return paths in modern ICs are primarily in the power distribution network and other neighboring wires [141]. The complexity of extracting the inductance is further exacerbated since the current return paths can be tens to hundreds of micrometers from

the wire under consideration. In addition, the return path does not have to be through a single wire, potentially leading to tens or even hundreds of return paths. A method proposed by Rosa [142] and further expanded upon by Ruheli [143] assumes that the current return path is at infinity. This approach provides an accurate estimate of inductance as compared to field solvers. Most closed-form expressions of the self- and mutual inductance of a TSV [112, 126, 132, 134] are modifications of the inductance expressions developed by Rosa [142].

a. Internal and external magnetic field

Magnetic flux exists both internal to and external to a current carrying wire. Determining the inductance can therefore be divided into internal and external components, which are summed to provide the total inductance of a wire. The internal component of the inductance is affected by the frequency of the signals propagating through the wire. At high frequencies, due to the skin affect (as discussed in Subsection 2.3.1a.), the internal component of the inductance is lower as most of the current flows at the surface of the conductor. As the frequency is increased, the internal component approaches zero as almost all of the magnetic flux is external to the conductor.

b. High frequency effects

As noted above, the density and cross-sectional profile of the current flowing through a wire affect the inductance of a wire. There are two effects at high frequency (typically exceeding 1 GHz) that change the current distribution within a conductor. The first effect is the skin effect, and the other effect, the proximity effect, is related to the skin effect.

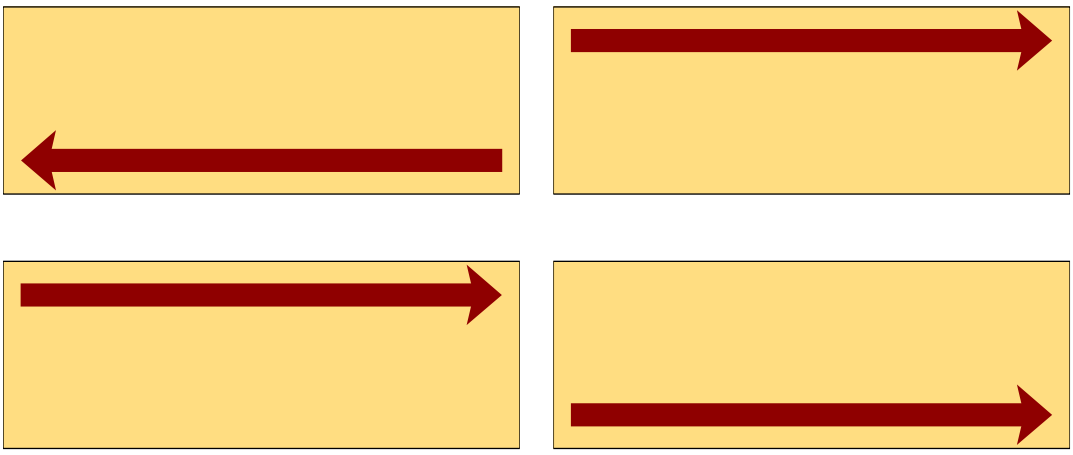


Figure 2.6: Current profile due to the proximity effect for (a) currents propagating in opposite directions, and (b) currents flowing in the same direction [137].

The proximity effect is a phenomenon that occurs between two current carrying wires. When current flows in opposite directions, such as the case of a conductor and the return path, the current crowds towards the location between the two conductors, as illustrated in Figure 2.6(a). In the case where the current flows in the same direction through two separate conductors, the current moves to the opposite side of each conductor, farthest from the neighboring wire. In both cases, current flowing in the same or opposite direction, the proximity effect reduces the overall inductance of a conductor.

A third high frequency effect that changes the inductance of a wire or TSV is due to multi-path current redistribution in the return path [141, 144]. There are many possible current return paths in an integrated circuit, including the power and ground networks, the substrate, and nearby signal lines. Each of these returns paths exhibits a frequency dependent impedance. Current is redistributed among the different return paths to minimize the total impedance of the path. Therefore, depending upon the frequency of the signal, certain return paths are more effective in minimizing the total impedance than other paths,

and the path does not have to remain the same from one frequency to another frequency.

c. Loop inductance

The loop inductance is induced by a current flowing through a wire and a return path. In the case of multiple return paths, the total loop inductance of a wire is the sum of the partial loop inductances with each return path segment. When the return paths are not known, a mathematical representation of the partial loop inductance elements is determined for all wires forming the current loop. There are two approaches to solving for the effective loop inductance of a path. The first method directly calculates the partial loop inductance between two coupled segments, while the second method uses mathematical representations of the self- and mutual inductance to all other segments before determining the loop inductance with each return path. An overview of each technique is described below, starting with the first method of directly solving for the partial loop inductance. The partial loop inductance of two segments is [145, 146]

$$L_{ab,partial} = \frac{\mu_o}{4\pi}\frac{1}{A_a A_b}\int_{A_a}\int_{\ell_a}\int_{A_b}\int_{\ell_b}\frac{d\vec{\ell_a}\cdot d\vec{\ell_b}}{|\vec{r_a}-\vec{r_b}|}dA_a dA_b, \tag{2.11}$$

where A_a and A_b are the cross-sectional area of, respectively, segments a and b, and ℓ_a and ℓ_b are the length of each respective segment.

The total loop inductance, which is the sum of all partial self- and mutual inductances, is therefore given by (2.12). The s_{ij} term in (2.12) is -1 when the currents flow in opposite directions and +1 when currents flow in the same direction.

$$L_{loop_1} = \sum_i \sum_j s_{ij} L_{ij,partial}. \tag{2.12}$$

The second method is based on the closed-form expressions of the self- and mutual inductance as developed by E. B. Rosa [142]. Rosa derived the self-inductance of a cylindrical wire of non-magnetic material in free space. An expression for the self-inductance is

$$L_{self} = \frac{\mu_o}{2\pi}\left[ln\left(\frac{\ell + \sqrt{\ell^2 + r^2}}{r}\right)\ell + r - \sqrt{\ell^2 + r^2} + \frac{\ell}{4}\right], \tag{2.13}$$

where r and ℓ are, respectively, the radius and length of the TSV, and μ_o is the permeability of free space. Most closed-form expressions of the inductance of a TSV are based on this expression by Rosa, with minor modifications. One example is the expression by Pucel [112, 147] in (2.14). Pucel removed the $\frac{\ell}{4}$ term, doubled the effective length of the conductor, and halved the resulting inductance to account for the image inductance caused by a vertical current proximate to a ground plane [148].

$$L_{self} = \frac{\mu_o}{4\pi}\left[ln\left(\frac{2\ell + \sqrt{(2\ell)^2 + r^2}}{r}\right)2\ell + r - \sqrt{(2\ell)^2 + r^2}\right]. \tag{2.14}$$

The mutual inductance between two cylindrical wires has also been described by Rosa [142], and is applied in the development of closed-form expressions for cylindrical TSVs. The mutual inductance is

$$L_{mutual} = \frac{\mu_o}{2\pi}\left[ln\left(\frac{\ell + \sqrt{\ell^2 + p^2}}{p}\right)\ell + p - \sqrt{\ell^2 + p^2}\right]. \tag{2.15}$$

The total loop inductance is calculated from the self- and mutual inductance of a wire and each of the return paths.

$$L_{loop_2} = L_{self_1} + L_{self_2} - 2L_{mutual_{21}}. \tag{2.16}$$

2.3.3 TSV capacitance

Unlike the resistance and inductance (assuming that all return paths are known for each wire) of a material, where accurate closed-form expressions have been developed that relate the material properties, physical characteristics, and dimensions, closed-form expressions of the capacitance require a model that incorporates additional effects beyond dimensional and material properties. One effect is the termination of the electric field lines on neighboring metals, which plays a significant role in the total capacitance of a conductor. Two important but different considerations include: 1) field lines emanating from the conductor (TSV) and terminating to a reference ground plane on the back metal of the silicon, and 2) field lines emanating from the TSV and terminating on another TSV within the silicon. In each case, a depletion region (or accumulation region depending on the applied voltage) forms around the TSV that needs to be incorporated into the closed-form expression for the capacitance of a TSV. A third effect is the fringe capacitance, as neglecting this component can lead to underestimating the total capacitance. Recent research has focused on terminating the electric field lines not only to other TSVs but also to the on-chip interconnect within a device plane [132, 149], further increasing the complexity of the compact model characterizing the capacitance of a TSV.

In the simplest of terms, the capacitance is the ability of a body to store electric charge, and is the ratio of the charge to the applied voltage. The capacitance between two conductors is

$$C \equiv \frac{Q}{\Phi_{12}} = \frac{\oint_S \vec{D} \cdot d\vec{A}}{-\int_A \sigma \vec{E} \cdot d\vec{\ell}} \ , \tag{2.17}$$

where Φ_{12} is the voltage difference between the two conductors, ℓ is any path from the negative to the positive conductor, and A is any surface enclosing the positively charged conductor. An analytic solution of Poisson's equation is necessary to determine the capacitance of a material, which is dependent on the location of the ground reference. As a result, extracting the capacitance with electric field solvers such as Ansys Q3D [113] and FastCap [119], although highly accurate, is often computationally expensive. Closed-form expressions are therefore critical to accurately and efficiently extract the capacitance of a TSV, while including physical effects such as the depletion region, fringe capacitance, and the location of the ground reference to enhance the accuracy of the model.

Early research in compact modeling of the capacitance of a TSV assumed the silicon substrate behaved similarly to a metal with a finite conductivity. Although this assumption is somewhat correct, the result is that all field lines emanating from a TSV are assumed to terminate on the other side of the oxide barrier layer (at the interface with the silicon). The assumption produces a worst case estimate of the capacitance of a TSV. Two expressions for estimating the TSV capacitance based on the assumption that silicon behaves like a highly resistive metal are

$$C = \frac{\epsilon_{ox}}{t_{ox}} 2\pi r\ell, \tag{2.18}$$

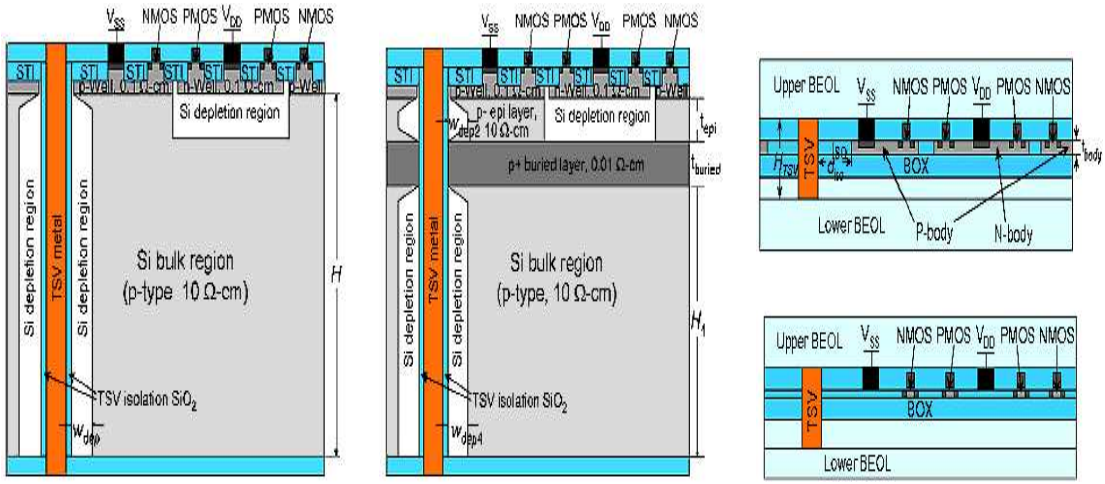
$$C = \frac{\epsilon_{ox}}{ln\left(\frac{r+t_{ox}}{r}\right)} 2\pi\ell, \tag{2.19}$$

where t_{ox} is the oxide thickness, r is the radius of the TSV, and ℓ is the length of the TSV. Equation (2.18) is used in [126] to compare this model with experimental data. The closed-form expression of (2.19) modifies (2.18) to more accurately represent the radius of the conductor within the dielectric without a change in the underlying assumption that the silicon behaves like a metal with low conductivity (1 to 10 S/m) [150].

From this early work, a more accurate model of the physical phenomena affecting the capacitance of a TSV has been developed. One of the most important of these phenomena is the formation of a depletion region around the TSV [112, 127, 129, 130, 132], such that the silicon can no longer be modeled as a highly resistive metal. Inclusion of the depletion region in a TSV model is further discussed in the following subsection.

a. TSV model including the effects of the depletion region

The through silicon via is surrounded by silicon. Similar to a MOS capacitor [151,152], the applied voltage on the TSV either depletes the silicon of mobile charge carriers or causes accumulation [127, 130]. The voltage on the TSV can also invert the surrounding silicon when a low frequency signal is passed through the TSV [130]. The silicon is typically depleted as a result of a positively applied voltage on the TSV and the p-type silicon body. The formation of a depletion region for bulk and SOI technologies is illustrated in



(a) (b) (c)

Figure 2.7: Cross-sectional view of different CMOS technologies with TSVs depicting the formation of a depletion region around the TSV in (a) bulk CMOS, and (b) bulk CMOS with p+ buried layer. TSVs in either PD-SOI (shown in (c) top) or FD-SOI (shown in (c) bottom) reveal minimal formation of a depletion region [127].

Figure 2.7.

As previously mentioned, the applied voltage determines whether the TSV operates in either the accumulation region, the depletion region, or in inversion, similar to a MOS capacitor. The ratio of the TSV capacitance to the oxide capacitance as a function of the applied voltage V_g is shown in Figure 2.8.

As noted in Figure 2.8, the capacitance in accumulation is the oxide capacitance [112],

$$C_{TSV_{acc}} = C_{ox} = \frac{2\pi\epsilon_{ox}\ell}{ln\frac{r_{ox}}{r_{metal}}}, \tag{2.20}$$

and the depletion capacitance is

$$C_{dep_{min}} = \frac{2\pi\epsilon_{ox}\ell}{ln\frac{r_{max}}{r_{ox}}}, \tag{2.21}$$

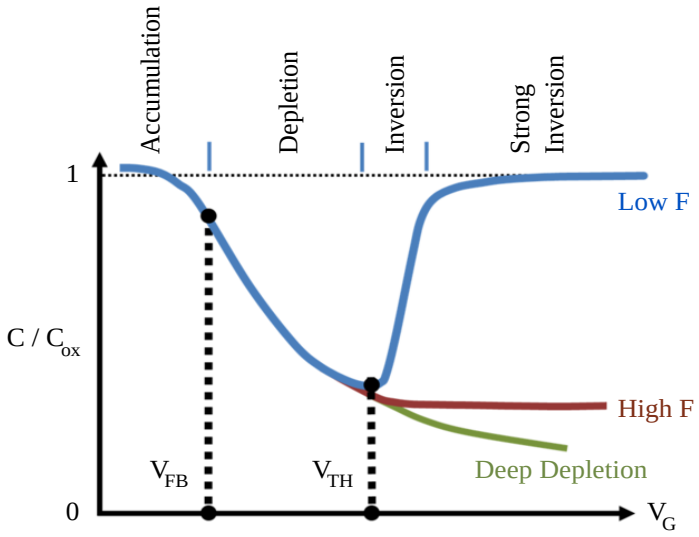


Figure 2.8: Ratio of the total TSV capacitance to the oxide capacitance as a function of applied voltage V_g.

where r_{max} is the maximum depletion radius, r_{metal} is the radius of the TSV, r_{ox} is the radius of the conductor with the oxide, and ℓ is the length of the TSV. Since the depletion capacitance is in series with the oxide capacitance, the total TSV capacitance is reduced. The series combination of the oxide and depletion capacitance is [112, 153]

$$C_{TSV_{dep}} = \frac{C_{ox}Cdep}{C_{ox}+Cdep}. \tag{2.22}$$

2.3.4 TSV conductance

There are two conductive components of a TSV. One component is due to the dielectric liner surrounding the through silicon via, and the second component is due to the lossiness of the silicon substrate with a typical conductivity of 1 to 10 siemens per meter. An expression for the conductance between two conductors in close proximity is

$$G = \frac{\ell \pi \epsilon}{ln\left[\frac{s}{2r} + \sqrt{\frac{s}{2r}^2 - 1}\right]}, \tag{2.23}$$

where, in addition to the radius r, length ℓ, and permittivity of the material, the spacing s between TSVs is also included [154, 155]. The conductance described by (2.23) is divided in half when two equal distant TSVs are in close proximity, and by four if there are four surrounding TSVs.

A second approach to model the conductance of a TSV considers the loss tangent [156] of an inhomogeneous or lossy material. In this case, the capacitance is strongly dependent on the frequency. Given that the complex permittivity is

$$\epsilon = \epsilon' - j\epsilon'' = \epsilon' - j(\epsilon_b + \frac{\sigma}{\omega}), \tag{2.24}$$

and the loss tangent is

$$tan\delta = \frac{\epsilon''}{\epsilon'} = \frac{\sigma}{\omega\epsilon'}, \tag{2.25}$$

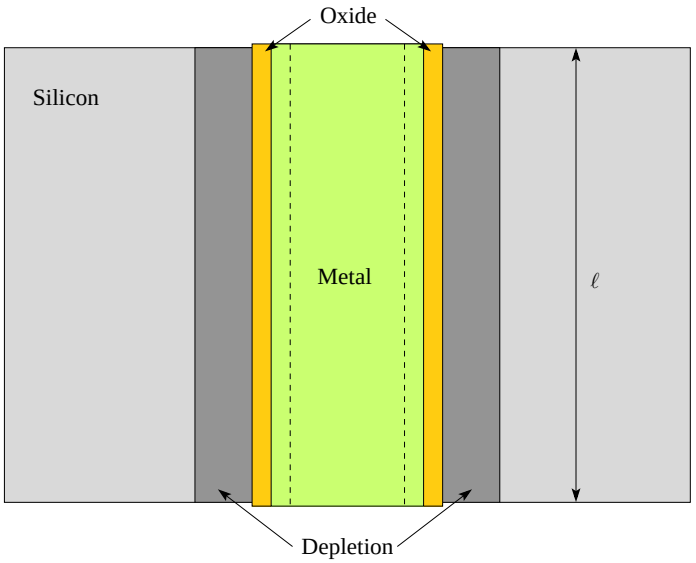
the conductance of the TSV is

$$G = \omega C tan\delta. \tag{2.26}$$

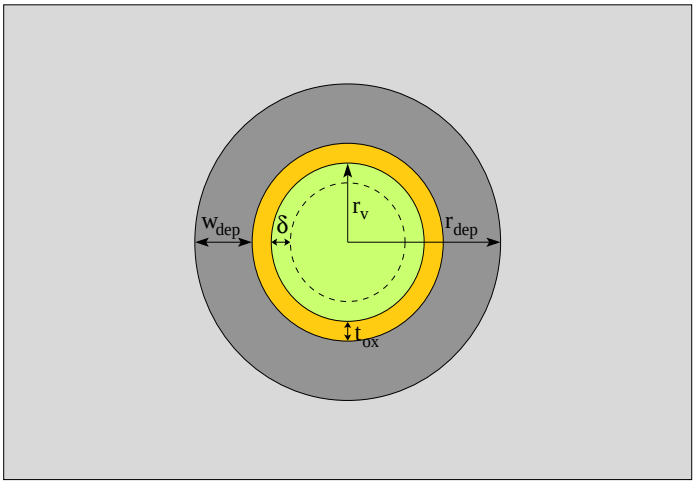
The loss tangent in (2.25) exhibits an inverse frequency dependence with ω as both ϵ and σ are also dependent on the frequency [156]. For this reason, many materials exhibit a constant dependence of $\frac{\sigma}{\omega\epsilon}$ on ω over a wide frequency range [156].

2.3.5 Comparison of TSV Impedance Models

Several compact models for the resistance and inductance of a TSV are listed in Table 2.2, and compact models of the capacitance and conductance are listed in Table 2.3. The relevant physical parameters that affect each electrical property of the TSV are also listed. The physical parameters include the TSV length ℓ, radius r_v, oxide thickness t_{ox}, TSV-to-TSV pitch p_v, depletion region width w_{dep}, and skin depth δ. Material parameters include the permittivity of the silicon (ϵ_{Si}) and oxide (ϵ_{ox}), conductivity of the silicon (σ_{Si}), resistivity of the metal ρ, and permeability of the metal μ_o. The depletion region width includes the dependence on the doping concentration of the silicon. In addition, a few expressions include the dependence on the applied voltage V and a TSV threshold voltage V_{th} similar to a MOS capacitor, which also exhibits a dependence on the material properties of the silicon. The various physical parameters and materials are illustrated in Figure 2.9.



(a)



(b)

Figure 2.9: The physical parameters and materials used in the compact models of the TSV listed in Table 2.2 are included in the (a) side and (b) top view of a single device plane.

Table 2.2: Comparison of closed-form expressions of the resistance and inductance of a TSV.

Label	Equation	Reference	Parameters	Notes
(2.27)	$R_{DC} = \frac{\rho\ell}{\pi r_v^2}$	[112, 131, 134, 129, 132, 157]	length ℓ, radius r_v, resistivity ρ	
(2.28)	$R_{DC} = \frac{\rho\ell}{\pi r_1(r_1+\ell r_2)}\left[1+\frac{1}{2}r_2^2\right]$	[98]	length ℓ, smallest TSV radius r_1, largest TSV radius r_2, resistivity ρ	tapered TSV
(2.29)	$R_{AC} = \frac{\rho\ell}{\pi\left(r_v^2-(r_v-\delta)^2\right)} = \frac{\rho\ell}{\pi(2r_v\delta-\delta^2)}$	[132, 134]	length ℓ, radius r_v, resistivity ρ, skin depth δ	[134] includes fitting parameter
(2.30)	$R_{AC} = \frac{\rho\left(2+r_2^2\right)}{4\pi r_2\delta}log\left(1+\frac{2r_2\ell}{2r_1-\delta}\right)$	[98]	length ℓ, smallest TSV radius r_1, largest TSV radius r_2, skin depth δ	tapered TSV
(2.31)	$L = \frac{\mu_o}{4\pi}\left[2\ell ln\left(\frac{2\ell+\sqrt{r_v^2+(2\ell)^2}}{r_v}\right)+r_v-\sqrt{r_v^2+(2\ell)^2}\right]$	[112]	length ℓ, radius r_v, permeability of metal μ_o	uses Goldfarb inductance [148]
(2.32)	$L = \frac{\mu_o}{2\pi}\left[\ell ln\left(\frac{\ell+\sqrt{r_v^2+\ell^2}}{r_v}\right)+r_v-\sqrt{r_v^2+\ell^2}+\frac{\ell}{4}\right]\alpha$	[98, 126, 134]	length ℓ, radius r_v, permeability of metal μ_o, fitting parameter α	[134] added fitting parameter to Rosa inductance [142]
(2.33)	$L = \frac{\mu_o}{4\pi}\left[2\ell ln\left(\frac{\ell+\sqrt{r_1^2+\ell^2}}{r_1}\right)+2r_1-2\sqrt{r_1^2+\ell^2}+\frac{\ell}{2} + \frac{4\beta\ell\left(\ell^2-(\ell+r_1)\left(\sqrt{\ell^2+r_1^2}-r_1\right)\right)}{\left(\sqrt{\ell^2+r_1^2}-\ell\right)(4r_1+\beta\ell)}\right]$	[98]	length ℓ, smallest TSV radius r_1, largest TSV radius r_2, permeability of metal μ_o, tapering angle relationship β ($\beta = tan\theta = \frac{r_2-r_1}{\ell}$)	(2.33) goes to (2.32) as β goes to 0
(2.34)	$L = \frac{\mu_o}{2\pi}\ell\left[ln\left(\frac{2\ell}{r_v}\right)-\frac{3}{4}\right]$	[126, 142]	length ℓ, radius r_v, permeability of metal μ_o	for $\ell \gg r_v$
(2.35)	$L = \frac{\mu_o}{2\pi}\ell\left[ln\left(\frac{2\ell}{r_v}\right)-1\right]\alpha$	[126, 134, 142]	length ℓ, radius r_v, permeability of metal μ_o, fitting parameter α	when skin depth dominates; [134] added fitting parameter
(2.36)	$L = 0.65\frac{\mu_o}{2\pi}\ell\left[ln\left(\frac{2\ell}{r_v}\right)-1\right]$	[126]	length ℓ, radius r_v, permeability of metal μ_o	modified [142] to match measurements
(2.37)	$L = \frac{\mu_o\ell}{2\pi}ln\left(1+\frac{2.84\ell}{\pi r_v}\right)$	[131]	length ℓ, radius r_v, permeability of metal μ_o	modified [142]; 20μm ≤ ℓ ≥ 140μm, 10μm ≤ r_v ≥ 40μm
(2.38)	$L_{21} = \frac{\mu_o}{2\pi}\left[\ell ln\left(\frac{\ell+\sqrt{p_v^2+\ell^2}}{p_v}\right)+p_v-\sqrt{p_v^2+\ell^2}\right]\beta$	[98, 134]	length ℓ, TSV-to-TSV pitch p_v, metal permeability μ_o, fitting parameter β	[134] added fitting parameter to Rosa inductance [142]
(2.39)	$L_{21} = \frac{\mu_o}{2\pi}\left[\ell ln\left(\frac{\ell+\sqrt{p_v^2+\ell^2}}{p_v}\right)+p_v-\sqrt{p_v^2+\ell^2} + \frac{2\beta\ell\left(\ell^2-(\ell+p_v)\left(\sqrt{\ell^2+p_v^2}-p_v\right)\right)}{\left(\sqrt{\ell^2+p_v^2}-\ell\right)(4p_v+\beta\ell)}\right]$	[98]	length ℓ, smallest TSV radius r_1, largest TSV radius r_2, permeability of metal μ_o, tapering angle relationship β ($\beta = tan\theta = \frac{r_2-r_1}{\ell}$)	(2.39) goes to (2.38) as β goes to 0

Table 2.3: Comparison of closed-form expressions of the capacitance and conductance of a TSV.

Label	Equation	Reference	Parameters	Notes
(2.40)	$C_{ox} = \frac{\epsilon_o \epsilon_{ox}}{t_{ox}} 2\pi r_v \ell$	[126, 157]	length ℓ, radius r_v, oxide thickness t_{ox}, oxide permittivity ϵ_{ox}	
(2.41)	$C_{ox} = \frac{\epsilon_o \epsilon_{ox}}{ln\left(\frac{r_v + t_{ox}}{r_v}\right)} 2\pi\ell$	[150, 158, 159, 112, 135]	length ℓ, radius r_v, oxide thickness t_{ox}, oxide permittivity ϵ_{ox}	
(2.42)	$C_{ox} = \frac{\epsilon_o \epsilon_{ox}(r_v - t_{ox})}{t_{ox}} 4\ell$	[160]	length ℓ, radius r_v, oxide thickness t_{ox}, oxide permittivity ϵ_{ox}	
(2.43)	$\frac{1}{C} = \frac{1}{2\pi\ell\epsilon_{ox}} ln\left(\frac{r_v + t_{ox}}{r_v}\right) + \frac{1}{2\pi\ell\epsilon_{Si}} ln\left(\frac{r_v + t_{ox} + w_{dep}}{r_v + t_{ox}}\right)$	[127]	length ℓ, radius r_v, depletion region width w_{dep}, oxide thickness t_{ox}, oxide permittivity ϵ_{ox}, silicon permittivity ϵ_{Si}	includes both C_{ox} and C_{Si}
(2.44)	$C = \frac{63.34\ell\epsilon_o}{ln(1 + 5.26\frac{\ell}{r_v})}$	[131]	length ℓ, radius r_v, permittivity of free space ϵ_o	no oxide thickness parameter, no depletion region
(2.45)	$C = \alpha\beta \frac{\epsilon_o \epsilon_{ox}}{t_{ox} + \frac{\epsilon_{ox}}{\epsilon_{Si}} w_{dep}} 2\pi r_v \ell$	[134]	length ℓ, radius r_v, depletion region width w_{dep}, silicon permittivity ϵ_{Si}, oxide permittivity ϵ_{ox}	fitting parameters α and β adjust for, respectively, distance to ground and portion of metal contributing to C
(2.46)	$C_{dep} = \frac{\epsilon_o \epsilon_{Si}}{ln\left(\frac{r_v + t_{ox} + w_{dep}}{r_v}\right)\sqrt{1 - \frac{V}{V_{th}}}} 2\pi\ell$	[135]	length ℓ, radius r_v, depletion region width w_{dep}, oxide thickness t_{ox}, silicon permittivity ϵ_{Si}, applied voltage V, threshold voltage V_{th}	
(2.47)	$C_{dep} = \frac{\epsilon_o \epsilon_{Si}}{ln\left(\frac{r_{dep}}{r_v + t_{ox}}\right)} 2\pi\ell$	[112]	length ℓ, radius r_v, oxide thickness t_{ox}, depletion region radius r_{dep}	
(2.48)	$C_{Si} = \frac{0.5\epsilon_o \epsilon_{Si} \alpha (8n_{bc} + 1)\frac{p_v}{r_v}}{ln\left(\frac{r_v + w_{bc}}{r_v}\right)} \pi\ell$	[135]	length ℓ, radius r_v, TSV-to-TSV pitch p_v, number of body contacts n_{bc}, distance to body contact w_{bc}, silicon permittivity ϵ_{Si}	includes fitting parameter α
(2.49)	$C_{Si} = \epsilon_o \epsilon_{Si} \frac{\frac{p_v}{2} + d_v + \beta}{\alpha p_v} \ell$	[157]	length ℓ, diameter d_v, TSV-to-TSV pitch p_v, scaling factors α and β, silicon permittivity ϵ_{Si}	
(2.50)	$C_{Si} = \frac{\ell\pi\epsilon_{Si}}{ln\left[\frac{p_v}{2r} + \sqrt{\left(\frac{p_v}{2r}\right)^2 - 1}\right]}$	[154]	length ℓ, radius r_v, TSV-to-TSV pitch p_v, silicon permittivity ϵ_{Si}	
(2.51)	$G_{Si} = \sigma_{Si} \frac{\frac{p_v}{2} + d_v + \beta}{\alpha p_v} \ell$	[157]	length ℓ, diameter d_v, TSV-to-TSV pitch p_v, scaling factors α and β, silicon conductivity σ_{Si}	fitting parameters are in μm
(2.52)	$G_{Si} = \frac{0.5\sigma_{Si}(8n_{bc} + 1)}{\alpha\frac{p_v}{r_v} ln\left(\frac{r_v + w_{bc}}{r_v}\right)} \pi\ell$	[135]	length ℓ, radius r_v, TSV-to-TSV pitch p_v, number of body contacts n_{bc}, distance to body contact w_{bc}, silicon conductivity σ_{Si}	includes fitting parameter α
(2.53)	$G_{Si} = \frac{\ell\pi\sigma_{Si}}{ln\left[\frac{p_v}{2r} + \sqrt{\left(\frac{p_v}{2r}\right)^2 - 1}\right]}$	[154]	length ℓ, radius r_v, TSV-to-TSV pitch p_v, silicon conductivity σ_{Si}	
(2.54)	$G_{dep} = \frac{\sigma_{Si}}{ln\left(\frac{r_v + t_{ox} + w_{dep}}{r_v}\right)\sqrt{1 - \frac{V}{V_{th}}}} 2\pi\ell$	[135]	length ℓ, radius r_v, depletion region width w_{dep}, oxide thickness t_{ox}, silicon conductivity σ_{Si}, applied voltage V, threshold voltage V_{th}	

a. Comparison of TSV resistance models

Compact models for the resistance consider two types of TSV structures. The first model characterizes the TSV as a cylinder with a constant radius, and the second model represents the TSV as a tapered cylindrical TSV where the top of the TSV is wider than the bottom. The per cent error for the DC resistance of a cylindrical TSV does not exceed 1%. The per cent error increases to a maximum of 4% for frequencies up to 10 GHz when applying (2.29) with a fitting parameter (a 7% error without the fitting parameter). A closed-form expression for a tapered cylindrical TSV at DC (listed as (2.28) in Table 2.2) produces less than 2% error. The compact model of the tapered cylindrical TSV produces less than 6% error for frequencies up to 10 GHz when a fitting parameter is utilized in (2.30) [98] .

b. Comparison of TSV inductance models

Two basic expressions are used to model the inductance of a TSV. The expressions are listed as (2.31) and (2.32) in Table 2.2 and are derived, respectively, by Goldfarb [148] and Rosa [142]. Goldfarb modified (2.32) to better match experimental data. Equation (2.31) was used by Katti *et al.* [112] to compare his model with experimental results. The per cent error ranged from 8% to 17.3% for TSVs with a length of 400 μm and a radius of 50 to 60 μm. Rosa also developed expressions for the inductance for the case where the length ℓ significantly exceeds the radius r_v which is listed as (2.34), and for the case where the skin depth dominates, see (2.35). Both [134] and [126] use (2.35) to model the high frequency inductance of a TSV. A fitting parameter is added to both (2.32) and (2.35) in [134] to

reduce the maximum error from greater than 20% to, respectively, less than 3% and 8.5%. Compact models of the DC and high frequency inductance given by, respectively, (2.32) and (2.35) are verified by an electromagnetic field solver for radii of 0.5 to 30 μm and aspect ratios ($\frac{\ell}{2r_v}$) of 0.5 to 9 [134]. Wu [126] altered (2.35) to better match experimental results with the modified expression of (2.36). Both (2.31) and (2.32) overestimate the measurements by Wu, producing errors, respectively, of 40% and 65% for an aspect ratio of 10. The errors exceed 50% and 75% for aspect ratios approaching 100. The modified expression, adding a multiplier of 0.65, reduces the maximum error to less than 25% for all aspect ratios (the radius is varied from 1 to 5 μm). An additional modified inductance expression is (2.37). Equation (2.35) is modified to better match simulated results from a 3-D/2-D quasi-static electromagnetic field solver. The inductance calculated by (2.37) is within 3% of the simulated result for lengths ℓ of 20 to 140 μm and radii r_v of 10 to 40 μm [131].

The effect of a tapered cylindrical shape on the inductance of a TSV is examined by [98]. An expression for the self-inductance is (2.33) and the mutual inductance is (2.39). Both (2.33) and (2.39) reduce to the Rosa expressions listed as, respectively, (2.32) and (2.38) as β, a measure of the tapering angle, approaches zero [98]. The tapering angle ranges from 0° to 15°, producing a maximum error of 5% at low frequencies and approximately 20% at high frequencies, which is reduced to 8% with a fitting parameter. Equations (2.38) and (2.39) depend upon the TSV-to-TSV pitch p_v. The p_v term is substituted for the radius of the TSV r_v when calculating the mutual inductance.

c. Comparison of TSV capacitance models

Several methods have been applied in the development of compact models of the TSV capacitance. These methods include 1) the worst case capacitance, the oxide capacitance C_{ox}, 2) a series representation of the oxide and depletion capacitance (C_{dep}), similar to a MOS capacitor, 3) a single expression encompassing both the oxide and depletion capacitance, 4) a series combination of the oxide and silicon capacitance, where a parallel plate capacitor is assumed to be formed between two TSVs in silicon [157], and 5) a series combination of the oxide, depletion, and silicon capacitance (C_{Si}) [135]. The oxide capacitance is represented by (2.40) - (2.42). Most research groups use (2.41) to model the oxide capacitance of a TSV, as (2.41) more accurately models the physical dimensions of a TSV and reduces the error as compared to simulated results from approximately 10% when using (2.40) to less than 0.5% [112].

The oxide capacitance is one component of the TSV capacitance; therefore, any model that ignores the depletion region and, to a certain extent, the silicon capacitance will overestimate the TSV capacitance. The depletion capacitance is determined by Salah *et al.* [135] and Katti *et al.* [112] and is listed, respectively, as (2.46) and (2.47) in Table 2.2. Katti does not provide simulation results for the depletion region capacitance in isolation since the oxide capacitance is also present, but does provide a comparison of the oxide and depletion capacitance between the analytic model and Raphael simulations [112]. A maximum error of 3.5% is reported for a TSV that includes a depletion region of the maximum radius r_{dep}. Additional TSV parameters include radii of 1 and 2.5 μm, and, respectively, an oxide

thickness of 50 nm and 120 nm. The length is set to 20 μm for both radii. Katti also compares the measured TSV capacitance, but with simulated geometries rather than based on a closed-form expression. However, an analysis of the oxide capacitance, based on (2.41), and the minimum TSV capacitance, using (2.47) (when both the oxide and maximum depletion capacitance are considered) as a function of the TSV diameter and oxide thickness is provided by Katti *et al.* [112]. The depletion capacitance developed by Salah *et al.* [135] is verified by electromagnetic field solvers (Q3D Extractor) by comparing the S parameters produced by the field solver with those produced by a lumped element model. The lumped element model includes the capacitance for the oxide (2.41), the depletion region (2.46), and the silicon capacitance (2.48). The silicon capacitance includes a fitting parameter α that ranges from 1.5 to 3.5, and is adjusted based on the number of surrounding TSVs. The lumped model also includes a compact model of the silicon and depletion region conductance, in addition to the resistance and inductance of the TSV [135]. Expressions for the silicon capacitance and conductance developed by Salah *et al.* also include parameters that account for the number of body contacts present in the silicon n_{bc} and the distance to the body contacts w_{bc}. The per cent difference between the S_{21} parameter results produced by the lumped model and simulation does not exceed 6%. S_{21} parameters for both the model and the simulation have been extracted for a frequency range of 1 to 10 GHz.

In addition to the silicon capacitance derived by Salah *et al.* [135], Kim *et al.* developed a compact model with the assumption that a parallel plate capacitor is formed between two TSVs in silicon [157]. The silicon capacitance model is listed as (2.49) in Table 2.2. The

capacitance model of the TSV includes the oxide capacitance (2.40) and the silicon capacitance (2.49). The compact model is verified by comparing S_{21} parameter values, similar to Salah *et al.* [135]. The lumped model of the TSV includes the resistance, inductance, and conductance of the TSV. The per cent difference between the model and the 3-D field solver ranges from 2 to 15% for frequencies up to 20 GHz. A comparison is made for TSV radii of 5, 10, and 15 μm, distance between TSVs of 100, 120, and 140 μm, and oxide thickness of 500, 700, and 900 nm. The greatest error occurs at the lower frequency (up to 3 GHz) where the large TSV capacitance dominates, demonstrating that the reduction in TSV capacitance due to the depletion capacitance is negligible.

Two expressions listed in Table 2.2 produce a total capacitance of the TSV by either empirically fitting numerical values to match simulation or by combining the oxide and depletion capacitance into a single expression (which includes the fitting parameters). These expressions are listed, respectively, as (2.44) and (2.45). Equation (2.43), derived by Xu *et al.* [127], is a series combination of (2.41) and (2.47). Both (2.44) and (2.45) produce errors of less than 8% as compared to results from electromagnetic field solvers. The TSV capacitance derived by Weerasekera *et al.* [131] in (2.44) is compared with simulations for a range of physical parameters describing a TSV. The parameters include TSV lengths ℓ of 20 to 140 μm and radii r_v of 10 to 40 μm. In comparison, Savidis *et al.* [134] examines radii of 10, 20, and 30 μm with aspect ratios ($\frac{\ell}{2r_v}$) ranging from 0.5 to 9. The primary difference between these two expressions is that (2.44) does not consider the depletion region, and therefore also ignores the oxide thickness. The 8% error reported in [131] does, however, include inaccuracies due to neglecting the oxide for thicknesses up to 1 μm. The fitting

parameters α and β are further discussed in Chapter 5 as (2.45) is the author's own work.

d. Comparison of TSV conductance models

The conductance of the silicon and depletion region is also listed in Table 2.2. The closed-form expression for the conductance of silicon is the same as the capacitance, but the permittivity of silicon is replaced by the conductivity. Equations (2.51) and (2.52) for the conductance of silicon are, respectively, by Kim *et al.* [157] and Salah *et al.* [135] for a lumped model representation of the TSV. The conductance model is not individually verified with electromagnetic simulation, but the S_{21} extracted from simulation is compared with the lumped model of the TSV. Note that the lumped model by Salah *et al.* also includes the depletion region conductance (as well as the capacitance), which is listed as (2.54) in Table 2.2.

A textbook definition of the silicon capacitance and conductance as applied to electromagnetic fields are listed, respectively, as (2.50) and (2.54) [154]. These two expressions are provided for comparison with the other capacitance and conductance expressions for silicon and other TSV models.

2.4 Summary

The electrical characterization of a wire or TSV requires field solvers to accurately determine the effects of various parameters including the physical structure and dimensions, frequency, and surrounding environment. These field solvers, however, require significant computational time.

Compact models provide an efficient and fairly accurate method for characterizing the electrical properties of a conductor. Compact models are often verified through field solvers and experimental data. Once verified, these models can be used to replace full wave simulations to determine the impedance of a TSV. An understanding of the physical phenomena is necessary to develop closed-form expressions of the resistance, inductance, capacitance, and conductance of a TSV. The complexity of these TSV models is dependent upon additional effects such as frequency dependencies (both proximity and skin effects), the depletion region, and temperature. The primary objective of these closed-form expressions for characterizing the TSV impedance is to provide a *RLCG* electrical representation of a TSV for circuit simulation and analysis. The electrical parameters of a TSV are reduced to a lump *RLCG* model for inclusion in circuit simulation.

Expressions for the resistance consider a cylindrical TSV, as well as tapered cylindrical vias, where the diameter of the TSV is reduced farther into the silicon. Models are provided for both DC and high frequencies, where the skin effect increases the resistance of a TSV. Similar to the resistance, closed-form expressions for the inductance are presented for both a cylindrical and tapered cylindrical TSV. Inductance expressions for the self- and mutual inductance at DC and high frequencies are described. Most expressions characterizing the TSV inductance are based on work completed by Rosa in 1908. Many capacitance models of the TSV have been developed as compared to either resistance or inductance. The capacitance expressions are of varying accuracy based on the number and type of physical effects that are considered. Closed-form expressions of the finite conductance of the silicon surrounding the TSV have also been developed.

Chapter 3

3-D Power Delivery: Generation and Distribution

Power delivery is a critical component in modern high performance integrated circuits. Any time a computer or mobile device such as a tablet or smartphone runs an application, hundreds of millions of transistors switch over a billion times per second, each drawing a small amount of current. The aggregate sum of these small currents requires high on-chip current. The power delivery network must supply this large current while satisfying the noise specifications of the IC. When an IC transitions from an idle state to an active state, there is a sudden draw of current that leads to a voltage drop on the power network. In addition, the large inductive component of the power network impedance leads to a voltage dependent change in current with respect to time ($L \cdot di/dt$ noise). Variations in the power supply current behave as noise, producing uncertainty in the delay of the clock and data signals. The effect of the supply noise on circuit operation is a lower clock frequency, and potential damage to the oxide of the MOS transistors [141, 161, 162].

The target application is typically high speed computing, where the dissipated power can exceed 100 watts, or mobile devices where battery lifetime is critical. In either case,

efficient power delivery with minimal noise due to *IR* drops and/or *L·di/dt* noise is necessary to operate correctly. Models of the power delivery network that properly account for noise generation and propagation are therefore critical.

A description of power delivery for 3-D integrated circuits is provided in this chapter. An overview of basic concepts that apply to both 2-D and 3-D power delivery is presented in Section 3.1. Recent research on modeling power delivery networks in 3-D ICs is discussed in Section 3.2, including work on noise modeling and the optimal placement of TSVs and decoupling capacitors within 3-D power networks. Compact models of the noise within a 3-D power delivery network are provided in Section 3.3. Research on 3-D power regulator techniques and topologies are described in Section 3.4.2. Some concluding remarks are offered in Section 3.5.

3.1 Basic concepts of power delivery

The current demand and noise requirements of a 3-D integrated circuit typically exceeds that of a 2-D IC for a given footprint. The potential to stack multiple device planes leads to high current densities, and therefore greater noise fluctuations. Although there are fundamental differences in the design of a 3-D power delivery system as compared to a 2-D power delivery system, significant overlap in basic concepts and principles exists. The power and ground lines are typically paired, resulting in an increase in capacitive coupling between each pair. An on-chip power delivery system is typically composed of a global and local power distribution network. The global power network distributes the supply current across the die and is typically placed on the higher, wider, and thicker on-chip metal levels

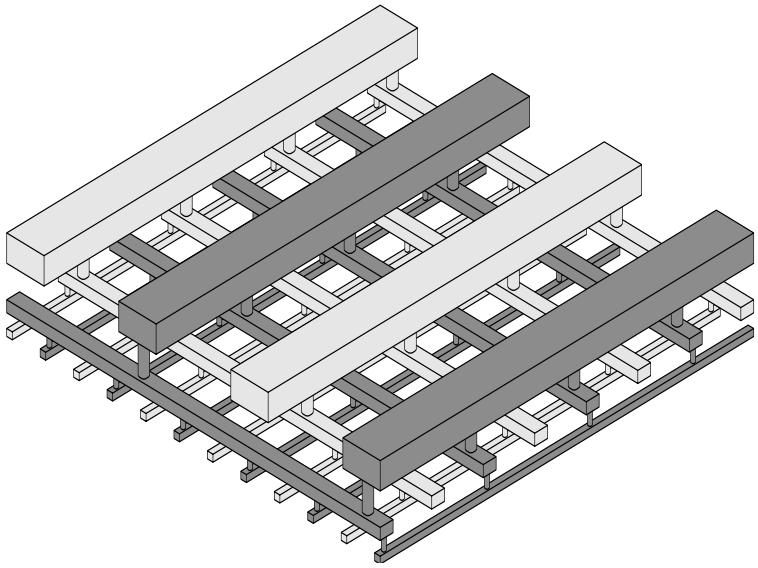


Figure 3.1: Interdigitated power/ground pair in global and local power distribution network [161].

(lower resistance, higher inductance lines). The local power distribution network delivers current from the global network to the active devices using thinner lines with a smaller pitch to increase the interconnect density for routing signals between devices (higher resistance, lower inductance lines). The paired power and ground lines as well as the hierarchical placement of the global and local power distribution networks are illustrated in Figure 3.1.

A power delivery network is composed of off-chip as well as on-chip components. One similarity between the 2-D and 3-D IC power delivery process is the method of delivering on-chip current from an off-chip source. High performance ICs are typically flip-chip bonded to reduce the parasitic impedance of the electrical paths connected to the die [163, 164]. Parasitic impedances are reduced by spreading the power and ground I/O pads along one surface of the die and directly bonding to the package pads, as shown in Figure 3.2.

The power and ground pads of an integrated circuit are connected to the global power

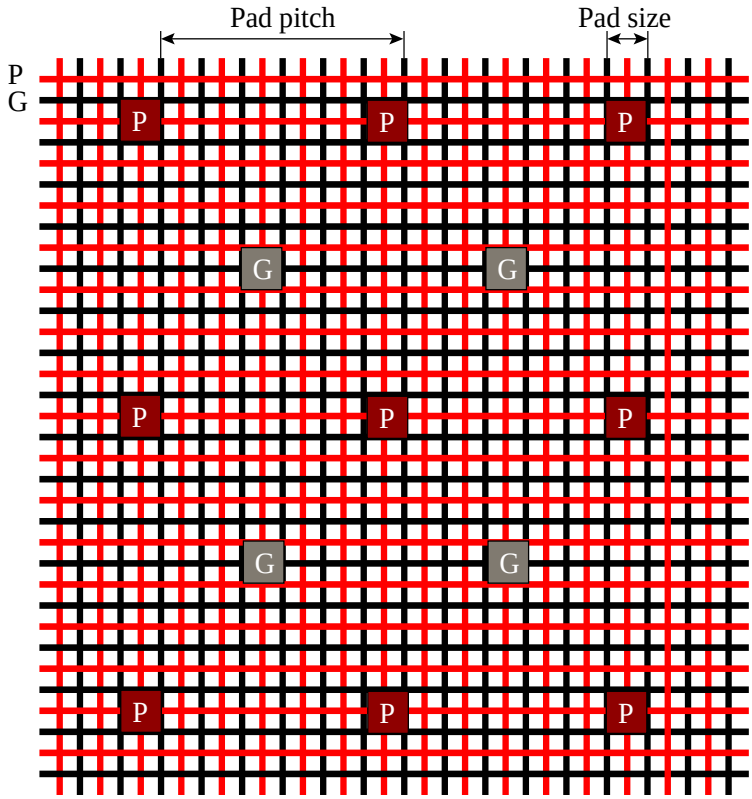


Figure 3.2: On-chip power and ground grid with I/O pads for flip-chip packaging. [80].

distribution network through on-chip vias. A global power distribution network with power and ground pads at intersection points is illustrated in Figure 3.2. Current flows from the package through the bond pads into the global power network on the integrated circuit before passing to the local power network and the on-chip devices. The current returns to the package through the ground interconnects (first through the local, and then the global network) and ground I/O pads. Note that for 3-D integrated circuits, the power/ground I/O pads are only located on one side of the stacked die, similar to a 2-D IC. A limit therefore exists on the total I/O bandwidth and pads available for power and ground, which in turn limits the total number of device planes that can be stacked.

An additional similarity between 2-D and 3-D integrated circuits is the model of the

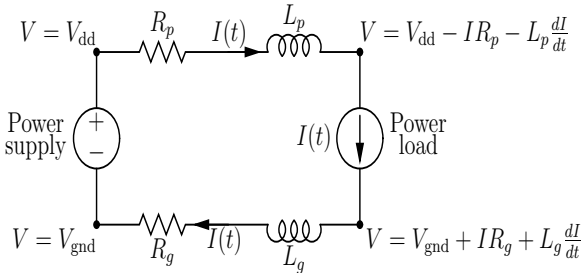


Figure 3.3: Representation of a power distribution network with a voltage source, current load, and characteristic impedance [141].

power distribution network. A simple model of a power delivery network including the voltage source, current loads, and interconnect between the source and load is shown in Figure 3.3. An ideal voltage source represents the power supply which provides a nominal power V_{dd} and ground V_{gnd} voltage. The load is modeled as a variable current source $I(t)$ and represents the current demand of an IC [141,161,162]. The model of the resistance and inductance for 3-D ICs depends on the number of device planes, process parameters, and the electrical impedance model of the TSV. Resistive voltage drops $\Delta V_R = IR$ and inductive voltage spikes $\Delta V_L = L\frac{dI}{dt}$ form across the interconnect lines as current is drawn from the power supply. The voltage levels at the load terminal are therefore affected by the noise generated by the impedance of the interconnect, device planes, and TSVs. The nominal voltage at the power terminal is reduced to $V_{dd} - IR_p - L_p\frac{dI}{dt}$ and the ground terminal is increased to $V_{gnd} + IR_g + L_g\frac{dI}{dt}$ [161,162].

3.2 Models of 3-D power distribution networks

Although many aspects of 3-D IC power delivery are similar to 2-D power delivery, there are significant differences. The most significant difference is simply the vertical bonding of multiple device planes. By stacking dies vertically, delivering current to each device plane creates both unique design challenges and opportunities for 3-D integration; 1) placement of TSVs on the power network to more effectively deliver current (while satisfying noise specifications) while not accruing a large area penalty, 2) potential to integrate power distribution topologies not available in 2-D ICs, 3) novel design techniques and methodologies for delivering power in 3-D ICs, and 4) development of EDA tools to accurately design and synthesize 3-D power networks. Although integration of disparate technologies provides unique systems level opportunities, challenges for properly delivering power to each of the stacked dies are further exasperated when integrating completely disparate technologies. Accurate and computationally efficient models and algorithms are therefore necessary to enhance the design of 3-D power networks.

Power grid analysis requires an accurate representation of various components of the power network. The model of the power network must consider the impedance of the power/ground grids, load locations, placement of the decoupling capacitors, and additionally, for 3-D circuits, the number and location of the TSVs, and the different process parameters encountered within the heterogeneous device planes. The number and location of TSVs depend on the current demand of each region within a 3-D IC. In addition, for those regions in a 3-D IC requiring high currents, thermal TSVs (discussed in Chapter 4) are necessary to reduce the thermal profile of the circuits, which in turn reduces the active

area available for placing circuit blocks.

Electrical models of a power delivery network for 3-D integrated circuits have been proposed. These models incorporate a variety of characteristics that affect 3-D power distribution networks. The electrical characteristics include the placement of decoupling capacitors [165, 166], placement of TSVs [167, 168], representation of the power and ground grids in three dimensions as either distributed or lumped, and the effect of current paths unique to TSV based 3-D circuits [169]. Physical characteristics included in some models are the TSV fabrication process (via-first, via-middle, or via-last) that affects the electrical characteristics of the power network [170], and the effect of the substrate [166, 170]. A comparison of several models of 3-D power distribution networks is provided in Table 3.1.

A distributed power and ground grid is used in the various models listed in Table 3.1. One difference among the models is the representation of the power network impedance as a simple resistive network, or as a resistive and inductive network. The inductance is not included in these models when only the DC *IR* drop of a power network is investigated [167, 171]. In general, the capacitance is included in the model when decoupling capacitors are placed, and ignored otherwise. One particular study examined the time and frequency domain characteristics of a 3-D power delivery network as a distributed impedance [172]. An *RLC* model is used to determine the noise behavior within a 3-D power network. A lumped model including the interplane connections is used to investigate the global effects of a 3-D power network. A frequency domain analysis is applied to determine the global and local resonances, while a time domain analysis predicts the worst case power noise [172].

Table 3.1: Comparison of 3-D power distribution network models.

Reference	Contribution	Results
[172]	• on-chip distributed model of 3-D power grids • lumped model proposed based on distributed model • Rents Rule power model for current modelling	• frequency and time domain analysis of power grids • global and local resonance modeled, worst-case supply noise determined • Rogue Wave concept introduced to explain spatial and temporal noise response
[170]	• comparison of power networks for via-first, via-middle, and via-last fabrication process • analyzed power delivery network for a 32-nm nine plane process-memory system • electrical model of 3-D power distribution network that includes the power/ground TSVs, power network within each device plane, substrate, and switching current	• area overhead of is ~9% for via-first process and < 2% for via-middle and via-last • Power network for via-first TSVs is typically overdamped, reducing the resonance of the network • via-last process produces underdamped power network that is sensitive to resonance and exhibits sensitivity to number of TSVs and decoupling capacitors
[169]	• investigate multiple current paths in 3-D circuits • examine *IR* reduction due to low impedance paths within 3-D circuit	• reduce number of intraplane vias by 22% or 25% reduction in decoupling capacitance • global and local resonance modeled, worst-case supply noise determined
[175]	• estimates static (*IR*) and dynamic ($L\frac{di}{dt}$) power grid noise	• less than 1% error between algorithm and SPICE simulation of interdigitated power grid with four 1 V power supplies and three 30 mA current loads (w = 10 μm, t = 2μm, pitch = 20μm)
[176]	• frequency domain model of network includes on-chip voltage regulators and effect of inductance • parallel computing algorithm compared against HSPICE simulations for power grid with 32×119 nodes and 5 to 15 stacked die in 10×10 mm^2 area footprint	• computational time is reduced by 22 times for single core and 430 times for 200 cores • at low frequency, high TSV density results in less output impedance • at high frequency, high TSV density has large impedance peaks
[166]	• 3-D power network model includes on-chip decoupling capacitors and silicon substrate • lumped *RLCG* model based on segmentation method and scalable equations	• model validated up to 20 GHz with 3-D EM simulation • analyze impact of decap position and silicon conductivity on impedance of 3-D power network
[165]	• decoupling capacitor allocation algorithm uses 3-D congestion analysis, linearized congestion models, and noise models based on adjoint sensitivity analysis to allocate CMOS and MIM decaps	• simulated results on four device planes with six tiers of distributed power grids • combination of MIM and CMOS capacitors necessary to eliminate voltage violations
[167]	• efficient estimation of power delivery requirements for supply noise limits • four algorithms to find minimum number of TSVs for target *IR* drops • trace selection technique uses relevant portion of power trace to represent worst load	• algorithms estimate number of TSVs for block-level model of 3-D IC • number of traces reduced by 51×, reducing SPICE simulation to 1 hour on a 40 node cluster • greedy algorithm that increments number of TSVs provides minimal solution
[177]	• analytical physical model derived to incorporate power supply noise	• model has less than 4% error as compared to SPICE simulations • design guidelines to reduce power supply noise through decap and TSV insertion
[178]	• model estimates peak-to-peak switching noise in 3-D stack of ICs • vary physical dimensions of TSVs and value of decoupling capacitors	• model is 2 to 3% of values produced by Ansoft Nexxim4.1 equivalent model • computationally 3 to 4× as fast as Nexxim4.1 with 2× less memory
[171]	• iterative DC *IR* drop solver (in Matlab) for 3-D power delivery network • non-uniform grid with location dependent resistivity is accounted for in a finite volume formulation • conjugate gradient (CG) method with diagonal pre-conditioner used to solve system equation $Yx = b$	• simulation of a 5 mm by 5 mm 3-D IC with four stacked dies compared with • solves for over 10 million unknowns (node resistivities, voltages, impedances) • computational time ranges from 590 to 1060 minutes on 3 GB memory

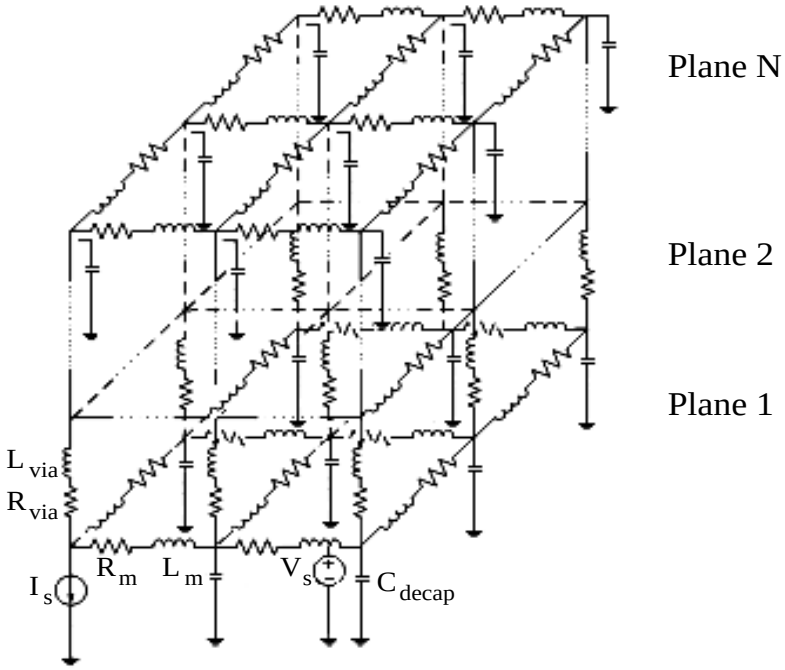


Figure 3.4: Equivalent model of an N-plane power delivery network for a 3-D integrated circuit [176].

A model of an N-plane power distribution network is shown in Figure 3.4. The model includes both the resistive and inductive components of the power/ground grids. The inductance and resistance of the TSVs is also included. The capacitance of the TSV is ignored under the assumption that the decoupling capacitors are much larger than the cumulative sum of the TSV capacitances. Note that the TSV capacitance is not always ignored in models of the power distribution network. A single current load I_s is shown in the figure, reducing the computational complexity of the model. Additional current loads are included across device planes to better match the current demands of the various circuits. Single current sources can be placed at each node of a power grid depending upon the current demands of each circuit block [172–174].

The distributed model described in [172] is reduced to a lump model representation of each device plane, similar to that depicted in Figure 3.5. A lump model includes the impedance of the TSV, series *RL* of a line, and *C* connected to the substrate. A lumped model of each device plane represents a two-port network. One port is formed by merging all of the power (or ground) connections on the redistribution layer (RDL), and the second port is formed by combining all of the power (or ground) nodes on metal 1. The model consists of passive *RLC* elements and current sources (loads) at each of the ports. The passive elements are determined by removing the current sources and solving for the impedance [172],

$$\begin{bmatrix} V_1(s) \\ V_2(s) \end{bmatrix} = \begin{bmatrix} Z_{11}(s)\, Z_{12}(s) \\ Z_{21}(s)\, Z_{22}(s) \end{bmatrix} \begin{bmatrix} I_1(s) \\ I_2(s) \end{bmatrix} , \tag{3.1}$$

where $Z_{12} = Z_{21}$. The impedance matrix is solved for Z_{11}, Z_{22}, and Z_{21}, and the values of the *RLC* components are individually determined.

A lump model of the 3-D power distribution network is used to examine the resonance phenomena for a 500 μm by 500 μm power distribution network composed of two device planes. A multi-tiered power grid is constructed on each device plane using metals 1, 3, 6 and an RDL [172]. The two device planes included decoupling capacitors, 5 nF/mm^2 and 50 nF/mm^2, placed on, respectively, plane one and two. The result is that at "mid-frequencies," the resonant peak of the second device plane is not present. The large capacitance of the second device plane effectively places the metal one resistance of the second plane in series with the effective series resistance of the package inductance.

As previously noted, the power or ground grid is modeled with a resistive network

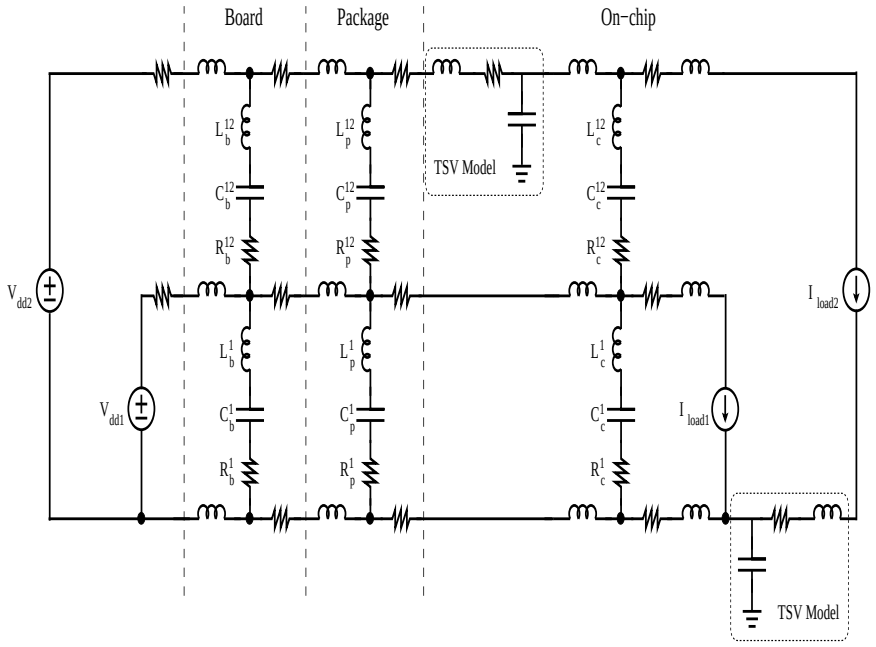


Figure 3.5: Lumped model of a power delivery network for two device planes including the impedance of the TSVs.

when a static DC analysis (*IR* drop) is desired. A closed-form expression for the effective resistance between any two nodes in a power distribution network is [175, 179, 180]

$$R_{eff} = \frac{1}{2\pi} ln\left(m^2 + n^2\right) + 0.51469, \tag{3.2}$$

where

$$m = |x_1 - x_2|, \tag{3.3}$$

and

$$n = |y_1 - y_2|. \tag{3.4}$$

A general algorithm that uses the effective resistance to determine the *IR* drop in a uniform power supply grid with multiple non-uniformly distributed voltage sources and switching current loads has been developed [179, 181, 182]. The closed-form expressions in [182] estimate the *IR* voltage drop within 1% of results produced by SPICE simulation.

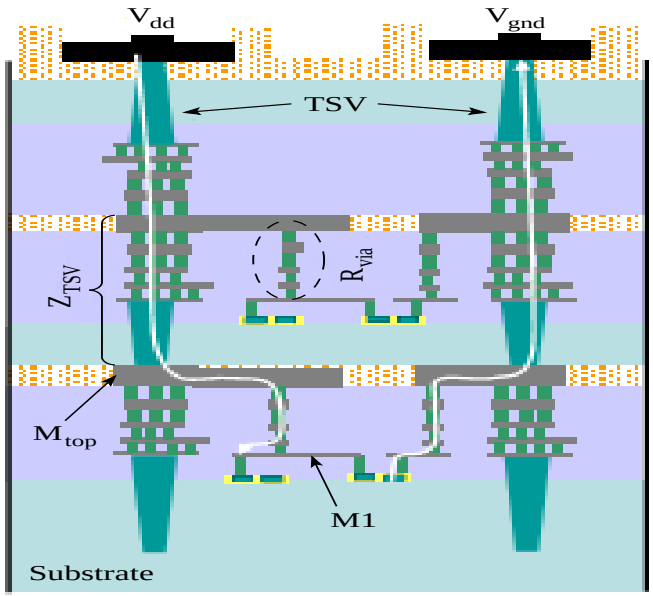
3.2.1 Multiple current paths in stacked dies

Additional work on 3-D power delivery models has explored different power distribution topologies and the effect of the TSV geometry on the 3-D power delivery system [169, 177, 183]. The presence of alternative low impedance current paths in 3-D ICs has been explored. For the example described in [169], these alternative current paths produce either a 22% reduction in the number of intraplane vias or a 25% reduction in the required decoupling capacitance. Two scenarios are considered in [169]: 1) current flow from the TSV and into the on-die topmost metal layer, and 2) current flow from the TSV into both the top metal layer and the first metal layer. The single and multi-current paths are shown in Figure 3.6, and an equivalent electrical model of the two scenarios is depicted in Figure 3.7.

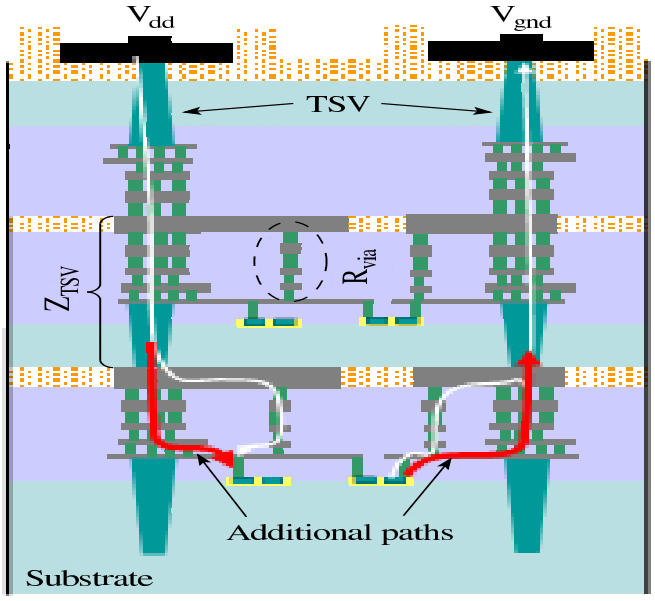
The maximum current of the single and double path models shown in Figure 3.7 for a maximum allowed voltage ripple of 5% of the nominal power supply is, respectively,

$$I_L \leq \frac{(1 - V_{ripple})V_{dd}}{Z_{TSV} + R_{MT} + R_{via2}}, \tag{3.5}$$

and



(a)



(b)

Figure 3.6: Current paths where (a) the TSV is connected to the top metal M_{top} (single path), and (b) the TSV is connected to both the top metal M_{top} and metal 1 M_1 (double path) [169].

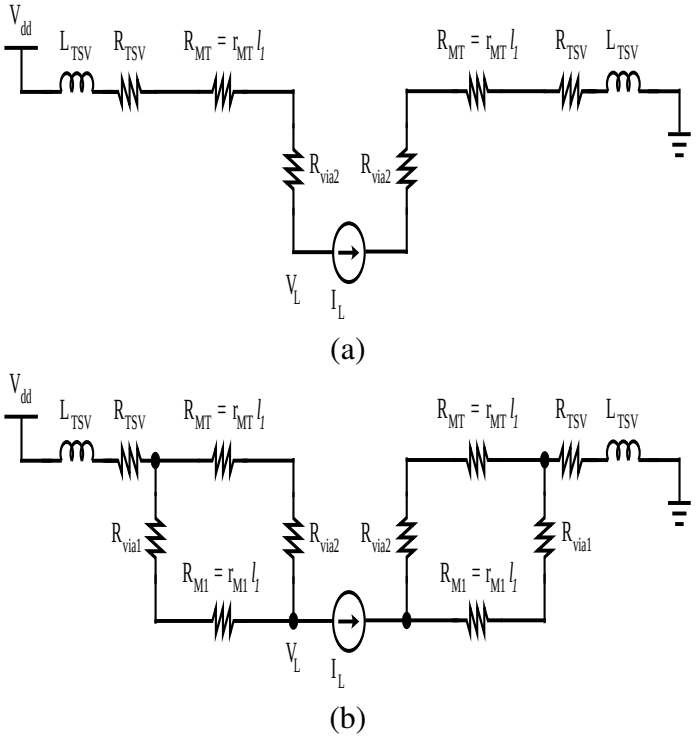


Figure 3.7: Equivalent model of the current paths shown in Figure 3.6. Current path through (a) top metal M_{top} only (single path), and (b) top metal M_{top} and metal 1 M_1 (double path) [169].

$$I_L \leq \frac{(1 - V_{ripple})V_{dd}}{Z_{TSV} + (R_{MT} + R_{via1}) \parallel (R_{M1} + R_{via2})}, \tag{3.6}$$

where Z_{TSV} is the impedance of the TSV, R_{via1} is the stack of vias included as part of the TSV, and R_{via2} is the impedance of a stack of vias connecting the topmost and lowest metal layers. The resistance of the topmost metal layer is R_{MT}, which is the product of the per unit length top metal resistance multiplied with the path length $r_{MT}l$. Similarly, the resistance of the bottom metal (metal 1) is R_{M1}, the product of the per unit length metal 1 resistance multiplied with the path length $r_{M1}l$. A conclusion of this study is that the path with TSVs requires fewer intra-plane vias as the structure of the TSV itself provides an additional current path to the load.

3.2.2 Decoupling capacitor placement for 3-D ICs

A model of the 3-D power distribution network also requires information describing the placement of the decoupling capacitors. Recent research on 3-D IC power deliver has focused on the effectiveness of the decoupling capacitor placement on power grid noise [165, 166]. The effects of the decoupling capacitor and silicon substrate on the 3-D IC power distribution network are analyzed in [165, 166], and these results have been incorporated into models of a 3-D power delivery network.

A combination of an *RLGC* lumped model based on the segmentation method [184], scalable equations derived from the physical configuration, and models of the on-chip decoupling capacitors is validated by 3-D electromagnetic (EM) simulation up to 20 GHz

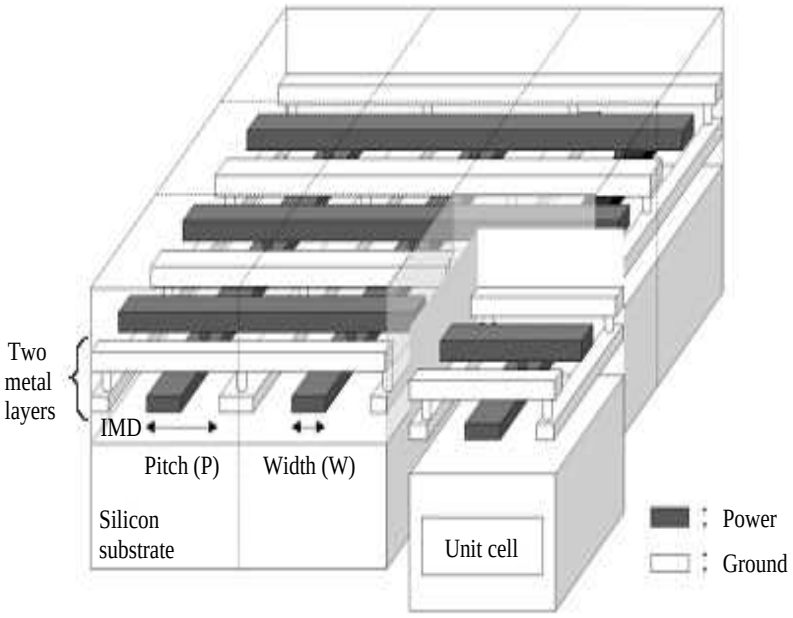


Figure 3.8: Segmentation of the power distribution network into unit cells. [166].

using CTS Microwave Studio (CTS MWS) [166]. The segmentation method [185] reconstructs the on-chip power distribution network from a unit cell. The unit cell decomposes the power distribution network into identical sub-blocks based on the physical parameters of the power network [166]. The power network examined in [166] includes two metal levels of interdigitated power and ground lines over a silicon substrate. The two metal level power delivery network and a unit cell of the network is illustrated in Figure 3.8. The effect of the inter-metal dielectric is included in the capacitance term of each unit cell. Face-to-face ports of adjacent unit cells are connected to reconstruct the power delivery network using the segmentation method [185]. The compact model reduces the computational time required to estimate the impedance of a power network.

The reconstructed power distribution network model is for a single device plane. An illustration of the unit cell decomposed into finite metal partitions is shown in Figure 3.9(a).

The electrical equivalent of each section of the unit cell (shown in Figure 3.9(b)) is depicted in Figure 3.9(c) and includes the effect of the substrate on the electrical properties of each section. The capacitance and conductance of the substrate (C_{sub} and G_{sub}) and the capacitance of the power and ground lines to silicon (C_{pSi} and C_{gSi}) are included in Figure 3.9(c). Silicon conductivities of 5, 10, 15, and 50 S/m are investigated by modifying the capacitance and conductance of the silicon within the model. The *RLGC* lumped model of a specific section is extracted from the structure by first determining the air-filled capacitance, assuming the structure is a coplanar waveguide. The air-filled capacitance replaces the waveguide structure with air where dielectric was once present. The inductance of a section is extracted from the air-filled capacitance using

$$L_{A_x} = \frac{1}{c^2 C'^{air}_{A_x}}, \tag{3.7}$$

where c is the speed of light in air, 3×10^8 m/s. The air filled capacitance is

$$C'^{air}_{A_x} = C'^{air}_1 + C'^{air}_3, \tag{3.8}$$

where C_1 and C_3 are

$$C'^{air}_i = 2\epsilon_o \frac{K(k_i)}{K(k'_i)}. \tag{3.9}$$

Each section of the power distribution network is vertically partitioned into three subsections. The first partition is composed of the substrate with height H_1, the second partition

includes the area from the substrate to metal 1 with height H_2, and the third partition includes the area from metal 1 to metal 2 with height H_3. The capacitance, given by (3.8), includes the first and third subsections.

An electrical model of each section includes the conductance G_{sub} and capacitance C_{sub} of the substrate. The substrate capacitance is

$$C_{sub} = 2\epsilon_{r1}\epsilon_o \frac{K(k_1)}{K(k_1')}, \epsilon_{r1} = \epsilon_{r,Si}. \tag{3.10}$$

The capacitance between the power and ground lines is

$$C_{pg(A_x)} = \epsilon_{r3}\epsilon_o C_3^{'air}, \epsilon_{r3} = \epsilon_{r,IMD}. \tag{3.11}$$

where $\epsilon_{r,IMD}$ is the permittivity of the intermetal dielectric and $C_3^{'air}$ is the air-filled capacitance of the third partition. The model also includes the capacitance describing the penetration of the electrical field lines into the silicon from the power and ground lines, respectively, by C_{pSi} and C_{gSi}. The expression accounting for the coupling between the power and ground lines with silicon is

$$C_{pSi\&gSi} = \epsilon_{r2}\epsilon_o \left[\left(\frac{W}{H_2}\right) + 0.77 + \left(\frac{W}{H_2}\right)^{0.25} + 1.06\left(\frac{T_1}{H_2}\right)^{0.5}, \epsilon_{r2} = \epsilon_{r,IMD} \right], \tag{3.12}$$

where W is the width of the line, T_1 is the thickness of metal 1, H_2 is the distance between the silicon and metal 1, and ϵ_{r2} is the permittivity of the intermetal dielectric.

An expression for the resistance R_{A_x} of each section is included in the model. The

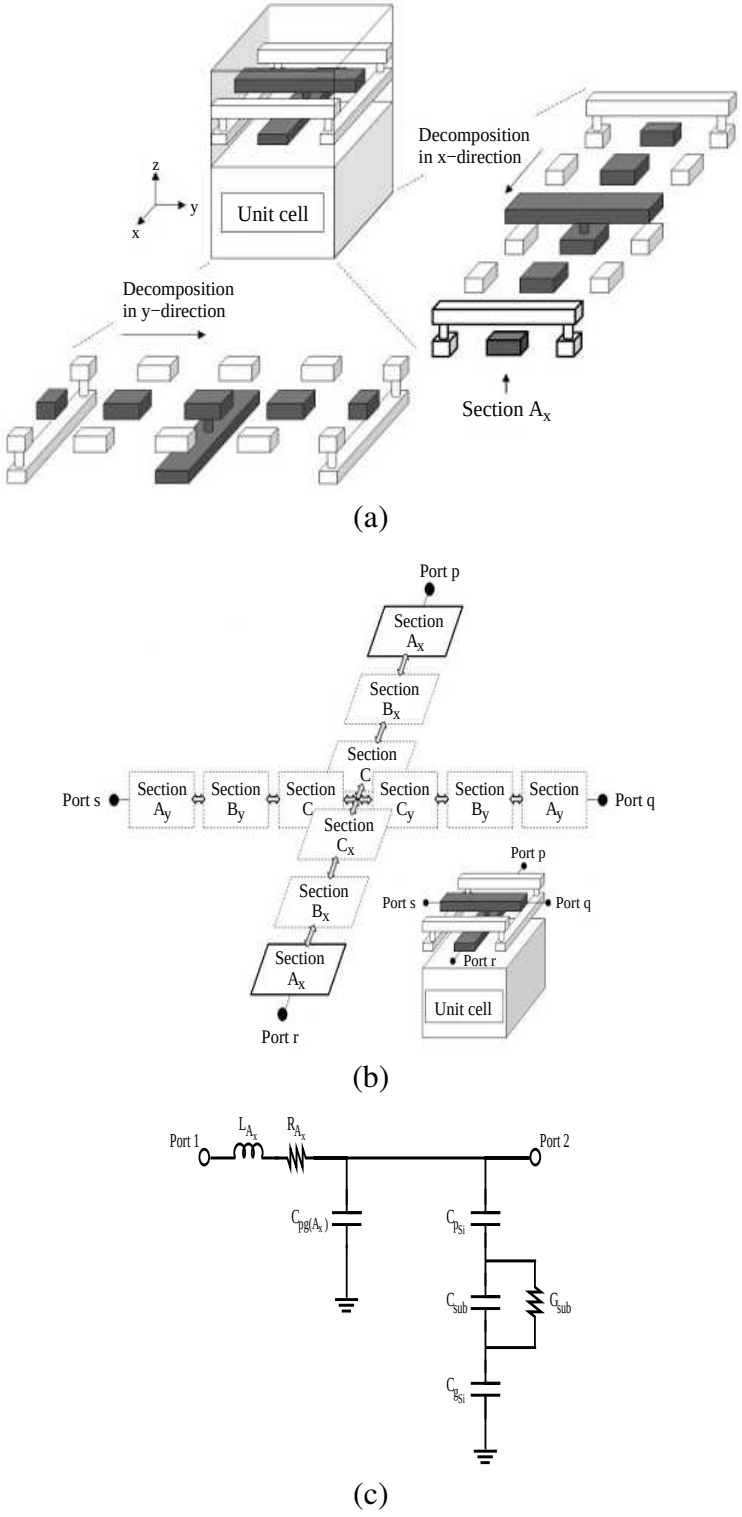


Figure 3.9: Unit cell of a power distribution network as a (a) decomposition in the x- and y-directions, (b) schematic representation of connected sections, and (c) electrical model of each section [166].

resistance of a section of a line is

$$R_{A_x} = \frac{1}{2}R_{M1} + \left[\frac{1}{2}R_{M1}\right] \| \left[\frac{1}{2}R_{M2}\right]. \tag{3.13}$$

An expression for the resistance of a section of a power distribution network is determined from the sheet resistances R_{M1} and R_{M2}, respectively, of metal 1 and metal 2. The width of metal W is also used to determine the resistance.

The power distribution network includes decoupling capacitors, often implemented as MOS devices. A model of a MOS decoupling capacitor including the equivalent series resistance ESR is shown in Figure 3.10. The on-chip decoupling capacitor is connected to a unit cell of the power distribution network (see Port N in Figure 3.10) [166]. The physical and electrical parameters of a single decoupling capacitor are listed in Table 3.2. Larger decoupling capacitors are multiples of a single capacitor.

Table 3.2: Physical and electrical parameters of the on-chip decoupling capacitor unit cell [166].

Parameter	Value
L	2 μm
W	85 μm
t_{ox}	7.3 nm
C_{ox}	0.815 pF/decap
ESR	17.65 Ω/decap

A model of a single plane power distribution network that includes the effects of the decoupling capacitors and the substrate has been verified with experimental data. A test circuit has been fabricated within decoupling capacitors embedded within the on-chip power distribution network. Measurements of the power distribution network impedance range from 300 KHz to 26.5 GHz. Less than 10% error is observed across all frequencies in the

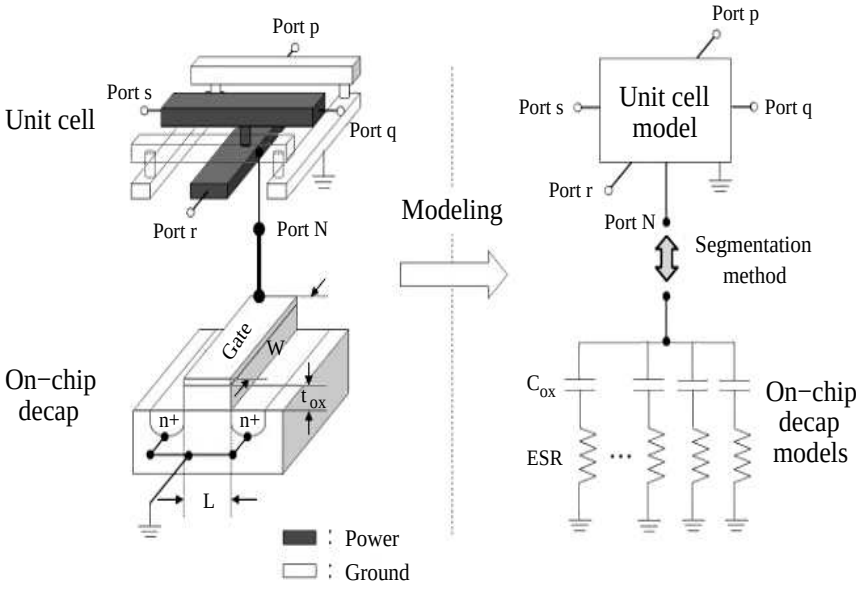


Figure 3.10: Lumped *RC* model of the decoupling capacitor unit cell connected to a unit cell within the power distribution network [166].

range of 0.1 GHz to 20 GHz.

A more complex model that includes the 3-D integrated physical device planes is described in [166]. This model includes the impedance of a TSV, three different device planes (each modeled as a single plane network), and different types of decoupling capacitors. The physical parameters of the TSV are a 50 μm length, 40 μm diameter, and TSV-to-TSV pitch of 100 μm. The impedance of the TSV is determined from the closed-form expressions described in [112]. An illustration of the stacked 3-D integrated circuit including the power and ground TSV pairs is shown in Figure 3.11. The segmentation method is applied to determine the location of the TSVs within the power distribution networks between device planes.

Two sets of investigations are described [166]. The first investigation examines the

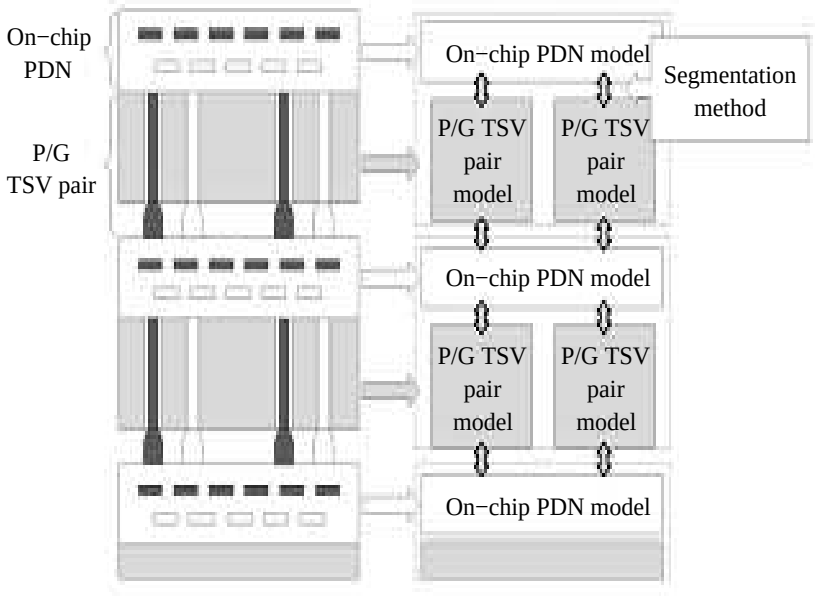


Figure 3.11: 3-D stacked power distribution network including pairs of power and ground TSVs. [166].

effect of the size of the decoupling capacitance on the impedance of the 3-D IC power distribution network, while the second study investigates the effects of decoupling capacitor location on the impedance of the power network. The three test cases examine the effects of the size of the decoupling capacitor, as shown in Figure 3.12(a). Three additional test cases explore the effects of decoupling capacitor placement, as shown in Figure 3.12(b) [166].

The value of the three on-chip capacitance in the experiment are 0.29 nF, 1.17 nF, and 2.35 nF. The extracted intrinsic capacitance of the power distribution network is 31.83 pF. The 1.17 and 2.35 nF capacitors produce similar network impedances for frequencies up to 20 GHz. The impedance of the power network is less than 5 Ω up to 2 GHz before increasing due to the loop resistance of the power network. The loop resistance of the power distribution network is approximately 5 Ω. The 0.29 nF capacitor significantly reduces the

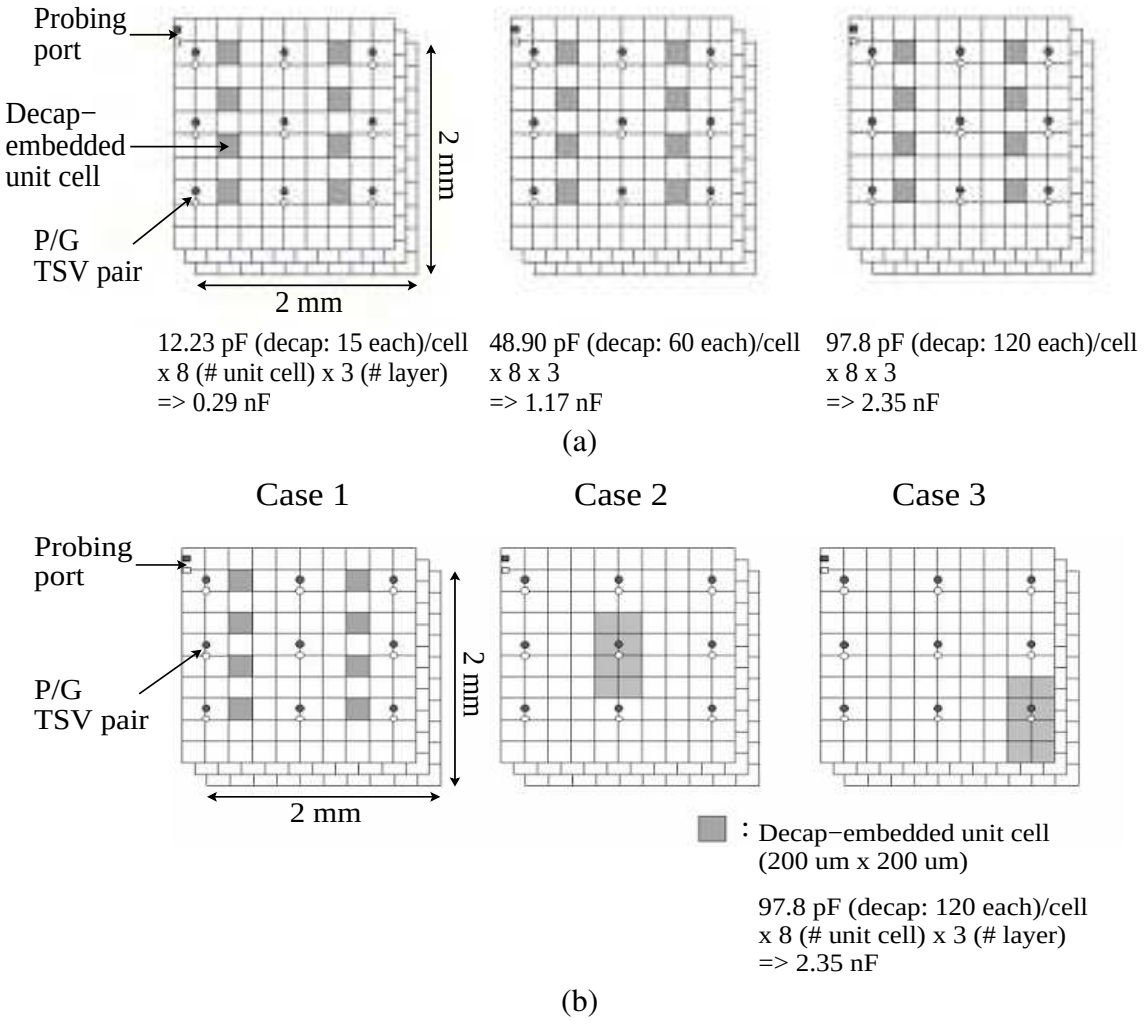


Figure 3.12: The effect on 3-D power network impedance when (a) decoupling capacitance is increased, and (b) location of the decoupling capacitor [166].

network impedance (from approximately 50 Ω to 7 Ω) at DC, with the impedance reducing slightly to 5 Ω at 200 MHz. From 200 MHz to 2 GHz, all three values of capacitance produce similar network impedances. The value increases from 2 GHz to 20 GHz with a network impedance of approximately 6 Ω at 2 GHz and 45 Ω at 20 GHz (due to the power network resistance) [166].

The simulations examine the effect of the location of the decoupling capacitor within the 3-D power distribution network. The network impedance includes a 2.35 nF capacitance for all of the configurations depicted in Figure 3.12(b). Each shaded square in Figure 3.12(b) represents a capacitance value of 97.8 pF. The eight shaded areas per device plane and three device planes produce a total capacitance of 2.35 nF. All of the three cases produce a constant network impedance from DC to approximately 2 GHz. The distributed decoupling capacitor (Case 1 shown in Figure 3.12(b)) produces the lowest network impedance of 4 Ω while both Case 2 and Case 3 exhibit an approximate 5 Ω impedance. The trend continues from 2 GHz to 20 GHz with the distributed decoupling capacitor scheme producing an approximate 40 Ω network impedance at 20 GHz. Cases 2 and 3 produce an approximate 50 Ω impedance at 20 GHz [166].

The allocation and placement of decoupling capacitors in 3-D integrated circuits is significantly different from 2-D ICs. Non-uniform decoupling capacitor placement is possible across device planes, particularly in the case of heterogeneous integration. The simulations examining decoupling capacitor placement reveal that the location of the decoupling capacitor affects the total power network impedance, with distributed capacitors producing the lowest impedance [186]. Novel decoupling capacitor placements are also possible, with the

potential to dedicate a full device plane to decoupling capacitance. Early work has shown up to a 36% reduction in noise when an entire device plane is dedicated to decoupling capacitance [186, 187]. Additional simulated and experimental investigations are needed to determine the optimal size, topology, and placement of the decoupling capacitors in 3-D power delivery networks under non-uniform interplane loads.

3.2.3 Power/ground TSV placement

Placing TSVs to efficiently deliver current among the different device planes requires algorithms and methodologies that consider the voltage and noise requirements of each plane, the active area lost to the TSV, and the effect of the TSV impedance on the power distribution network. Approximately 30 to 40% of the available TSVs are used for power delivery [176, 188], which depends on the total number of device planes, number of unique voltages, and the current demand. The area overhead of placing TSVs within a 3-D power distribution network ranges from 2 to 9% [170]. In addition, a TSV allocation strategy has been developed for microprocessor and DRAM 3-D ICs, where the effect of noise and area on the location of the TSVs placed through the DRAM substrate is examined [168].

The through silicon via density affects the output impedance of the power distribution network [176]. For low frequencies, the higher via density produces a smaller output impedance as the resistance of the power network is reduced. However, at high frequencies, the power distribution network with a higher TSV density produces a higher impedance. The higher impedance peaks at high frequencies due to the increase in the Q factor as the resistance of the network is reduced, leading to greater anti-resonant behavior [176].

The relationship between the TSV density and power network noise is not conclusive, as described in [187]. Although a small drop in the peak power noise occurs when the number of TSVs is increased from 2,048 to 30,000 for a fixed number of power/ground I/O pads (2,048), the primary cause of power grid noise is the limited number of power and ground I/O pads. A maximum 70% reduction in the peak noise is observed when the number of power and ground I/O pads is increased from 2,048 to 30,000 for a fixed number of TSVs [187]. These results indicate a tradeoff between the number of I/O pads necessary to minimize noise on the power network with the number of I/O pads available for signaling, in particular since only one side of the 3-D IC is available for I/O placement.

Algorithms have been developed to estimate and minimize the substrate area dedicated to the power delivery network. During the early design phase of the power distribution network, the minimum number of TSVs necessary to deliver power are estimated, while remaining within specified noise margins [167, 189]. These algorithms reduce the number of *IR* analyses by 51, significantly lowering the computational time [167]. The TSV minimization problem is formulated as

$$min \sum_{i=1}^{M} n_i, \tag{3.14}$$

constrained by

$$\forall_{i,j,k} : |V_{ref} - v_{i,j}^{k}| \leq V_b, \tag{3.15}$$

and

$$\forall_i : n_i \geq 0, \tag{3.16}$$

where n_i is the total number of TSVs, V_{ref} is the input voltage reference, V_b is the maximum allowable noise voltage, and $v_{i,j}^k$ is the voltage between two adjacent nodes i and j on device plane k [167]. Four separate optimizations are performed: 1) reduce the maximum slack, 2) reduce an arbitrary slack, 3) reduce the local slack, and 4) improve any worst case noise violations. The voltage slack is the difference between $v_{i,j}^k$ and $(V_{ref} - V_b)$. An iterative approach is applied to achieve each optimization objective. The total number of TSVs is gradually incremented at each grid node, providing a solution with the minimum number of TSVs as compared to other algorithms that start with a large number of TSVs and iteratively lower the number of TSVs [167]. The algorithm runs in approximately one hour on a cluster of 40 processing nodes for a circuit with three stacked dies and 256 TSV grid points.

Due to mechanical stress from the TSV, a keep out zone is required around the TSV where no devices can be placed, requiring additional area [167, 189]. The total area of the TSVs on a power network is

$$A_{TSV} = S \times \sum n_i, \tag{3.17}$$

where S is the size of the TSV including the keep out zone, and n_i is the total number of TSVs. The area penalty for including TSVs and the power consumed by the TSVs in a DRAM stack is examined in [168]. The area occupied by the TSVs in a homogeneous 3-D IC, where all device planes include identical logic is

$$A_{TSV} = L_v W_{pch} N_{tsvpch_h} + L_h W_{pch} N_{tsvpch_v} + N_{tsvs} P_s^2, \tag{3.18}$$

where L_v is the length of the die along the vertical direction, L_h is the length of the die in the horizontal direction, W_{pch} is the width of the through die power TSV channel, P_s is the pitch of the signal TSV, and N_{tsvpch_h}, N_{tsvpch_v}, and N_{tsvs}, are respectively, the number of horizontal power TSV channels, vertical power TSVs channels, and signal TSVs. The third term in (3.18) considers the area penalty due to the signal TSVs [168].

The power consumed by the TSVs within a power network must also be included when placing TSVs within a power network. Although the resistance of the power grid is reduced, each TSV consumes a finite amount of power. Assuming that a uniform current passes through all TSVs within a 3-D power network, the current flowing through each TSV is

$$I_{tsv_p} = \frac{P_{die}}{V_{dd} \frac{N_{tsv_p}}{2}}, \tag{3.19}$$

where P_{die} is the total power consumption of the die, V_{dd} is the power supply voltage of the die, and N_{tsv_p} is the number of TSVs within the power network. The power consumption of each TSV is therefore

$$P_{tsv_p} = N_{tsv_p} I_{tsv_p}^2 R_{tsv_p}, \tag{3.20}$$

and the total power consumed by all of the TSVs within the power network is

$$P_{tsvp_{total}} = \frac{16}{\pi}\frac{\rho N P_{die}^2 T_t}{N_{tsvp} V_{dd}^2 D_{tsv}^2} + \frac{16}{\pi}\frac{P_{die}^2 (2N-1)\rho_c}{N_{tsvp} V_{dd}^2 D_{tsv}^2}. \tag{3.21}$$

The parameters in (3.21) include the diameter of the TSV D_{tsv}, the total thickness of each die T_t (which includes the thickness of the adhesive layer, metal layers, active silicon layer, and substrate), the total number of dies N, the resistivity of the material used to form the TSV ρ, the specific contact resistance ρ_c, and the total number of TSVs within the power network N_{tsv} [168].

The placement of TSVs within a power network affects both the impedance of the power network and the power consumed by a 3-D integrated circuit. A tradeoff therefore exists between the power network resistance and power consumption when increasing the total number of TSVs used within a power distribution network. Placement algorithms must consider both of these effects when determining the optimal number and location of TSVs within a power network.

3.3 Closed-form expressions of power noise

Analytic models provide a computationally efficient method for estimating power supply noise in 3-D ICs [177]. Compact models of the DC *IR* drop within a 3-D power delivery network with a non-uniform grid impedance have been developed [171]. Mathematical models of the switching *L·di/dt* noise have also been described [178, 187].

A compact model of static *IR* noise includes the location dependent resistivity using a finite volume formulation, which permits a heterogeneous power delivery network to

be evaluated with differing resistivities [171]. The governing expression for the voltage distribution within a 3-D power network is

$$\nabla \cdot \left(\frac{1}{\rho(x,y,z,T)} \nabla \phi(x,y,z) \right) = 0, \tag{3.22}$$

where ϕ(x,y,z) represents the voltage distribution across an entire 3-D network, and ρ(x,y,z,T) is the temperature dependent electrical resistivity of the conductors [171]. Equation (3.22) is solved for the noise voltage distribution based on power supply and current source boundary conditions and other homogeneous boundaries within a specific power distribution network. The power network of each device plane is partitioned into a non-uniform mesh, as depicted in Figure 3.13. The finite volume method is applied to solve the differential equation in (3.22). The voltage distribution is solved using local dependent resistivities between nodes within the power network, as shown in Figure 3.13. A node at grid point (i,j) is surrounded by four other nodes with nodal distances, Δx_1, Δx_2, Δy_1, and Δy_2. The resistivity of the four cells surrounding the grid point (i, j) does not need to be identical, and are ρ_1, ρ_2, ρ_3, and ρ_4 (see Figure 3.13) [171].

The node (i, j) is surrounded by a finite volume cell enclosed by the dashed lines shown in Figure 3.13. The voltage at node (i, j) is given by $\Phi_{i,j}$. The finite volume method is applied to the enclosed cell delimited by the dashed lines. The corners of the dashed line intersect the center points of the four cells surrounding the node (i, j) [171]. Integrating (3.22) over the dashed cell and applying the divergence theorem results in

$$\int_{\substack{dashed \\ line}} \frac{1}{\rho(x,y,z,T)} \nabla \phi(x,y,z) \cdot \hat{n} dl = 0. \tag{3.23}$$

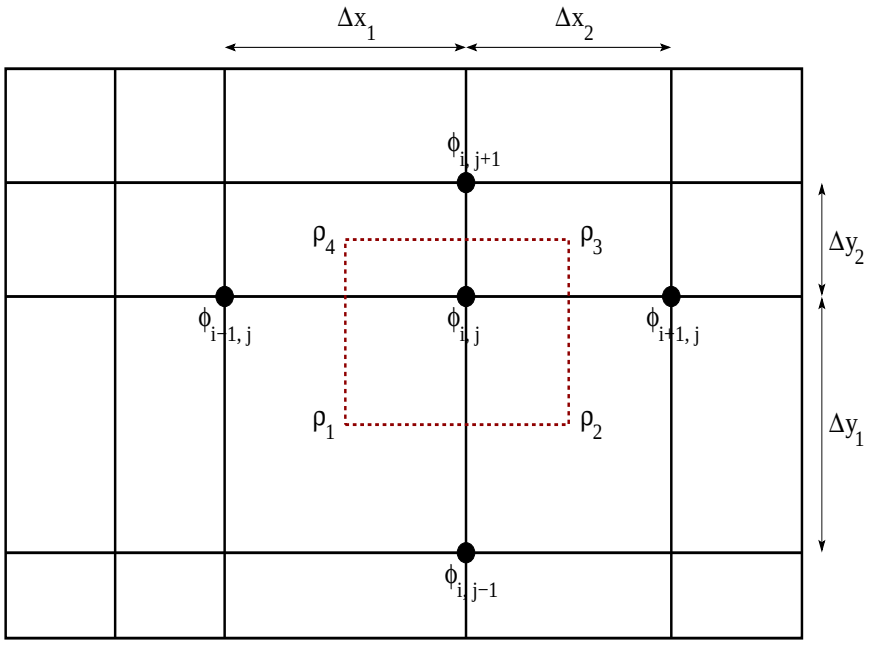


Figure 3.13: Rectangular grid of a 2-D power delivery network to determine voltage distribution [171].

The inhomogeneous resistivity of the power distribution network is considered by including the location dependent resistivity of the surrounding cells ρ_1, ρ_2, ρ_3, and ρ_4. The finite volume method with non-uniform resistivity has a general solution,

$$\begin{aligned}\left(\frac{\Delta y_1}{\rho_1 \Delta x_1}+\frac{\Delta y_2}{\rho_4 \Delta x_1}\right)\left(\phi_{i,j}-\phi_{i-1,j}\right)+\left(\frac{\Delta y_1}{\rho_2 \Delta x_2}+\frac{\Delta y_2}{\rho_3 \Delta x_2}\right)\left(\phi_{i,j}-\phi_{i+1,j}\right)+\\ \left(\frac{\Delta x_1}{\rho_1 \Delta y_1}+\frac{\Delta x_2}{\rho_2 \Delta y_1}\right)\left(\phi_{i,j}-\phi_{i,j-1}\right)+\left(\frac{\Delta x_1}{\rho_4 \Delta y_2}+\frac{\Delta x_2}{\rho_3 \Delta y_2}\right)\left(\phi_{i,j}-\phi_{i,j+1}\right)=0.\end{aligned} \tag{3.24}$$

The DC *IR* drop is converted into a system of equations of the form $Yx = b$ from the finite volume formulation of (3.24). The impedance matrix Y is sparse and symmetric positive definite (SPD) [190].

Compact models of the transient power noise have also been developed [178, 187]. The power and ground networks are approximated as a single continuous planar surface as the

number of power/ground grids is typically large [187]. The frequency characteristics of the power network noise $V_i(x,y,s)$ for each node on the i^{th} device plane is determined by solving the partial differential equation, which is a combination of Poisson and Helmholtz equations, given by

$$\nabla^2 V_i(x,y,s) = R_{si}J_i(s) + 2V_i(x,y,s)\cdot sR_{si}C_{di} + \Phi_i(x,y,s), \tag{3.25}$$

where $\Phi_i(x,y,s)$ is the source function for the partial differential equation of all layers i except for the first layer and is

$$\Phi_i(x,y,s) = R_{si}\sum_{k=1}^{N_{via}}\left(\frac{V_{(i-1),viak}-V_{i,viak}}{sL_{viak}+R_{viak}} - \frac{V_{i,viak}-V_{(i+1),viak}}{sL_{viak}+R_{viak}}\right)\cdot\delta(x-x_{viak})\,\delta(y-y_{viak}). \tag{3.26}$$

The source function for the first device plane is

$$\Phi_1(x,y,s) = -\frac{R_{s1}}{4sL_p}V_{pad}(s)\cdot\delta(x)\delta(y) + \sum_{k=1}^{N_{via}}\left(-\frac{V_{1,viak}-V_{2,viak}}{sL_{viak}+R_{viak}}\right)\cdot\delta(x-x_{viak})\,\delta(y-y_{viak}). \tag{3.27}$$

Equations (3.26) and (3.27) are dependent on the number of TSVs on each device plane N_{via} and the voltage at the location where via k connects to layer i. The source functions described by (3.26) and (3.27) also connect device plane *(i-1)*, *i*, and *(i+1)*. The source function for the first device plane includes the effect of the package inductance on the generated noise, where the noise voltage of the power and ground I/O pads is included as

V_{pad}. No current flows normal to the cell boundaries and therefore the boundary conditions are [177, 187]

$$\frac{\partial V_i}{\partial y} I_{x=0} = 0, \frac{\partial V_i}{\partial y} I_{x=a} = 0, \frac{\partial V_i}{\partial x} I_{y=0} = 0, \frac{\partial V_i}{\partial x} I_{y=a} = 0, \tag{3.28}$$

where a is the cell size.

The differential equations of (3.26) and (3.27) are transformed into Helmholtz equations [191] and are solved analytically by applying the boundary conditions described by (3.28). The power supply noise for the i^{th} device plane and the first device plane are, respectively,

$$V_i(x,y,s) = R_{si} \sum_{k=1}^{N_{via}} \left(\frac{V_{(i-1),viak} - V_{i,viak}}{sL_{viak} + R_{viak}} - \frac{V_{i,viak} - V_{(i+1),viak}}{sL_{viak} + R_{viak}} \right) \cdot G(x,y,x_{viak},y_{viak},s) - \frac{J_i(s)}{2sC_{di}} \tag{3.29}$$

and

$$V_1(x,y,s) = -\frac{R_{s1}}{4sL_p} V_{pad}(s) \cdot G(x,y,0,0,s) + \sum_{k=1}^{N_{via}} \left(-\frac{V_{1,viak} - V_{2,viak}}{sL_{viak} + R_{viak}} \right) \cdot G(x,y,x_{viak},y_{viak},s) - \frac{J_1(s)}{2sC_{d1}}, \tag{3.30}$$

where $G(x,y,\xi,\eta,s)$ is the Green's function of the Helmholtz equation, where an expression for the Green's function is found in [191]. The pad voltage V_{pad} and TSV voltages $V_{i,viak}$ are unknowns that can be solved from (3.26) and (3.27).

The time domain transient noise is obtained by performing an inverse Laplace transform of (3.26) and (3.27). The peak noise is the summation of the transient noise on the power and ground networks. The model has been validated against SPICE for a 3-D high

performance microprocessor with five device planes in 45 nm CMOS. The top two metal levels are allocated to the power delivery network. Five identical critical blocks are placed on each device plane, and a current density of 100 A/cm^2 per die is assumed. Decoupling capacitors account for 20% of each die area, and 49 nodes are assigned to each power and ground network. Each power/ground pad includes a 0.5 nH inductance that connects the pad to the package. A lumped impedance model for a 50 μm diameter and 200 μm long TSV is used. The difference in peak noise obtained from (3.26) and (3.27) and SPICE simulations is less than 4% across all device planes [187].

3.4 Voltage regulators for 3-D ICs

The integration of disparate technologies such as analog, RF, digital, MEMS, mixed-signal, memory, and communications often require a variety of power supply voltages. Two methods of delivering these unique voltages include: 1) dedicated power distribution networks for each voltage, and 2) on-chip conversion of a source voltage into two or more voltages. Dedicated power delivery networks require additional metal resources. In particular, delivering the power to a specific device plane requires dedicated TSVs for each power network. The device plane closest to the I/O pads incurs the largest area penalty as only one side of a 3-D IC is typically used for I/O; all unique voltages therefore pass through TSVs on this plane.

3-D power delivery utilizing on-chip voltage conversion is an appealing alternative to distribute current to the disparate device planes. Conventional DC-DC voltage converters are categorized into three categories: 1) switching, 2) switch capacitor, and 3) linear [192,

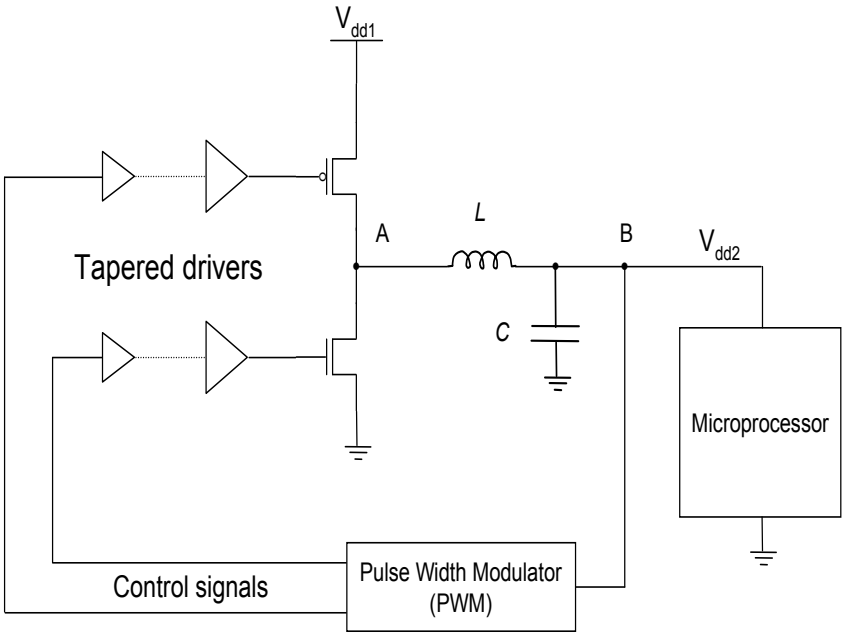


Figure 3.14: Conventional buck converter. The inductor and capacitor are typically placed off-chip [192].

193]. High performance microprocessors typically use switching DC-DC buck converters due to the high power efficiency and good output voltage regulation characteristics [80, 192, 194]. A conventional buck converter, which down converts an input voltage, is shown in Figure 3.14. The capacitor and inductor shown in Figure 3.14 are often placed off-chip due to the large area requirement for integrating these passive impedances on-chip. The basic operating principles of a buck converter are presented in Subsection 3.4.1, and a comparison of buck converter based circuits as applied to 3-D power delivery is described in Subsection 3.4.2.

3.4.1 Basic operation of buck converter

The basic operation of a buck converter includes power MOSFETs that produce a continuous sequence of pulses at node A, as shown in Figure 3.14. The input voltage to the *LC* rectifier is converted to a DC voltage at the output based on the duty cycle, period, input voltage, and rise and fall time of the input waveform. The inductor produces a ripple current (and therefore voltage) at the output node B. This voltage is typically constrained to within 10% of the output voltage [195–197]. The input voltage waveform is shown in Figure 3.15. The DC component of the output is

$$\frac{1}{T_s}\int_0^{T_s} V_A dt = \frac{V_{dd1}\left(2DT_s - t_{tr} + t_{fp}\right)}{2T_s}, \tag{3.31}$$

where D is the duty cycle, V_{dd1} is the input voltage, T_s is the period, and t_{rp} and t_{fp} are, respectively, the rise time and fall time of the pulse. When the rise and fall times are equivalent, the DC component described by (3.31) reduces to

$$V_{dd2} = V_{DC|t_{rp}=t_{fp}} = DV_{dd1}. \tag{3.32}$$

The output node B is fed back to the power transistors through a pulse width modulator (PWM), as shown in Figure 3.14. The PWM produces the pulse chain that controls the power MOSFETs by adjusting the duty cycle of the waveform. Due to the PWM, the output voltage is maintained at a specified voltage while compensating for variations in the load current and input voltage [198].

The primary power loss components of the buck converter are capacitive switching loss

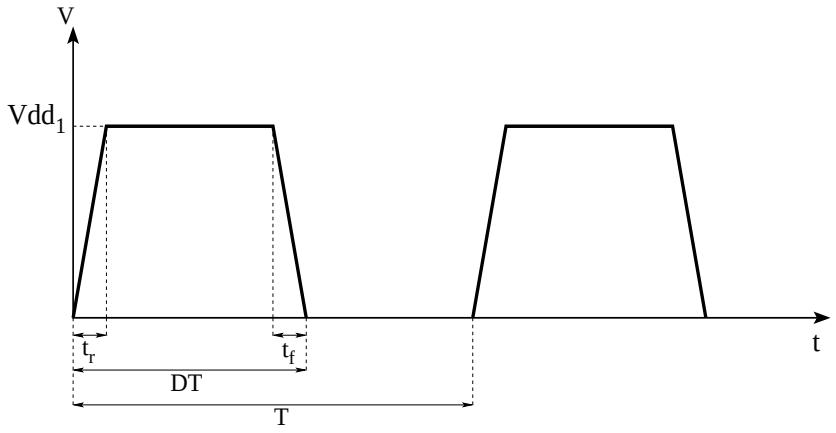


Figure 3.15: Input voltage waveform at node A (shown in Figure 3.14) to the *LC* rectifier [192, 198].

and resistive loss at node A and losses due to the inductor (the resistor is not shown in Figure 3.14, but is in series with the inductor). The total loss is [80]

$$P_{loss} = P_{cap} + P_{res} + P_{ind}, \tag{3.33}$$

The capacitive switching losses are

$$P_{cap} = C_b V_{in}^2 f = W C_o V_{in}^2 f, \tag{3.34}$$

where V_{in} is the input voltage, f is the frequency of the input voltage, and W is the combined width of the NMOS and PMOS transistors. The capacitance shown in Figure 3.14 is a combination of the output capacitance of the transistors C_b and the decoupling capacitance added to the power network. The capacitance C_b in (3.34) is

$$C_b = W_N C_N + W_P + C_P = W C_o, \tag{3.35}$$

where C_o is

$$C_o = \frac{C_N + \alpha C_P}{1+\alpha}. \tag{3.36}$$

The effective switching capacitance of the NMOS and PMOS transistor are, respectively, C_N and C_P.

The power loss due to the inductor equals the product of the intrinsic inductor resistance R_i multiplied by the square of the root mean square current I_{rms} through the inductor, and is

$$P_{ind} = R_i I_{rms}^2 = \frac{V_{in} D(1-D)}{2 f \tau_i I_R}\left(I_L^2 + \frac{I_R^2}{3}\right). \tag{3.37}$$

The ripple current created by the inductor is

$$I_R = \frac{V_{in} D(1-D)}{2 f L_i}, \tag{3.38}$$

and the intrinsic inductor resistance is

$$R_i = \frac{V_{in} D(1-D)}{2 f \tau_i I_R}. \tag{3.39}$$

The ripple current is dependent on the input voltage V_{in}, the frequency of the waveform at node A f, the duty cycle D, and the inductance L_i. The inductor time constant τ_i is

$$\tau_i = \frac{L_i}{R_i}, \tag{3.40}$$

and the root mean square current is

$$I_{rms}^2 = \left(I_L^2 + \frac{I_R^2}{3} \right), \tag{3.41}$$

The root mean square current is dependent on both the ripple current I_R and the load current I_L. The capacitors within the power network decouple the ripple current from the load current.

The third component that produces power loss is the resistance of the power transistors. The effective resistance of the power transistors R_b multiplied by the square of the root mean square current I_{rms}, quantifies the power loss across the transistors. The loss across the power transistor is

$$P_{res} = R_b I_{rms}^2 = \frac{R_o}{W} \left(I_L^2 + \frac{I_R^2}{3} \right), \tag{3.42}$$

where R_b is the effective resistance of the power transistors,

$$R_b = \frac{R_o}{W} = \frac{R_N}{W_N}(1 - D) + \frac{R_P}{W_P}(D). \tag{3.43}$$

3.4.2 3-D IC power converters

Recent work on 3-D integrated power generation has focused on switching converters, with a primary focus on buck converter topologies [195–201]. One research area examines the topological placement of the buck converters within a device plane adjacent to the load [197, 199–201]. A second research area focuses on novel techniques that exploit the

impedance of the 3-D power distribution network to reduce the size of the *LC* filter [195, 196, 198].

Placement of the buck converter, specifically the inductor and capacitor, on an adjacent device plane supports point-of-load (PoL) power delivery. Research on integrating a buck converter with the CMOS [197, 199] and SiGe [201] device planes has been described. Experimental verification of the buck converter is listed in Table 3.3. The experimental results demonstrate that both CMOS and SiGe can be integrated with a device plane dedicated to power conversion.

One research study explores buck converters on a silicon interposer layer and integrates the interposer with CMOS circuitry [199]. The CMOS LSI includes active components such as CMOS switches, pulse wave modulators, and driver circuits. The passive components of the buck converter, namely the metal-insulator-metal (MIM) capacitor and spiral inductor, are integrated into the silicon interposer. Silicon interposer technology is considered a precursor to full TSV-based 3-D integration, and is often referred to as 2.5-D technology [202–204]. The interposer and CMOS device plane are connected through metal bumps, producing a small parasitic capacitance and resistance between device planes [199].

The silicon interposer technology realizes a fine pitch design rule of 30 μm line width and spacing, a 30 μm via hole diameter, and a metal thickness of 15 μm. The metal thickness is three times that of conventional interposers, and therefore significantly boosts the performance of the inductor. A 2,000 by 2,000 μm^2 test circuit that includes a 6 turn square spiral inductor experimentally verifies the 3-D buck converter. The MIM capacitor is placed underneath the spiral inductor within the interposer plane. An input voltage of 1.8 volts is

converted to 1.0 volts. The peak power efficiency of the buck converter with a 15 μm thick inductor is 77% with an output current of 70 mA, while a 5 μm thick inductor produces a 66% efficiency with an output current of 60 mA [199].

A second research result explores the integration of a fully monolithic interleaved DC-DC buck converter for point-of-load power delivery [201]. The power stage of the buck converter includes gate drivers, a control switch, a synchronous rectifier, an output filter inductor and capacitor, and an internal active load. The buck converter uses a tapered buffer to drive the power MOSFETs, where a PMOS transistor is used as a pull-up control switch and an NMOS transistor as a pull-down synchronous rectifier. A fabricated prototype includes two buck converters operating in parallel in a 0.18 μm SiGe BiCMOS process. Each cell delivers a nominal output current of 500 mA and converts a 1.8 volt input to 0.9 volts. The switching frequency of the buck converter is 200 MHz. An on-chip filter includes an air-core spiral inductor with a width of 25 μm, minimum spacing of 5 μm, and a total of 3.5 turns. The total inductance is 2.14 nH and the outer diameter is 290 μm. The resistance of the winding inductor is 201 mΩ. A 8.22 nF capacitor (output capacitor) is a MOS capacitor rather than an MIM or PIP capacitor to reduce area [201].

The prototype buck converter circuit matches the footprint of the CMOS device plane to achieve wafer level 3-D integration. The total area of the die is 10 mm^2. 57.9% of the total area is dedicated to the capacitors, with the input decoupling capacitor occupying 30.8% and the rectifier capacitance occupying 27.1% of that 57.9%. On-chip active loads are used for dynamic response testing [201]. The power efficiency of the buck converter is relatively independent of the switching frequency, ranging from 160 to 220 MHz, with a

slight decrease in output current. A maximum efficiency of 64% is achieved at 200 MHz for a 500 mA output current [201]. The dynamic response test reveals an approximately 88 mV overshoot voltage which returns to the desired level in 86 ns when a 225 mA current step with a slew rate of 10 A/μs is applied [201]. A 6 mV drop in output voltage is observed at higher current levels due to the relatively small DC gain of the compensator (34 dB).

A third research result investigates the integration of a buck converter with a CMOS device plane, producing results similar to the two aforementioned prior efforts. These results are summarized in Table 3.3. The primary result is the integration of a buck converter with a maximum efficiency of 62% at 70 mA output current with a switching frequency of 200 MHz. The input voltage of 3.3 volts is converted to 2.3 volts. The power MOSFETs are 1000 μm and 500 μm, respectively, for the pull-up and pull-down transistors [205].

Another research path on 3-D power generation eliminates the discreet *LC* filter and replaces that filter with the intrinsic impedance of the metal interconnect and TSV impedance [195, 196, 198]. The distributed filter topology removes the need for dedicated inductors and capacitors, although the decoupling capacitors of the power delivery network remain present. The per unit length resistance, inductance, and capacitance of the interconnect (Metal 3) and the TSV are listed in Table 3.4. The impedances are determined from the predictive technology model [206].

The low pass bandwidth properties of transmission lines is exploited to simultaneously generate and distribute current to different device planes. A distributed buck converter filter is modeled based on the MITLL 3-D CMOS technology [82]. The power efficiency of the buck converter ranges from 60 to 90% for an output-to-input voltage ratio of, respectively, 0.3 to 0.9 for a 200 mA load current and switching frequency of 1 GHz [195]. The

Table 3.3: Comparison of buck converter topologies for 3-D power delivery.

Reference	Experimental results	V_{in} (V)	V_{out} (V)	Maximum efficiency	Conditions of maximum efficiency	Frequency range	Output current	Notes
[199]	Yes	1.8 V	1.0 V	77%	36 MHz/70 mA	—	10-110 mA	use of silicon interposer for inductor and capacitor placement; 12% increase in power efficiency for 15 μm thick inductor (compared to 5 μm thickness)
[201]	Yes	1.8 V	0.9 V	64%	200 MHz/500 mA	120-220 MHz	500-800 mA	constant power efficiency from 160 MHz to 220 MHz across all current loads; reduced inductor and capacitor sizes; 0.18 μm SiGe BiCMOS process
[197]	Yes	3.3 V	1.8-2.4 V	62%	200 MHz/70 mA	100-400 MHz	20-70 mA	stack IC implementation of buck converter; gate widths of 1000 μm (PMOS) and 500 μm (NMOS); 0.35 μm CMOS process

Table 3.4: *RLC* impedance of aluminum interconnect and TSV [195].

	R [mΩ/μm]	L [pH/μm]	C [fF/μm]
Al Interconnects	14.5	1.3	0.5
Cu Interconnects	10.4	1.3	0.89
TSV [134, 207]	20.4	0.55	0.37

maximum current load is 700 mA at an output voltage of 1.2 volts.

The on-chip inductor of a rectifier can range from 50 to 10,000 nH for switching frequencies up to 1 GHz [195]. The large range of inductance values is due to the wide range of switching frequencies applied to the power MOSFETs. When the frequency is 3 GHz, the inductor is less than 10 nH. At low switching frequencies of up to several megahertz [208], a conventional *LC* filter cannot be integrated on-chip [195]. In the MHz frequency range, the capacitor can range between 100 pF and 100 nF, while the size of the inductor ranges between 60 nH and 10 μm [195, 196].

The performance characteristics of the distributed rectifier circuit are analyzed over a wide range of design parameters [195, 196, 198]. The distributed filter is designed to produce a voltage ripple of less than 5% for an input voltage of 3.3 volts. The power efficiency is considered for six different switching frequencies: 1 MHz, 10 MHz, 100 MHz, 500 MHz, 1 GHz, and 3 GHz. The efficiency is 80% at a switching frequency of 1 MHz and 98% for frequencies between 500 MHz and 3 GHz [195]. One limitation of the study is that the power efficiency is solely based on the filter characteristics, ignoring the power MOSFETs. The efficiency at higher switching frequencies is therefore dependent on the input impedance of the distributed filter and does not change significantly at higher frequencies. In addition, the input impedance remains approximately constant at higher frequencies

since the distributed capacitance of the filter contributes less to the input impedance. The total area and capacitance of the distributed filter therefore decreases as the frequency increases [195, 196, 198].

A primary benefit of 3-D integration, as suggested by the research results described in this section, is the potential to integrate disparate device planes. Focus has been placed on integrating switching based power converters in those device planes directly adjacent to the current load. Additional work examines topological configurations not available in 2-D ICs, such as the distributed rectifier of the buck converter. The objective is an integrated point-of-load power generation and delivery system that supports tens of voltages while satisfying the current demands of each device plane.

3.5 Summary

Power delivery in 3-D integrated circuits is a challenge due to increasing operating frequencies and power densities with decreasing supply voltages. A 3-D power delivery system requires additional considerations as compared to 2-D ICs. The most important consideration is that vertically stacking dies requires through silicon vias, which occupy active area previously reserved for transistors. An additional consideration unique to 3-D stacking is integrating disparate technologies with different process characteristics, current demands, and power network requirements. The integration of heterogeneous device planes results in varying impedance characteristics, noise requirements, and current loads, which produces varying simultaneous switching and *IR* noise characteristics within a 3-D power network.

Characterization and modeling of the 3-D power network is necessary to design 3-D power delivery systems that effectively distribute current to each device plane while satisfying the noise requirements of each plane. Models of a 3-D power distribution network have been developed, with certain models incorporating the effect of the TSV on the noise generation and propagation characteristics. Research on power network models that include the TSV also consider the tradeoff between silicon area lost due to the TSVs and the lower impedance of the 3-D power network.

The potential to integrate multiple device planes in a small form factor supports novel circuit topologies. One particular research focus explores integrating point-of-load power delivery with CMOS to form a 3-D IC. Preliminary test circuits integrating buck converters on a device plane separate from the CMOS circuits have been tested, indicating that high efficiency point-of-load buck converters can be incorporated into a small footprint without occupying area on a CMOS device plane.

In addition, novel power generation methodologies that explore replacing the rectifier circuit of the buck converter with the intrinsic impedances of the metal interconnect and TSVs have been developed. The overall objective of power delivery in 3-D ICs is similar to that of 2-D systems-on-chip: To efficiently and effectively generate and distribute multiple supply voltages while satisfying the current demands of the many millions to billions of devices within a 3-D integrated system.

Chapter 4

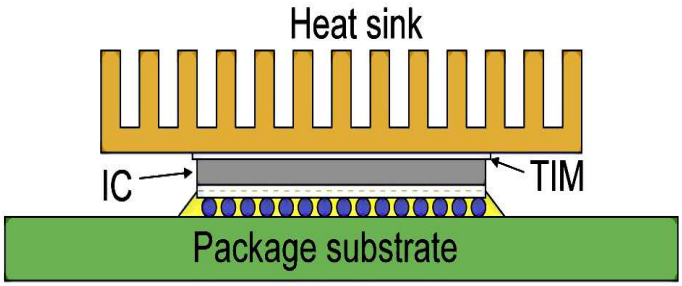
Heat Generation and Thermal Mitigation Techniques in 3-D Systems

The formation of thermal hotspots is not unique to 3-D integrated circuits. Due to the increased current density within 2-D integrated circuits, hotspots detrimentally affect the reliability of a circuit and limit the maximum frequency of operation [209]. The reliability of interconnect is particularly affected by increases in die temperature [209, 210] as the interconnect lifetime is exponentially dependent on the inverse of the temperature $\frac{1}{T}$ [211] and open-circuit metal failure is more likely at higher circuit temperatures [212]. In addition, the induced self-heating, which increases the resistance of an interconnect, limits the maximum allowable current density through a wire. Thermal aware design is therefore of significant importance in modern integrated circuits.

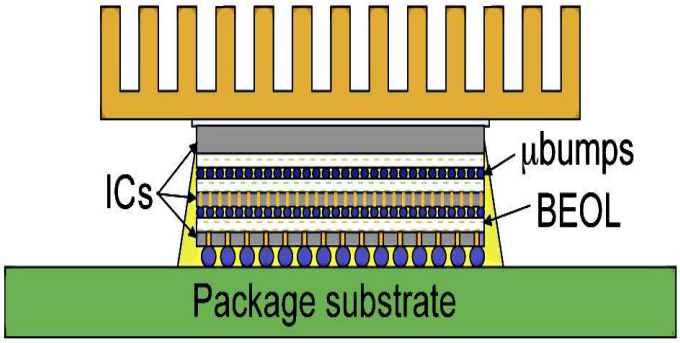
Thermal effects in 3-D integrated circuits present thermal management issues that are expected to greatly exceed the deleterious effects currently experienced with 2-D ICs [213–215]. There are three primary reasons for these anticipated increases in temperature: 1) A much higher power density due to an increase in the number of devices stacked within a smaller footprint, 2) the smaller footprint provides less surface area to propagate heat to

the heat sink, and 3) an increase in the distance of the thermal conductive paths from the devices to the heat sink [210, 216, 217]. An illustration of the differences between a packaged 2-D IC and 3-D IC including the heat sink is shown in Figure 4.1. Not only higher temperatures can result, which can potentially exceed the thermal limits of existing packaging technologies [17] but also greater thermal gradients can be produced. There are two essential components to properly manage thermal effects within 3-D integrated circuits. The first component is the proper characterization and modeling of hotspot formation and thermal propagation, and the second component is a set of design techniques and physical methods that reduce the peak temperature and minimize thermal gradients among the device planes within a 3-D IC. The objective is to produce highly accurate thermal models that include the effects of multi-plane thermal coupling as well as the effects of several mitigation techniques (such as thermal TSV placement, block placement, and active cooling techniques) while maintaining a low computational design time. These design methodologies and cooling techniques increase design complexity. The primary objective is to ensure that the on-chip temperature satisfies specifications with minimum impact on design time and physical area.

Current research on the characterization and compact modeling of thermal effects in 3-D ICs is discussed in this chapter. Research on modeling heat flow within a 3-D stack is described in Section 4.1. Both compact and mesh-based thermal models are discussed. Both passive and active techniques to minimize the peak temperature within 3-D ICs are described in Section 4.2. The passive techniques include the insertion of thermal TSVs and thermal aware floorplanning, while active liquid cooling through microchannels is also

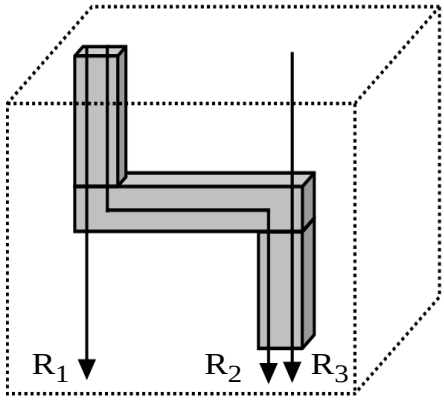


(a)



(b)

Figure 4.1: Schematic representation of a package with a heatsink for (a) a 2-D IC, and (b) a 3-D IC.



R_1: Metal dielectric interface

R_2: Metal (horizontal and vertical)

R_3: Dielectric metal interface

Figure 4.2: Vertical heat transfer path within a 3-D IC [220].

described. Some concluding remarks are provided in Section 4.3.

4.1 Thermal models of 3-D integrated circuits

Heat generation within 2-D and 3-D integrated circuits is a result of current flow through the transistors (and other active devices such as diodes) and the interconnect. Active devices behave like heat sources. Both active devices and interconnect also self-heat as a result of Joule heating which can significantly increase the circuit temperature [218, 219]. Three potential heat transfer paths in 2-D and 3-D ICs are shown in Figure 4.2. The first path is a metal to dielectric (or silicon) heat transfer path, the second path is a dielectric to metal heat transfer path, and the third path is a metal to metal heat transfer path.

The temperature distribution within any portion of a die can be described by the heat equation which is a parabolic partial differential equation (PDE) described by [221–223]

$$\rho c_p \frac{\partial T(r,t)}{\partial t} = k_t \nabla^2 T(r,t) + g(r,t), \tag{4.1}$$

where r represents the Cartesian coodinates of the point at which the temperature is determined, T is the temperature in Kelvin, ρ is the density of the material in $\frac{kg}{m^3}$, t is time in seconds, c_p is the heat capacity of the material in $\frac{J}{kg \cdot K}$, k_t is the thermal conductivity of the material in $\frac{W}{m \cdot K}$, and g is the power density per unit volume in $\frac{W}{m^3}$ [221–223]. The solution to (4.1) corresponds to the transient thermal response. In the steady state, all derivatives with respect to time are set to zero. A steady-state analysis of the heat transfer function therefore corresponds to solving Poisson's equation, the partial differential equation described by [223, 224]

$$k_t \nabla^2 T(r) = -g(r). \tag{4.2}$$

Equations (4.1) and (4.2) are solved by thermal field solvers for, respectively, time-dependent and steady-state conditions. Both expressions require a solution over the entire volume, which is a computationally expensive result. Therefore, to reduce the complexity of the thermal model, methods based on finite difference, finite element, and boundary element approaches have been adopted to evaluate the temperature of a 3-D circuit. In addition, analytic expressions have been developed to characterize the thermal profile within a 3-D IC.

4.1.1 Compact thermal models

The granularity of a compact thermal model depends upon the volume representation of a 3-D integrated circuit. In one approach, the entire device plane can be modeled as a single node within a vertical stack of nodes, while alternatively, each device plane can be partitioned into smaller representative blocks of nodes. For a single node representation of

a device plane, a uniform power density is assumed across the entire die. The entire die is therefore at one temperature. With finer block granularity, a more accurate representation of the distribution of the power density and therefore the thermal gradients across a device plane as well as between device planes is produced.

A model of the entire device plane as a single node provides a course estimate of the thermal behavior of a 3-D circuit. In this case, the total number of nodes represents the total number of planes. These first order models of the die temperature are easily described in analytic form [17,213,225–227]. A few assumptions are made for these first order models. In particular, due to the short height of the 3-D IC, one-dimensional heat flow is assumed. Since with this assumption heat flows only in the vertical dimension, the lateral boundaries of the 3-D IC are considered to behave adiabatically. No heat is therefore exchanged with the ambient surroundings through the sidewalls. The second assumption is that no heat is exchanged at the top surface of a 3-D IC [213], leaving only the bottom surface, the surface connected to a heat sink, for extracting the heat.

The maximum rise in temperature on the upper planes of a 3-D circuit is estimated from a closed-form compact model based on one-dimensional heat flow. The closed-form expression for the heat flow within a 3-D IC utilizes an expression for the temperature increase ΔT in a 2-D circuit,

$$T_{die} - T_o = R_n\left(\frac{P}{A}\right), \tag{4.3}$$

where P is the power consumption, A is the area ($\frac{P}{A}$ is the power density of the die), T_o is the ambient temperature of 25°C, and R_n is the thermal resistance between the package and

ambient environment [213, 226]. An analytic model to estimate the rise in temperature of the j^{th} active layer within an n-layer 3-D IC is [213, 226]

$$\Delta T_j = \sum_{i=1}^{j} \left[R_i \left(\sum_{k=i}^{n} \frac{P_k}{A} \right) \right], \tag{4.4}$$

where P_k is the power consumed in the k^{th} layer, R_i is the thermal resistance between the i^{th} and the $(i-1)^{th}$ layers, n is the total number of device layers, and A is the area footprint of the 3-D IC. From (4.4), the temperature difference between two adjacent device layers is [228]

$$T_i - T_{i-1} = R_i \sum_{k=i}^{n} \frac{P_k}{A}. \tag{4.5}$$

The rise in temperature at the uppermost n^{th} layer in an n-layer 3-D IC is [213]

$$\Delta T_n = \frac{P}{A} \left[\frac{R}{2} n^2 + \left(R_{ps} - \frac{R}{2} \right) n \right]. \tag{4.6}$$

The assumption is that all of the device planes dissipate the same power and all thermal resistances R between layers are therefore identical (for all $i \neq 1$). In addition, interconnect and device Joule heating is not considered. The R_{ps} term in (4.6) is the thermal resistance between the first layer and the heat sink and is

$$R_{ps} = \frac{t_{si_1}}{k_{si}} + \frac{t_{pkg}}{k_{pkg}}, \tag{4.7}$$

where t_{si1} is the thickness of the silicon substrate of the first plane and k_{si} is the thermal conductivity of the silicon. The thickness and thermal conductivity of the package are

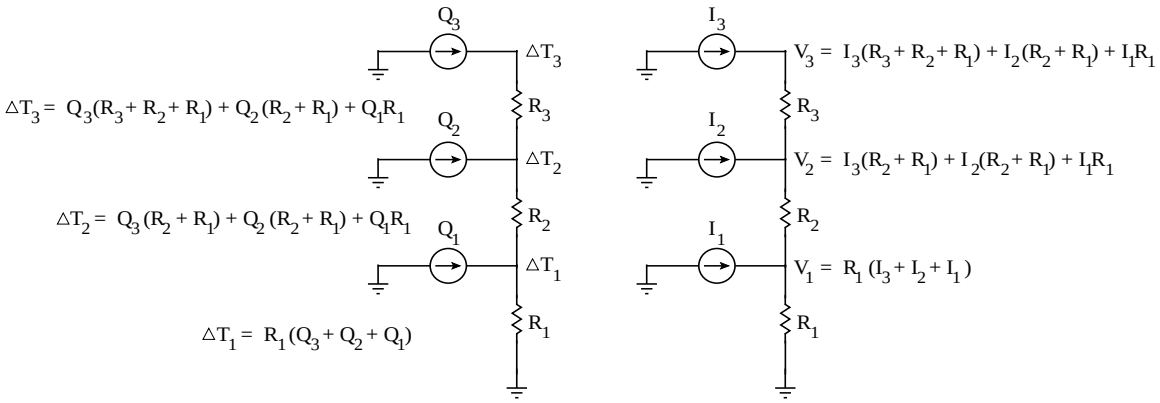


Figure 4.3: Equivalent representation of steady-state thermal properties based on an electrical analog [17]. The change in nodal temperatures is relative to ground.

described by, respectively, t_{pkg} and k_{pkg}. The thermal resistance of device planes k beyond the first plane is

$$R_k = \frac{t_{si_k}}{k_{si_k}} + \frac{t_{diel_k}}{k_{diel_k}} + \frac{t_{inter_k}}{k_{inter_k}}, \tag{4.8}$$

where t_{si_k}, t_{diel_k}, and t_{inter_k} are the thickness, and k_{si_k}, k_{diel_k}, and k_{inter_k} are the thermal conductivity of, respectively, the silicon substrate, dielectric, and bonding interface for the k^{th} plane.

Early work on modeling heat flow within a 3-D IC focused on steady-state conditions, as is evident from the analytic expressions where no time dependent component is included. The simplest representation of this model for steady-state conditions is shown in Figure 4.3. The heat sources are represented as current sources and the thermal resistance as an electrical resistor [17, 227]. The temperature at different nodes of this thermal network corresponds to the node voltages in an electrical network, where an Elmore delay based model is used to calculate the nodal temperatures [17].

Once a basic model of the steady-state thermal flow within a 3-D IC is determined, the model can be extended to include a time-varying component. The dynamic thermal model of a 3-D IC includes a time-varying component, which is typically represented by a capacitor at the node representing each plane (see Figure 4.3). An analytic expression of the thermal transients is approximated by [213, 229]

$$T(t) - T_i = \sum_{i=1}^{j} \left[\left(1 - e^{\frac{t}{\tau_i}} \right) \left(\sum_{k=i}^{n} \frac{P_k}{A} \right) \right], \tag{4.9}$$

where $T(t)$ is the time varying temperature in °C, T_i is the reference temperature in °C, $\frac{P_k}{A}$ is the power density of the k^{th} device plane, R_i is the thermal resistance between the i^{th} and $(i-1)^{th}$ layers, and τ_i is the time constant between the i^{th} and $(i-1)^{th}$ layers. The time constant is given by $R_i C_i$, where C_i is dependent on the thermal propagation properties of the materials.

Thus far this discussion on the thermal profile of a 3-D IC has only considered a vertical thermal path through the device planes. Two additional assumptions are 1) a uniform power density across the entire device plane, and 2) no horizontal thermal spreading. The activity of the circuit blocks is not uniform across device planes, or even across a particular die. In addition, thermal paths such as intraplane metal interconnect spread heat horizontally from the heat sources in addition to the vertical paths through TSVs and the silicon. Both a non-uniform power density and horizontal thermal spreading are considered by a simple manipulation of the model. Rather than a single resistor (and capacitor for dynamic temperature analysis) to represent a device plane, the plane can be sub-divided into nodes interconnected through resistors (and capacitors) depending upon the thermal conductivity

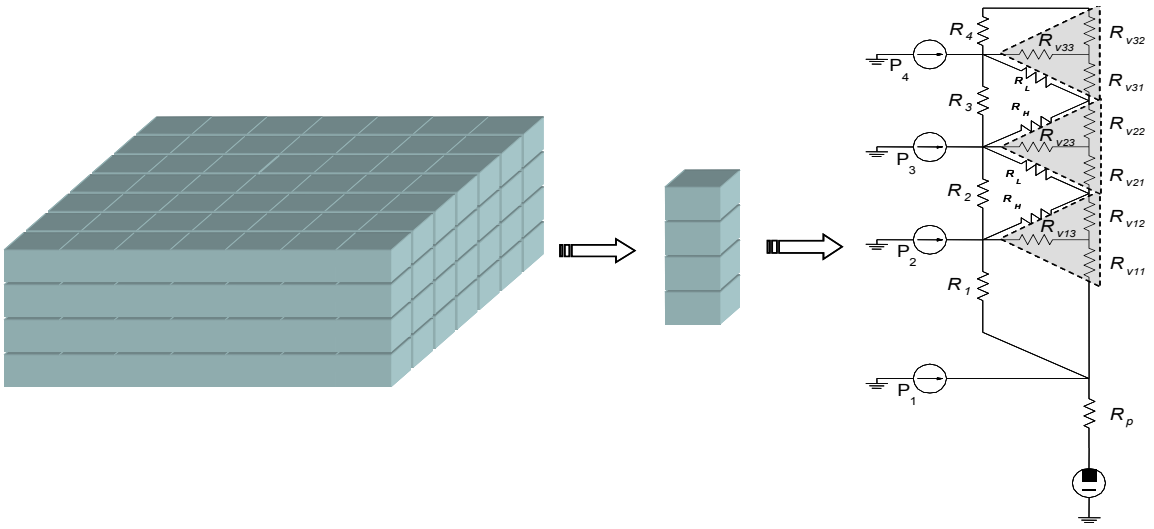
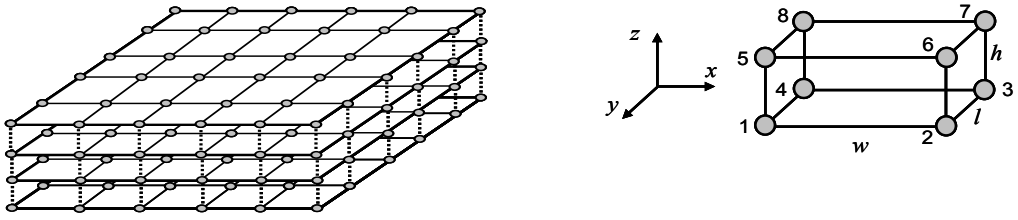


Figure 4.4: A fine grained representation of four device planes for compact thermal modeling [17, 231].

of the material between the nodes. The cost is an increase in the computational complexity of the analytic model representing the intra- and inter-plane heat flow.

The improved model supports power density and temperature vectors that are dependent on all three directions rather than just the vertical z direction. As an example of this model, assume four physical device planes form a 3-D IC, and that the 3-D system is modeled as a thermal resistive stack, as shown in Figure 4.4. The 3-D IC is partitioned into single pillars [230, 231], also shown in Figure 4.4. Each pillar is modeled by a one-dimensional thermal network including the heat sources and thermal resistors. Additional resistors are added to the model to incorporate horizontal heat flow between pillars (not shown in Figure 4.4). The thermal paths through the interplane vias are also included as resistors, as illustrated by the gray shaded triangles shown in Figure 4.4. The heat sources include all of the heat generated by all of the devices contained within each tile of the pillar.



(a) (b)

Figure 4.5: Mesh-based node representation of (a) a four plane 3-D circuit, and (b) a minimum parallelepiped unit block to model thermal effects based on a finite-element method [17].

The basic premise for these more complex compact models and the 1-D model in the closed-form expressions is the same. Both models consist of thermal resistors and heat sources modeled as current sources. The difference in complexity arises due to the finer granularity of the placement of heat sources as well as the inclusion of horizontal thermal paths. This approach raises the issue as to the proper granularity of the model to accurately determine the profile of a 3-D IC without greatly increasing the computational time of the thermal field solver.

4.1.2 Meshed based models

Mesh based models more accurately represent the thermal profile of a 3-D IC. Thermal models based on meshes can be applied to any complex geometry and are not dependent on boundary conditions. A 3-D stacked IC can be decomposed into a three-dimensional structure consisting of finite hexahedron (parallelepiped) elements, as shown in Figure 4.5. The mesh can also be non-uniform for the regions of complex geometry or even utilize non-hexahedron structures such as triangular prisms. In addition, non-uniform power densities can be applied to more accurately model the activity of the circuit.

The temperature at each node, the vertices of the parallelepiped structure, is determined by evaluating (4.1). As previously mentioned, numerical methods including the finite element method (FEM), the finite difference method, the finite volume method, and the boundary element method have been adopted to solve this partial differential equation [230, 232, 233]. The temperature at any other point within a 3-D mesh can be determined from interpolation.

Mesh-based thermal models are more accurate than compact thermal models, as compact thermal models consider a combination of interconnected resistors and current mirrors (heat sources). The temperature at any point within a 3-D stack can be determined by interpolating between nodal positions of the parallelepiped structure. Mesh-based thermal models are usually integrated into thermal field solvers due to the greater accuracy and the ability to consider the thermal properties of different materials. These models, however, are computationally expensive.

4.2 Thermal management techniques

An increase in temperature can severely affect the performance of a 3-D IC. Performance degrades as the resistance of the wires increases, leakage power increases (due to an exponential dependence on temperature), and the drive current of the transistors is reduced. In addition, higher operating temperatures degrade the reliability of an integrated circuit, potentially reducing circuit lifetime.

Thermal management in 3-D integrated circuits adds an additional order of complexity as compared to 2-D ICs. Although heat transfer in a 2-D IC is a three-dimensional process,

the heat originates from the transistors and wires which can be viewed as two-dimensional heat sources. The increase in complexity arises when the stacked die include sources in different planes in the z-direction. A more fundamental difference between 2-D and 3-D ICs is that 3-D ICs suffer from a higher power density. Removing heat from those dies farthest from the heat sink is non-trivial, requiring novel physical and circuit design techniques to achieve this objective.

Several passive and active techniques to remove heat have been developed that consider the higher temperatures and thermal gradients within a 3-D IC. In the following subsections, thermal TSVs and microfluidic cooling are described as a means to remove heat from those planes deep within a 3-D stack. In addition, thermal aware floorplanning can be used to reduce hotspots based on local power densities.

4.2.1 Thermal through silicon vias

A thermal through silicon via is a relatively novel concept that emerged during the past 15 years. The concept of a thermal via to reduce thermal effects was first applied to the design of packaging and circuit board technologies [234]. Early work on packaging and circuit board thermal vias produced two significant results that are applicable to 3-D IC thermal TSVs, namely: 1) increasing the size or number of thermal vias removes more heat at a cost of less area for routing [235], and 2) vertical heat transfer is more efficient through the thickness of the package/board than spreading the heat laterally [236].

The greater thermal effects in 3-D ICs is primarily due to the greater insulation, and

therefore higher thermal resistance due to the increased number of dielectric layers. Interconnect structures, such as thermal vias and in-plane interconnect, create thermal conduits to spread and remove heat from the hot spots by decreasing the effective thermal resistance of the bonding and dielectric materials within a 3-D system [234, 237–240]. The first proposed use of thermal TSVs (TTSV), or dummy thermal vias, was applied to reduce the effective thermal resistance to the substrate [241]. These dummy thermal vias are completely electrically isolated with no current flow.

An important focus is to design efficient heat conduction paths within an integrated circuit to reduce the peak temperature at localized hot spots. Research describing methods to integrate thermal vias on-chip to reduce thermal issues has been reported [225, 230, 241, 242]. Algorithms to properly place TTSVs within the on-chip white space during 3-D IC floorplanning have recently been reported [234, 238–240, 243, 244]. The design objective for these thermal via placement algorithms is to place thermal TSVs at the proper density while minimizing the effect on both intra- and inter-plane interconnect routing. The algorithms adjust the density of the TTSVs to satisfy thermal and routing constraints [234, 238]. The placement of TTSVs within a 3-D IC depicting the loss of routing area is schematically shown in Figure 4.6.

Inserting thermal TSVs significantly affects the thermal conductivity within a 3-D IC. TTSVs facilitate the transfer of heat in the vertical direction by increasing the thermal conductivity of the vertical heat transfer paths. Thermal via insertion satisfies a number of design objectives [234] including 1) reducing the maximum or average thermal gradients,

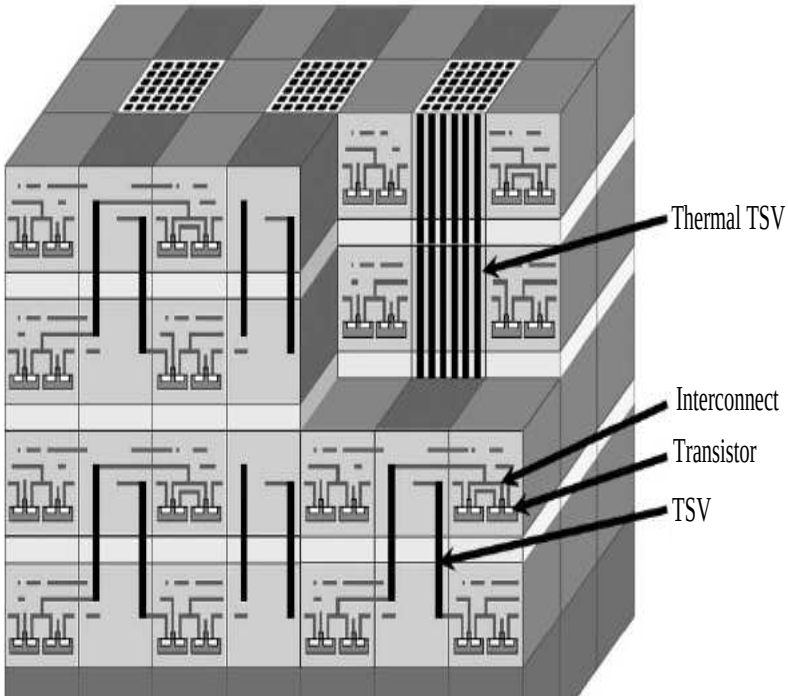


Figure 4.6: Placement of thermal TSV bundles within a 3-D stack for removing heat from internal device planes [234].

2) reducing the maximum or average on-chip temperature, and/or 3) reducing the maximum or average thermal via density [17]. The number of required thermal TSVs for a 3-D IC is determined from the temperature at specific nodes within the volume. The nodal temperatures within a 3-D grid are determined by solving the heat diffusion equation, (4.1), through finite-element based methods [17, 234]. An iterative approach is applied to determine the thermal conductivity of the elements, similar to that shown in Figure 4.5(b). When initializing the optimization procedure, an ideal value, which depends upon the target objective as well as the initial temperature profile, are used to characterize the thermal gradients within each of the thermal via regions. The thermal conductivity of these regions is set to the minimum value, a worst case value equivalent to no thermal vias in a region. The thermal conductivity of the regions is iteratively modified, and the temperature at the nodes is recalculated. The algorithm terminates when any additional change to the thermal conductivity produces minimal improvements to the targeted thermal objective [17]. A flow diagram of the thermal TSV placement algorithm is shown in Figure 4.7.

4.2.2 Floorplanning with thermal considerations

As previously mentioned, the primary causes of an increased thermal profile within 3-D ICs as compared to 2-D ICs is the increased power density and thermal resistance due to the greater distance to the heat sink. An additional cause of hotspot formation is the alignment of highly active circuit blocks in the vertical z-direction, where heat generated by each block produces a local hotspot greater in temperature than if only a single circuit block is active. This effect is in addition to any hotspot formation caused by the placement of highly

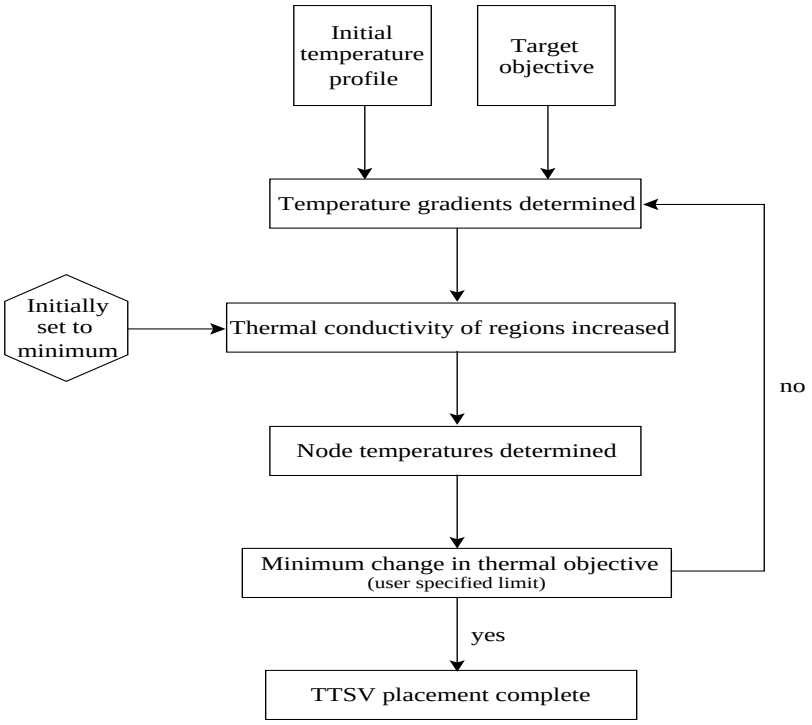


Figure 4.7: Flow diagram of the thermal TSV placement algorithm.

active circuit blocks in the xy-plane such as seen in standard 2-D circuits. An effective technique to minimize the formation of these thermal hotspots is to move the circuit blocks to reduce the vertical overlap of the highly active circuits with adjacent blocks that are particularly sensitive to thermal gradients [245–247]. The effect on the intra-plane thermal profile must be considered when moving these vertically aligned blocks.

In 2-D floorplanning, typical objective functions include minimizing the total area of a circuit or the total wirelength of the interconnects. 3-D IC objective functions must also minimize the number of interplane vias to decrease the fabrication cost and silicon area, while also accounting for thermal-driven design constraints [17, 243, 247, 248]. An objective function that includes both 2-D objectives and 3-D objectives is

$$cost = c_1 \cdot w_l + c_2 \cdot area + c_3 \cdot via_i + c_4 \cdot g(T), \tag{4.10}$$

where c_1, c_2, c_3, and c_4 are weight factors for, respectively, the normalized wirelength w_l, physical area, number of interplane vias v_i, and block cost for a specific temperature $g(T)$ [243, 248]. The cost objective function described by (4.10) does not minimize the temperature of a circuit but rather constrains the temperature within a specified level.

A simulated annealing algorithm has been applied to perturb the block placements until the solution converges to the desired freezing temperature [249]. Legal perturbations for 3-D IC floorplanning include 1) *rotation*, block level rotation, 2) *swap*, block swaps within one device plane, 3) *reverse*, exchange of the relative position of two blocks within a device plane, 4) *move*, moves a block within a device plane, 5) *interlayer swap*, swaps two blocks between two different device planes, 6) *z-neighbor swap*, swaps two blocks from two different device planes that are in close proximity, and 7) *z-neighbor move*, moves a block from one device plane to another plane while maintaining the current x-y position [248]. The last three operations are unique to 3-D ICs. The two interplane block swapping operations and the move operation are only executed if movement of the blocks does not significantly alter the x-y coordinates of the block.

As the blocks are moved around by the simulated annealing algorithm, the differential equation described by (4.1) is recomputed. With finite difference approximations, (4.1) can be written in matrix form as $T = R \cdot P$, where R is the node-to-node thermal resistance matrix, T is the node temperature matrix, and P is the node power dissipation matrix. Any modification to the placement of the blocks causes all three of these matrices to change.

A single block movement requires recomputing the resistance matrix R and power density matrix P, producing a new T matrix that is used to determine if the thermal objectives are satisfied. This recursive algorithm is computationally expensive, and therefore, techniques such as the force-directed method for block placement [250] and analytic block placement methods [251] have been developed to lower the computational time.

4.2.3 Active cooling through microchannels

Convective heat transfer through forced fluid flow is an effective means to remove heat [221, 222, 252, 253]. In addition to thermal-aware design, innovative heat removal techniques are being pursued to mitigate the high thermal profile within 3-D integrated circuits. Although package level microfluidic cooling for 2-D ICs was developed in the 1980's [254, 255], microfluidic based cooling through microchannels for application to 3-D integrated circuits has been considered only during the past ten years [256–261]. Compact thermal models have recently been developed that include the heat removal process of microfluidic channels [262–264].

The fundamental benefit of microfluidic channels is extracting heat in the lateral direction. As previously mentioned, the primary heat conduction path with conventional heat transfer is in the vertical z-direction. By incorporating microchannels between and through the different device planes of a 3-D IC, horizontal conductive paths for removing heat are made available. The microchannels are conduits with widths ranging from 1 to 1000 μm through which fluid (or gas) is forced to flow [228]. The electrical energy dissipated by the circuit as heat is absorbed by the fluid, which carries this energy to a heat exchanger, where

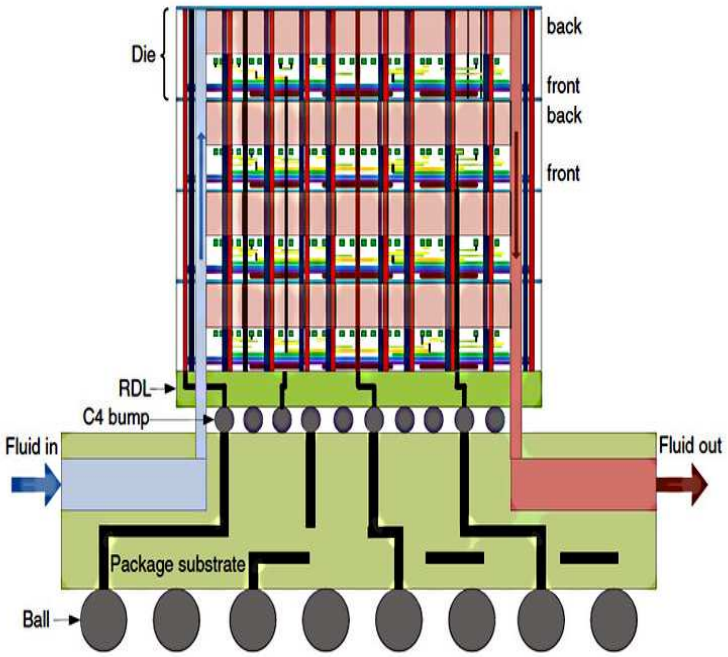


Figure 4.8: Microfluidic cooling within a 3-D integrated circuit with cool water flowing from the left and heated water removed from the right [78].

the energy is transferred out of the 3-D system. Once the liquid returns to the ambient temperature, it is recirculated through the 3-D IC. An example of the exchange of heat through microfluidic channels is depicted in Figure 4.8, where cool fluid is inserted at the left side of the stack, flows through four device planes, and exits on the right side of the stack [78].

This type of heat exchange system can be categorized as either single-phase or two-phase. The difference between the two systems is that a single-phase exchanger transfers heat primarily through convection [254, 255, 265], while a two-phase exchanger uses both convection and evaporation [266]. A heat sink with a fan is an example of a conventional gas-phase forced convection system [228].

The placement of the microfluidic channels depends on the particular fabrication process. It is possible to embed the microchannels underneath the cell rows, between the top

layer of metallization of one plane and the buried oxide of another plane [228]. The flow of fluid, in this case, is parallel to the cell rows. As the microchannels occupy additional area, the interconnect area for routing is once again more highly congested. Placement of the electrical interconnect and microchannels must therefore be designed concurrently.

The benefit of microchannel cooled 3-D integrated circuits is the reduction in the thermal resistance as compared to conventional air cooled devices. A reduced thermal resistance translates to improved IC performance and lower power consumption. A simple model, assuming a uniform power density across each device plane, that quantifies the relationship between cooling and the overall IC temperature is provided in (4.11) and (4.12) [259, 267]. The effective thermal resistance R_{th} is lowered to account for the active cooling of the 3-D IC due to the microfluidic channels.

$$P = \frac{T - T_{amb}}{R_{th}}, \tag{4.11}$$

$$P = \alpha C_{total} [V_{dd}(T)]^2 f + N_{gates} V_{dd}(T) I_{leak,o} e^{\frac{V_t(T)+\Delta V_t}{nkT/q}}, \tag{4.12}$$

where T is the on-chip temperature, T_{amb} is the ambient temperature, R_{th} is the thermal resistance, P is the total power dissipation, α is the activity factor, f is the clock frequency, C_{total} is the total capacitance, I_{leak} is the leakage current coefficient, n is the subthreshold slope, and ΔV_t is the variation in threshold voltage. The optimal supply voltage $[V_{dd}(T)]$ and threshold voltage $[V_t(T)]$ are chosen to minimize the power for a specific technology and circuit speed [268]. The values derived for V_{dd} and V_t in [268] are substituted into (4.12) [259, 267], producing a set of two equations, (4.11) and (4.12), and two unknowns,

the power and temperature. Solving and simplifying these expressions produces a cubic equation which yields closed-form expressions for the total power and temperature for a specific thermal resistance, device technology, and frequency [259, 267].

4.3 Summary

Thermal effects in 3-D integrated circuits present thermal management issues that if not properly addressed during design and floorplanning will lead to deleterious on-chip hotspot formation. The primary reasons for the anticipated increase in temperature are much higher power dissipation, a smaller footprint, and an increase in the length of the thermal conductive paths to the heat sink. There are two essential components to properly manage thermal effects in 3-D integrated circuits. First, proper characterization and modeling of hotspot formation and the thermal propagation paths are required. Second, design techniques and physical methods that reduce the peak temperature and minimize thermal gradients among the device planes within a 3-D IC need to be applied. Characterization and modeling of thermal effects as well as heat mitigation techniques are therefore critical research topics for 3-D ICs.

A review of current research on the characterization and compact modeling of thermal effects in 3-D ICs is provided in this chapter. Research on modeling heat flow within a 3-D stack is described. Both compact and mesh-based thermal models are discussed. The objective is to produce highly accurate and computationally inexpensive thermal models that include the effects of multi-plane thermal coupling. Passive techniques such as inserting thermal TSVs and thermal aware floorplanning are discussed. In addition, active

liquid cooling through microchannels is also reviewed. Design methodologies and cooling techniques that manage thermal effects in 3-D ICs increase circuit complexity. The primary objective is to ensure that the on-chip temperature satisfies specifications with minimal cost in design complexity and circuit area.

Chapter 5

Electrical Properties of Through Silicon Vias

This chapter on the electrical properties of a through silicon via is composed of two distinct sections. The first section describes the electrical characterization of a TSV based on the MIT Lincoln Laboratory second multi-project wafer (MPW) sponsored by the Defence Advanced Research Project Agency (DARPA) and which was code named 3DM2. A basic understanding of the electrical properties of these TSVs is necessary for both modeling the clock distribution network described in Chapter 6 and for analyzing the 3-D power delivery networks described in Chapter 7. The second component of this chapter focuses on closed-form expressions of the TSV resistance, capacitance, and inductance.

Electrical characterization of the resistance, capacitance, and inductance of inter-plane 3-D vias is presented in this chapter. Both capacitive and inductive coupling between multiple 3-D vias is described as a function of the separation distance and plane location. The effects of placing a shield via between two signal vias is investigated as a means to limit the capacitive coupling. The location of the return path is examined to determine a preferable placement of a 3-D via to reduce the overall loop inductance. Based on the extracted

resistance, capacitance, and inductance, the *L/R* time constant is shown to be much larger than the *RC* time constant, demonstrating that the 3-D via structure is inductively limited rather than capacitively limited.

Closed-form expressions of the resistance, capacitance, and inductance for inter-plane 3-D vias are also presented in this chapter. The closed-form expressions account for the 3-D via length, diameter, dielectric thickness, and spacing to ground. 3-D numerical simulations are used to quantify the electromagnetic behavior of the resistance, capacitance, and inductance for comparison with the closed-form expressions, revealing good agreement between simulation and the physical models. The maximum error for the resistance, capacitance, and inductance is less than 8%.

The work reported in this chapter examines the *RLC* impedances of a 3-D via based on the MIT Lincoln Lab 3-D integration process [89, 269]. Background information describing the MIT process and the Ansoft Quick 3-D toolset used in the via analysis is presented in Section 5.2. The extracted resistance, capacitance, and inductance of a single via are presented in Section 5.3. Capacitive and inductive coupling between two 3-D vias is examined in Section 5.4 for different separation distances and structural topologies. Inter-plane coupling is also analyzed in Section 5.4. This analysis is followed in Section 5.5 by an examination of the effects of placing a 3-D shielding via between two signal vias. The effect on inductive coupling and the loop inductance as a result of the via placement is discussed in Section 5.6. This section concludes the first portion of the chapter on the electrical characterization of the TSV used within the MIT Lincoln Laboratory 3-D fabrication process.

The second portion of the chapter focuses on closed-form expressions of the electrical characteristics of the through silicon via. Before the closed-formed expressions of the resistance, capacitance, and inductance are presented, the technique of modeling the 3-D vias as simple cylinders without a top and bottom copper landing is reviewed in Section 5.7. Background information describing the geometry of the 3-D vias used for comparison with the models are also presented in Section 5.7. Closed-form expressions of the through silicon via resistance for both DC and higher frequencies are provided in Section 5.8. Expressions characterizing the 3-D via inductance at both DC and high frequency are provided in Section 5.9. A model of the mutual inductance between two 3-D vias is also provided in this section. Models of the 3-D via capacitance are provided in Section 5.10. A few concluding remarks are provided in the final section of the chapter.

5.1 3-D via electrical modeling

The primary technological innovation required to exploit the benefits of 3-D integration is the development of a 3-D through silicon via (TSV) technology. Much work has been achieved to properly characterize and model these interplane through silicon vias.

Accurate closed-formed models of the 3-D via impedance provide an efficient method to characterize the performance of signal paths containing through silicon vias. These closed-formed expressions are similar in form to the models developed in [270] and [271] that characterize the capacitance and inductance [143] of on-chip interconnect for VLSI circuits. Most previous work characterizing 3-D vias has focused on bulk silicon and emphasized the experimental extraction of the via resistance and capacitance. Due to the large

variation in the 3-D via diameter, length, dielectric thickness, and fill material, a wide range of measured resistances, capacitances, and inductances has been reported in the literature. Single 3-D via resistance values vary from 20 mΩ to as high as 350 Ω [207, 272–276], while reported capacitances vary from 2 fF to over 1 pF [126, 150, 207]. A few researchers have reported measured via inductances that range from as low as 4 pH to as high as 255 pH [96, 207, 273]. Alternatively, preliminary work on modeling 3-D vias has primarily focused on the resistance and capacitance of simple structures to verify measured *RLC* impedances while providing some physical insight into 3-D via to 3-D via capacitive coupling [126, 150, 207, 272]. In addition, preliminary work in electrical modeling of bundled TSVs has been reported in [277–279]. This chapter expands on this early work, and presents closed-formed expressions for the resistance, inductance, and capacitance of the through silicon vias (3-D vias), accounting for via length, via diameter, dielectric thickness, and fill material. The models and closed-form expressions described in this chapter add new insight towards shielding both capacitive and inductive coupling amongst TSVs, and provide the necessary electrical characterization of the TSV.

5.2 MIT fabrication and Ansoft tools

A review of the MIT Lincoln Laboratory TSV based 3-D integration process is provided in this section. Background information describing the Ansoft Quick 3-D tool for full wave electromagnetic simulation is also described.

5.2.1 MIT Lincoln Laboratory 3-D integration process

A basic understanding of the MIT Lincoln Laboratory 3-D integration process is provided as a means to understand the via topologies being evaluated. Lincoln Laboratory has established a 0.18 μm low power, fully depleted silicon-on-insulator (FDSOI) CMOS process where three independently patterned wafers are physically bonded to form a 3-D integrated structure [89, 269]. Each wafer, described as either a plane or tier, has three metallization layers for routing. The top two planes have an additional backside metal for both routing and to form the 3-D vias. The critical dimensions of a 3-D via are depicted in Figure 5.1 [89]. Note that there are two inter-plane 3-D vias present in the figure, one that passes from Tier 1 (T1) to Tier 2 (T2) while the other passes from T2 to Tier 3 (T3). By overlaying the 3-D vias in this fashion, a signal may propagate from the bottom plane to the top plane. It is also acceptable for either the top via (T2 to T3) or bottom via (T1 to T2) to propagate a signal. Note also that the bottom plane (T1) includes a 675 ± 25 μm thick silicon substrate, not shown in the figure, which is used for wafer handling after the top two planes are bonded to T1. As a final note, the total length of a 3-D via starting from plane T1 and terminating on plane T3 is 17.94 μm, requiring wafer thinning of planes T2 and T3 to approximately 10 μm [89]. Each 3-D via is surrounded by a combination of two dielectric layers, SiO_2 and tetraethylorthosilicate (TEOS). Both of these dielectric layers have similar relative permittivities, 3.9 and 4.0, respectively. For simulation purposes, the average of the two permittivities is used in the Ansoft Quick 3-D toolset.

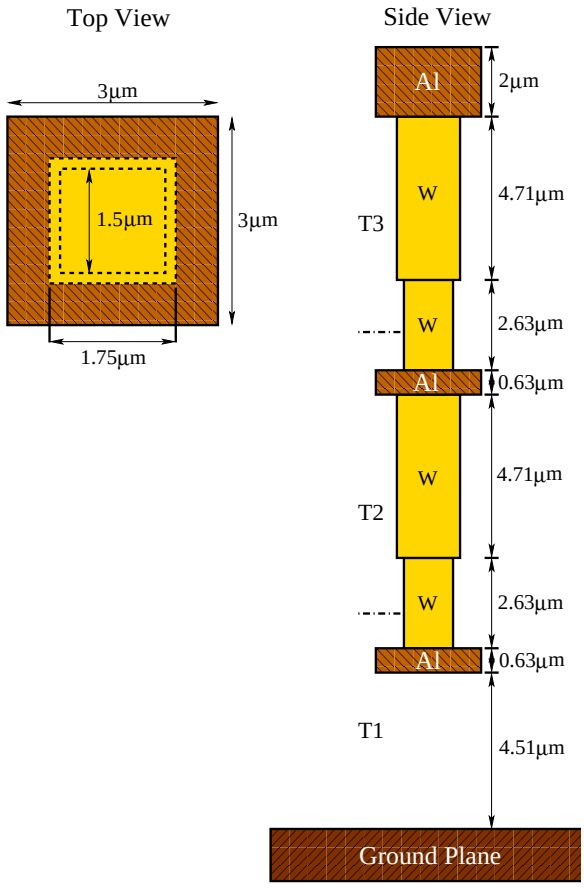


Figure 5.1: Coupling capacitance between two 3-D vias over a ground plane (D = 20 µm).

5.2.2 Ansoft Quick 3-D Electromagnetic-field Solver

Ansoft Quick 3-D (Q3D) software is used to examine the *RLC* impedances, coupling, and shielding behavior of the through silicon vias [280]. Quick 3-D is a 3-D/2-D quasi-static electromagnetic-field simulator used to extract parasitic impedances and electrical parameters. The tool solves Maxwell's equations by applying the Finite Element Method (FEM) [281] and the Method of Moments (MOM) [282,283] to compute the *RLC* or *RLCG* parameters of a 3-D structure. With regard to the inductance, Q3D is used to determine the asymptotic DC and high frequency (f_{asym}) inductance depicted in Figure 5.2. Q3D simulations do not provide the inductance for frequencies between the DC and high frequency extrema; therefore, the closed-form expression for inductance is applicable for frequencies lower and higher than the transitional frequency range (200 to 800 MHz for this particular technology).

5.3 *RLC* extraction of a single 3-D via

Multiple 3-D via configurations have been investigated with and without the ground plane, as illustrated in Figure 5.1. These configurations include T1 to T2, T2 to T3, and T1 to T3 signal propagation paths. The 2 μm thick aluminum ground plane is located at the top of the silicon substrate. Results from these analyses are listed in Table 5.1.

Tabulated results indicate that, in all cases, the *L/R* time constant is approximately four orders of magnitude greater than the *RC* time constant. In addition, the ground plane does

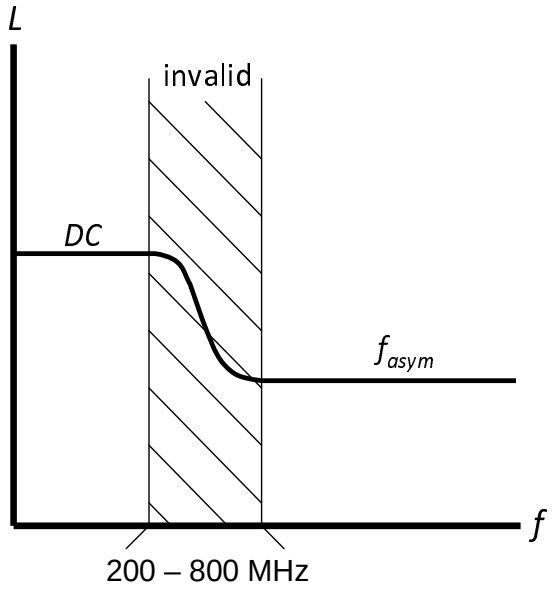


Figure 5.2: Frequency range for which Q3D models and closed-form inductance expressions are applicable.

not affect either the resistance or inductance. The total via capacitance, however, is significantly affected, with the capacitance increasing by as much as 37%. For this reason, the *L/R* time constant remains relatively unchanged with a maximum variation of 1.4%, whereas the *RC* time constant varies by as much as 37%.

5.4 *RLC* coupling between two 3-D vias

Once a single 3-D via is electrically characterized, the capacitive and inductive coupling between two 3-D vias can be investigated. This analysis has been performed on several configurations, each with increasing separation between the two 3-D vias. A ground plane is present in all cases. A graphical depiction of the simulation setup and the extracted *RLC* impedances are illustrated in Figure 5.3a. The extracted impedances for varying separation

Table 5.1: *RLC* impedances and related time constants for a single 3-D via.

	Configuration	C_t (fF)	C_{gc} (fF)	DC *RL*		AC *RL* (1 GHz)		DC time constants		AC time constants	
				R (mΩ)	L (pH)	R (mΩ)	L (pH)	τ_{RC} (fsec)	$\tau_{L/R}$ (psec)	τ_{RC} (fsec)	$\tau_{L/R}$ (psec)
With ground plane	T1T2	1.43	1.28	147.55	3.93	165.92	2.93	0.21	26.62	0.24	17.68
	T2T3	1.34	1.39	152.14	4.68	171.69	3.52	0.20	30.74	0.23	20.52
	T1T3	2.04	1.89	296.17	10.6	335.10	8.51	0.60	35.87	0.68	25.38
Without ground plane	T1T2	1.15	-	147.62	3.92	165.85	2.92	0.17	26.56	0.19	17.58
	T2T3	1.16	-	151.35	4.69	170.33	3.45	0.18	31.01	0.20	20.25
	T1T3	1.49	-	296.44	10.7	335.21	8.45	0.44	35.92	0.50	25.22

between the two 3-D vias are listed in Table 5.2.

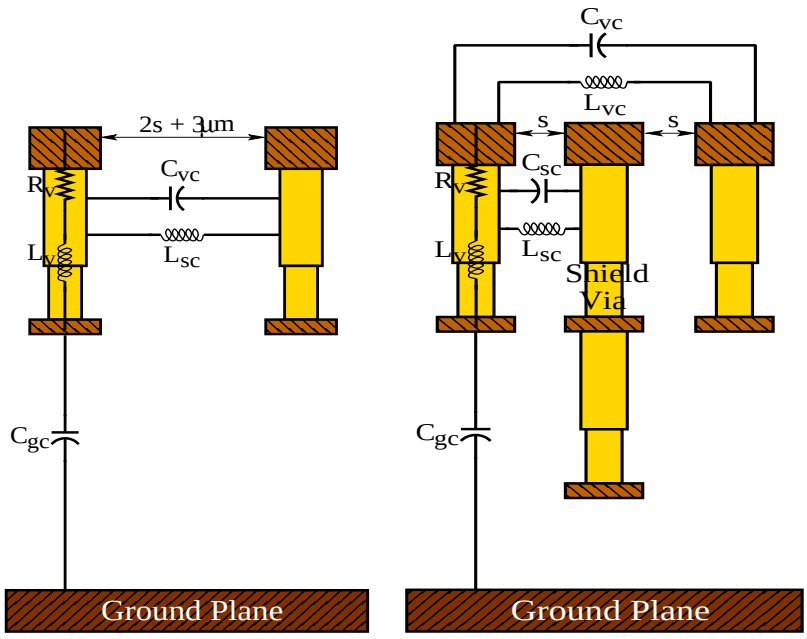
The tabulated results reveal several interesting characteristics in addition to the expected trend of decreasing coupling capacitance (C_{vc}) and inductance (L_{vc}) with increasing via separation. As the separation between the two 3-D vias increases, the AC resistance, AC self-inductance (L_v), and capacitance of the vias approach the extracted values listed in Table 5.1 for a single 3-D via. The DC resistance and self-inductance (L_v) are not affected by the physical separation, and are therefore equivalent to a single 3-D via. Simulations of the inter-plane coupling reveal a similar characteristic where the *RLC* impedance approaches the values extracted for a single 3-D via with increasing spacing. Assuming an ideal oxide-to-oxide bond between wafers, the inter-plane coupling (the row labeled as T2T3-T1T2 in Table 5.2) indicates that this form of coupling can not be neglected. Comparing the results for T2T3-T2T3 with T2T3-T1T2, the extracted electrical parameters converge to a similar value with increasing spacing, which supports the need to properly model the inter-plane coupling.

The loop inductance is also included in Table 5.2. The loop inductance is,

$$L_{loop} = L_{11} + L_{22} - 2 \times L_{21}, \tag{5.1}$$

where L_{11} and L_{21} are, respectively, L_v and L_{vc} from Table 5.2. The self-inductance of the second via L_{22} is not included in either Figure 5.3 or Table 5.2, but can be determined from (5.1). If the two 3-D vias connect the same two planes, (5.1) reduces to

$$L_{loop} = 2 \times L_{11} - 2 \times L_{21}, \tag{5.2}$$



(a) (b)

Figure 5.3: Circuit model for *RLC* extraction of (a) two 3-D vias, and (b) two 3-D vias with a shield via between the two signal vias.

since both vias have approximately the same self-inductance.

5.5 Effects of 3-D via placement on shielding

A primary focus of section 5.4 is to examine the extracted coupling capacitance and inductance between two 3-D vias as a function of the physical spacing. The effect of inserting a shield via between two signal vias on the coupling capacitance and inductance is discussed in this section. The simulation setup is depicted in Figure 5.3b, and the resulting extracted parameters are listed in Table 5.3.

The reduction in capacitance is determined by noting the difference in coupling capacitance between two 3-D vias with and without a shield. The coupling capacitance, assuming

Table 5.2: Extraction of *RLC* parameters for increasing separation between two 3-D vias.

Configuration	Separation (μm)	C_t (fF)	C_{vc} (fF)	C_{gc} (fF)	DC *RL*				AC *RL* (1 GHz)			
					R (mΩ)	L_v (pH)	L_{vc} (pH)	L_{loop} (pH)	R (mΩ)	L_v (pH)	L_{vc} (pH)	L_{loop} (pH)
T2T3-T2T3	1	2.05	1.16	0.87	152.36	4.69	2.05	5.27	173.71	3.40	1.63	3.55
	3	1.58	0.60	0.97	152.12	4.70	1.52	6.35	172.12	3.49	1.22	4.54
	5	1.45	0.42	1.05	152.23	4.69	1.20	6.99	171.93	3.51	0.97	5.08
	7	1.40	0.31	1.04	152.03	4.69	0.98	7.41	171.88	3.52	0.80	5.44
	9	1.34	0.25	1.04	152.04	4.69	0.84	7.71	171.58	3.53	0.68	5.68
	11	1.32	0.20	1.10	152.31	4.69	0.72	7.94	171.89	3.53	0.59	5.88
	13	1.31	0.17	1.11	152.10	4.70	0.64	8.12	171.73	3.53	0.52	6.01
T1T3-T1T3	1	3.2206	1.8429	1.3836	297.92	10.662	5.3366	10.63	341.316	8.199	4.5435	7.3176
	3	2.4023	1.0010	1.4351	297.49	10.654	4.1699	12.977	237.81	8.3573	3.5556	9.6252
	5	2.2407	0.6628	1.5088	297.50	10.654	3.4101	14.489	336.938	8.4217	2.9282	10.982
	7	2.1416	0.4852	1.8334	297.25	10.660	2.8868	15.530	336.329	8.4463	2.4878	11.914
	9	2.0860	0.3792	1.6312	297.45	10.647	2.4975	16.298	336.382	8.4557	2.1554	12.594
	11	2.0650	0.3038	1.7465	297.47	10.657	2.1993	16.913	336.42	8.4648	1.8871	13.158
	13	2.056	0.2479	1.7369	297.56	10.652	1.9624	17.396	336.461	8.4714	1.6882	13.558
T2T3-T1T2	1	1.57	0.51	1.02	152.27	4.69	1.07	6.49	171.91	3.52	0.90	4.68
	3	1.43	0.33	1.06	152.31	4.69	0.91	6.80	171.88	3.52	0.76	4.95
	5	1.39	0.26	1.12	152.20	4.69	0.79	7.04	171.77	3.52	0.65	5.16
	7	1.37	0.20	1.14	152.17	4.70	0.69	7.24	171.69	3.52	0.57	5.33
	9	1.37	0.16	1.16	152.28	4.70	0.62	7.39	171.79	3.52	0.51	5.46
	11	1.28	0.14	1.10	152.14	4.70	0.55	7.52	171.65	3.52	0.45	5.57
	13	1.35	0.11	1.22	152.32	4.69	0.50	7.62	171.85	3.53	0.41	5.66

a 1 μm space with a shield, is compared to a 5 μm space without a shield (see Table 5.2). The capacitive coupling (C_{vc}) between two signal vias is reduced by as much as 77% by including a third shield via. In addition, the effectiveness of the shield diminishes with increasing spacing, as additional electrical field lines terminate on the second signal line. Alternatively, the inductive coupling (L_{vc}) between the two signal lines increases due to the high mutual inductance between the shield and signal vias. The negative inductance listed in Table 5.3 indicates this increase in inductive coupling. Note that the increase in inductance when an inserted shield via is present only applies to the AC inductance, and that the maximum increase in inductive coupling is about 4.3%. As previously mentioned, an increase in the mutual inductance produces a decrease in the loop inductance. The effect of the return path on both the DC and AC inductance is described in the following section.

5.6 Effect of the return path on the 3-D via inductance

The return path can significantly affect the induced loop inductance. In addition, the distance between the signal via and the return path also contributes to the loop inductance. As the space between the signal and return path increases, the mutual inductance (L_{21}) decreases. From (5.1), a decrease in the mutual inductance increases the loop inductance. An analysis of a three T2T3 3-D via structure illustrates the effect of the return path: with the return path through the middle via (see Figure 5.4a) and with the return path through a side via (see Figure 5.4b). The inductance listed in Table 5.4 only includes a single row of the 3 x 3 matrix formed with these three conductors, corresponding to the self- (L_{11}) and mutual (L_{12} and L_{13}) inductance of conductor 1 shown in Figure 5.4. The remaining two

Table 5.3: L and C shielding with the addition of a third shield via.

Configuration	Separation (µm)	C_t (fF)	C_{vc} (fF)	% C_{vc} shielded	DC L			AC L		
					L_v (pH)	L_{vc} (pH)	% L_{vc} reduced	L_v (pH)	L_{vc} (pH)	% L_{vc} reduced
T2T3-T1T3-T2T3	1	2.07	0.107	74.39%	4.68	1.19	0.63%	3.40	1.01	-4.30%
	2	1.73	0.093	70.29%	4.69	0.98	0.17%	3.46	0.82	-2.52%
	3	1.58	0.079	68.94%	4.68	0.83	0.40%	3.50	0.69	-1.33%
	4	1.50	0.069	65.79%	4.68	0.72	-0.01%	3.51	0.59	-1.05%
	5	1.39	0.056	66.13%	4.69	0.64	-0.16%	3.52	0.52	-0.39%
T1T3-T1T3-T1T3	1	3.2267	0.1643	75.21%	10.665	3.4051	0.147%	8.1160	3.0119	-2.858%
	2	2.6774	0.1393	71.29%	10.652	2.8886	-0.062%	8.2591	2.5264	-1.551%
	3	2.4483	0.1177	68.97%	10.655	2.4992	-0.068%	8.3323	2.1710	-0.724%
	4	2.3127	0.0993	67.31%	10.645	2.2009	-0.073%	8.3657	1.8853	0.095%
	5	2.2023	0.08761	64.66%	10.667	1.9636	-0.061%	8.3679	1.6768	0.675%
T2T3-T1T3-T1T3	1	2.07	0.059	76.92%	4.69	0.79	0.37%	3.41	0.67	-2.74%
	2	1.74	0.054	73.35%	4.69	0.69	0.40%	3.47	0.58	-2.00%
	3	1.59	0.049	69.58%	4.69	0.61	0.24%	3.50	0.51	-1.08%
	4	1.51	0.044	67.68%	4.69	0.55	0.19%	3.51	0.46	-1.00%
	5	1.42	0.037	66.71%	4.68	0.50	0.23%	3.52	0.41	-0.43%

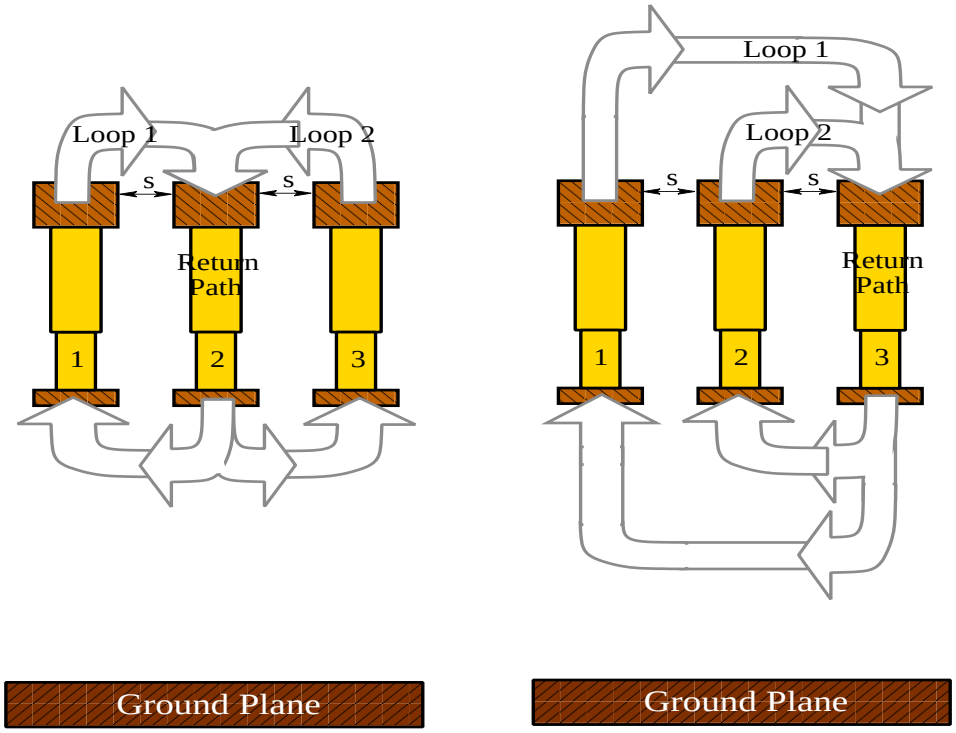


Figure 5.4: Effect of a return path on the loop inductance for (a) a return path placed on 3-D via 2, and (b) a return path placed on 3-D via 3.

rows are the self- and mutual inductance of the other two conductors.

These results are listed in Table 5.4, and are briefly addressed here. Placing the return path between the other two 3-D vias induces two symmetric inductance loops, as indicated by the similar loop inductances in the row labeled, ”return path conductor 2” in Table 5.4. Alternatively, placing the return path on either side of the three 3-D via system produces a 32% increase in the loop inductance between the two side vias. The mutual loop inductance between the two loops (L_{loop21}) increases by as much as 96% when the return path is placed on one of the side vias rather than the central via. Based on these results, proper placement of the return paths is critical in minimizing the loop inductance of a 3-D via.

Table 5.4: Inductive coupling and loop inductance for three T2T3 3-D vias.

		DC L			AC L (1 GHz)		
	Separation (μm)	L_{11} (pH)	L_{21} (pH)	L_{31} (pH)	L_{11} (pH)	L_{21} (pH)	L_{31} (pH)
Partial inductances	1	4.69	2.05	1.20	3.36	1.57	0.99
	2	4.69	1.75	0.98	3.44	1.36	0.81
	3	4.69	1.51	0.84	3.47	1.19	0.68
	4	4.69	1.34	0.72	3.49	1.06	0.59
	5	4.69	1.19	0.65	3.50	0.95	0.52
		DC L			AC L		
	Separation (μm)	L_{loop1} (pH)	L_{loop2} (pH)	L_{loop21} (pH)	L_{loop1} (pH)	L_{loop2} (pH)	L_{loop21} (pH)
Return path conductor 2	1	5.28	5.28	1.79	3.51	3.51	1.13
	2	5.90	5.89	2.19	4.11	4.10	1.48
	3	6.36	6.36	2.50	4.52	4.52	1.74
	4	6.71	6.71	2.75	4.82	4.84	1.93
	5	6.99	6.99	2.94	5.07	5.07	2.09
Return path conductor 3	1	6.70	5.28	3.50	4.75	3.51	2.37
	2	7.41	5.89	3.70	5.25	4.10	2.62
	3	7.71	6.36	3.86	5.57	4.53	2.79
	4	7.93	6.71	3.97	5.80	4.84	2.91
	5	8.11	6.99	4.05	5.97	5.07	2.98

5.7 Modeling 3-D vias as cylinders

The structural complexity of a 3-D via is treated as an equivalent cylindrical structure. A 3-D via typically includes a tapered cylindrical TSV with both a bottom and top metal landing. The electrical parameters produced by the Ansoft electromagnetic simulation tool for a 3-D via represented as a cylinder with a top and bottom copper landing (see Figure 5.5(a)) is compared to an equivalent length cylindrical via without the metal landings (see Figure 5.5(b)). A comparison of the simulation results for these two TSV configurations reveals less than a 7% difference in the resistance, inductance, and capacitance. The only exceptions arise for the DC resistance and high frequency inductance when the aspect ratio is between 0.5 and 1, as listed in Table 5.5. This behavior implies that the use of a simple cylindrical structure without metal landings is sufficient to represent a 3-D via.

The dimensions of the 3-D via are highly technology dependent. For an SOI process, where the buried oxide behaves as a natural stop for wafer thinning, the length $\mathfrak{L}$ of the

Table 5.5: Per cent error between the 3-D via model and the equivalent cylindrical model. Diameter = 20 μm

	R		*L*		*C*
a.r.	DC	1 GHz	DC	f_{asym}	
0.5	41.6	3.6	0	-16.4	6.4
1	16.9	-2.0	0	-13.6	6.2
3	5.4	0.7	-0.04	-0.6	5.4
5	2.9	-0.2	-0.04	-1.4	4.9
7	2.2	-0.4	0.01	-1.4	4.0
9	1.9	0.4	0.02	-0.03	3.1

through silicon vias are much shorter than the bulk counterparts. The diameter D of the vias follows a similar pattern, where an SOI process utilizes a smaller diameter than in bulk technologies. Therefore, these 3-D via models for an SOI process consider diameters of 1, 5, and 10 μm, while the 3-D vias in bulk silicon have diameters ranging between 20 μm and 60 μm. A 20 μm diameter is chosen for the comparison between the 3-D via model and the equivalent cylindrical structure. In all cases, the via lengths maintain an aspect ratio $\mathfrak{L}/D$ of 0.5, 1, 3, 5, 7, and 9. In addition, the dielectric thickness for a bulk process is 1 μm, and the material filling the 3-D vias in both SOI and bulk technologies is tungsten.

5.8 Closed-form resistance model of a 3-D via

Closed-formed expressions of the 3-D via resistance at DC and 1 GHz are presented below, respectively, as (5.3), and (5.4) and (5.5). The DC resistance is only dependent on the length $\mathfrak{L}$, radius $\Re$, and conductivity σ_W of tungsten. The resistance at a frequency of 1 GHz, however, is also dependent on the skin depth, the depth below the surface of a conductor where the current density has dropped by a factor of e [284]. Two guidelines are provided in [284] that ensure the high frequency resistance is valid when considering skin

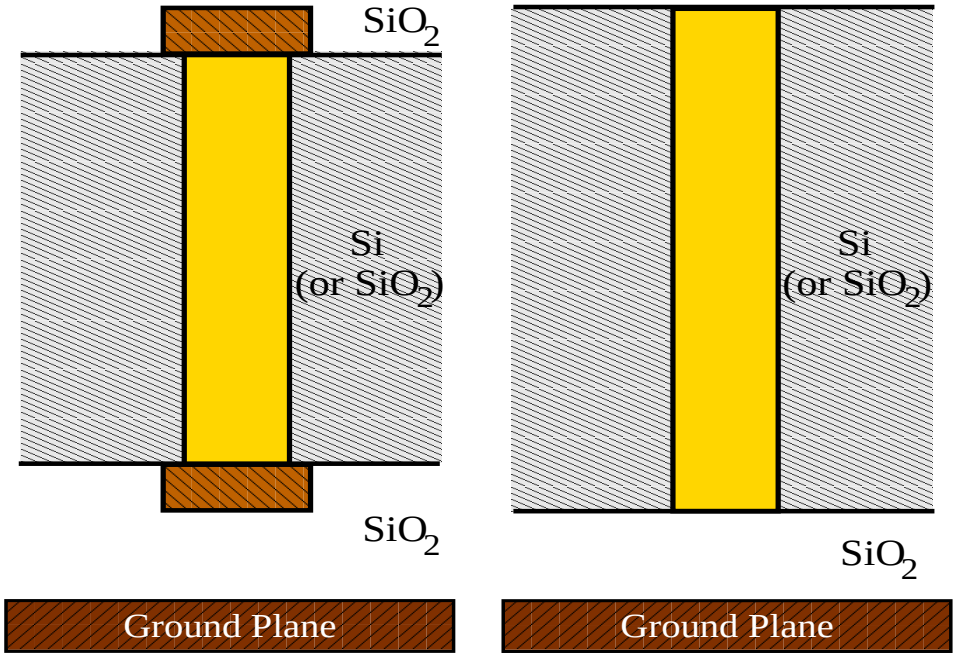


Figure 5.5: 3-D via structure: (a) 3-D via with top and bottom copper landings, and (b) equivalent structure without metal landings.

effects. These two guidelines are: 1) the return paths are infinitely distant, and 2) all radii of curvature and thicknesses are at least three to four skin depths. The second guideline necessitates the use of the fitting parameter α in (5.4) and (5.5) for SOI processes. In addition, α accounts for current losses in the substrate in bulk processes.

$$R_{DC} = \frac{1}{\sigma_W}\frac{\mathfrak{L}}{\pi\Re^2}, \tag{5.3}$$

$$R_{1GHz} = \begin{cases} \alpha\frac{1}{\sigma_W}\frac{\mathfrak{L}}{\pi[\Re^2-(\Re-\delta)^2]}, & \text{if } \delta < \Re \quad (5.4) \\ \alpha\frac{1}{\sigma_W}\frac{\mathfrak{L}}{\pi\Re^2}, & \text{if } \delta \geq \Re. \quad (5.5) \end{cases}$$

The effect of the skin depth is included in the 1 GHz resistance as a reduction in the

cross-sectional area of the 3-D via. The skin depth δ is

$$\delta = \frac{1}{\sqrt{\pi f \mu_o \sigma_W}}, \tag{5.6}$$

where f is the frequency, and μ_o is the permeability of free space. The permeability of free space is used as neither silicon nor silicon dioxide are magnetic materials.

As noted earlier, the α term in (5.4) and (5.5) is an empirical constant used to fit the resistance to the simulations and is

$$\alpha = \begin{cases} 0.0472 D^{0.2831} ln(\frac{\Omega}{D}) + 2.4712 D^{-0.269}, \text{ if } \delta < \Re & (5.7) \\ 0.0091 D^{1.0806} ln(\frac{\Omega}{D}) + 1.0518 D^{0.092}, \text{ if } \delta \geq \Re. & (5.8) \end{cases}$$

For frequencies other than DC and 1 GHz, (5.4) and (5.5) can be adjusted to other frequencies using (5.9). Equation (5.9) has been evaluated at a frequency of 2 GHz. The resistance of a 3-D via with diameters of 5, 20, and 60 μm at frequencies of DC, 1 GHz, and 2 GHz are plotted in Figure 5.6.

$$R_{f_{new}} = (R_{1GHz} - R_{DC}) \sqrt{\frac{f_{new}}{f_{1GHz}}} + R_{DC}. \tag{5.9}$$

The proximity effect [21] is examined through simulations of two 3-D vias over a ground plane. These simulations reveal less than a 0.25% change in the resistance at all examined frequencies as compared to a single 3-D via over a ground plane; these results are therefore not included. This characteristic implies that the proximity of the second 3-D

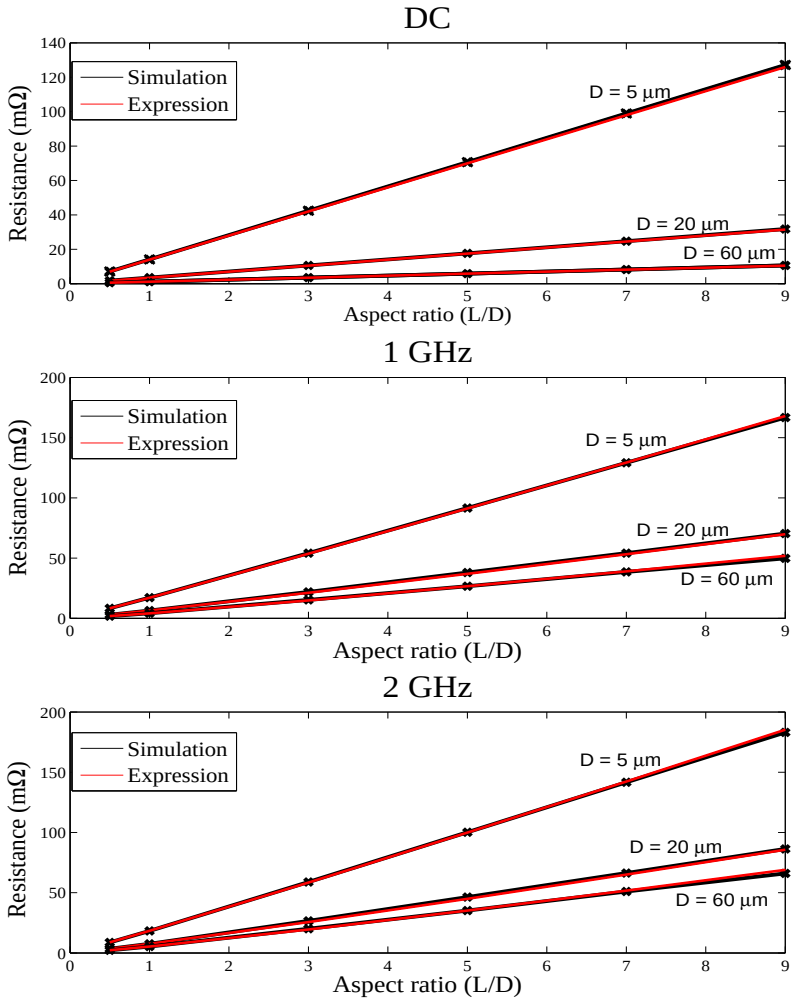


Figure 5.6: Resistance of a cylindrical 3-D via at DC, 1 GHz, and 2 GHz.

via as a return path does not significantly affect the resistance of the first via. The insignificance of the proximity effect is primarily a consequence of the relatively short length of the vias and the small space between the 3-D vias which is at least equal to the diameter of a single 3-D via.

An analysis of the per cent error between simulation and the closed-form resistance expressions as a function of frequency is provided. These results indicate less than 5% error between simulation and the closed-form expressions for all frequencies between DC and 10 GHz. The per cent error for two diameters, 5 and 20 μm, and three aspect ratios, 1, 5, and 9, is depicted in Figure 5.7.

The per cent variation between simulation and the model at frequencies of DC, 1 GHz, and 2 GHz is included in Section 5.11.1, respectively, as Tables 5.6, 5.7, and 5.8. The DC resistance based on the closed-formed expressions produces less than 2% error, and the 1 GHz and 2 GHz resistance produces less than a 5.5% difference from the equivalent model characterized by Q3D.

5.9 Closed-form inductance model of a 3-D via

The DC and high frequency self- and mutual inductance of two equal length TSVs is included in 5.9.1. In Section 5.9.2, the DC and high frequency mutual inductance between two non-equal length TSVs is described. Non-equal length vias are used to approximate the mutual inductance between a single TSV and a stack of TSVs propagating a signal across multiple planes.

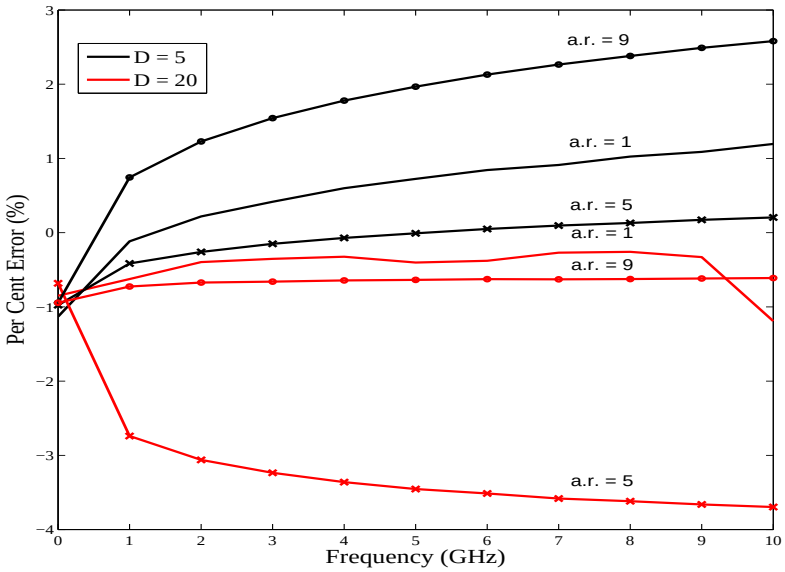


Figure 5.7: Per cent error as a function of frequency for the resistance of a 3-D via (a.r. = aspect ratio).

5.9.1 Inductance of equal length 3-D vias

Expressions for the DC and high frequency partial self- L_{11} and mutual L_{21} inductances are provided, respectively, in (5.10) and (5.11), and (5.12) and (5.13). The modeled DC and high frequency inductances are asymptotic values of the inductance. These expressions do not account for the transition in inductance from low frequency to high frequency (in the 200 to 800 MHz range for this particular technology). Note that the inductance transitions smoothly from low frequency to high frequency, indicating that the inductance of a 3-D via is bound by the DC (upper bound) and asymptotic (lower bound) values within this transitional range. The inductance models are based on [142] with a fitting parameter to adjust for inaccuracies in the Rosa expressions [142]. The expressions derived by Rosa assume that the length $\mathfrak{L}$ is much larger than both the radius $\mathfrak{R}$ (or diameter D) and the pitch P between the conductors. As $\mathfrak{L}$ is not larger than $\mathfrak{R}$ or P in all of the 3-D via structures

examined in this chapter, α and β are used to adjust the partial inductances in [142].

$$DC: \begin{cases} L_{11} = \alpha\frac{\mu_o}{2\pi}[ln(\frac{\mathfrak{L}+\sqrt{\mathfrak{L}^2+\mathfrak{R}^2}}{\mathfrak{R}})\mathfrak{L} + \mathfrak{R} - \sqrt{\mathfrak{L}^2+\mathfrak{R}^2} + \frac{\mathfrak{L}}{4}] & (5.10) \\ L_{21} = \beta\frac{\mu_o}{2\pi}[ln(\frac{\mathfrak{L}+\sqrt{\mathfrak{L}^2+P^2}}{P})\mathfrak{L} + P - \sqrt{\mathfrak{L}^2+P^2}] & (5.11) \end{cases}$$

$$f_{asym}: \begin{cases} L_{11} = \alpha\frac{\mu_o}{2\pi}|ln(\frac{2\mathfrak{L}}{\mathfrak{R}})\mathfrak{L} - 1| & (5.12) \\ L_{21} = \beta\frac{\mu_o}{2\pi}[ln(\frac{\mathfrak{L}+\sqrt{\mathfrak{L}^2+P^2}}{P})\mathfrak{L} + P - \sqrt{\mathfrak{L}^2+P^2}] & (5.13) \end{cases}$$

where

$$\alpha = \begin{cases} 1 - e^{\frac{-4.3\mathfrak{L}}{D}}, & \text{if } f = \text{DC} \quad (5.14) \\ 0.94 + 0.52e^{-10|\frac{\mathfrak{L}}{D-1}|}, & \text{if } f > f_{asym} \quad (5.15) \end{cases}$$

$$\beta = \begin{cases} 1, & \text{if } f = \text{DC} \quad (5.16) \\ 0.1535ln(\frac{\mathfrak{L}}{D}) + 0.592, & \text{if } f > f_{asym}. \quad (5.17) \end{cases}$$

The α parameter used to adjust the partial self-inductance approaches unity at DC and 0.94 at high frequencies with increasing aspect ratio $\frac{\mathfrak{L}}{D}$. β, used to adjust the partial mutual inductance, is unity at DC and ranges between 0.49 and 0.93 at high frequencies as the aspect ratio increases, respectively, from 0.5 to 9. The Rosa expressions are most inaccurate when calculating the mutual inductance of a small aspect ratio 3-D via operating at high frequencies. A comparison of both the partial self- and mutual inductances for the adjusted

Rosa expressions with Q3D simulations is provided, respectively, in Figures 5.8 and 5.9.

5.9.2 Inductance of non-equal length 3-D vias

An expression for the mutual inductance at DC between two TSVs with non-equal lengths is provided in (5.18). As for the previous topologies, the DC mutual inductance is the worst case inductance, as shown in Figure 5.2. Note also that the delay characteristics of a conductor are weakly correlated to the inductance [285]; therefore, the DC and f_{asym} inductance produces similar delay effects.

$$L_{21} = \alpha\beta\frac{\mu_o}{2\pi}[ln(\frac{\Omega_g+\sqrt{{\Omega_g}^2+P^2}}{P})\Omega + P - \sqrt{{\Omega_g}^2+P^2}], \tag{5.18}$$

$$\alpha = 0.8 + 0.1945\frac{S}{D}^{0.52}, \tag{5.19}$$

$$\beta = \beta_1 - 2e^{-0.3\frac{\Omega_y}{\Omega_g}+\beta_2}, \tag{5.20}$$

$$\beta_1 = 2.1 + 4e^{-0.375\frac{\Omega_g}{D}-0.1}, \tag{5.21}$$

$$\beta_2 = e^{-0.21\frac{\Omega_g}{D}+0.6} - 0.57. \tag{5.22}$$

The α parameter adjusts the partial mutual inductance between two non-equal length 3-D vias, accounting for the separation between vias. The β parameter adjusts the mutual

inductance based on the ratio of the larger length 3-D via $\mathfrak{Q}_y$ (which represents multiple vertically stacked 3-D vias) to the shorter length via $\mathfrak{Q}_g$ (a single 3-D via). $\mathfrak{Q}_y$ is therefore always an integer multiple of $\mathfrak{Q}_g$. The β term is dependent on the aspect ratio $\frac{\mathfrak{Q}_g}{D}$ of a single TSV (non-stacked), as shown by (5.21) and (5.22). A comparison between the mutual inductance produced by (5.18) with Q3D simulations is provided in Figure 5.10.

The high frequency inductance f_{asym} is calculated by multiplying the value produced by (5.18) with

$$\gamma = 0.955 - 1.1e^{-0.75-0.5\frac{\mathfrak{Q}_g}{D}}. \tag{5.23}$$

The γ term is only dependent on the aspect ratio of a single non-stacked TSV.

The self- and mutual inductance equations produce less than an 8% variation from electromagnetic simulations. The per cent variations between simulation and the model of the self- and mutual inductances are listed in Section 5.11.2, respectively, in Tables 5.9 and 5.10. The DC and f_{asym} mutual inductance of two non-equal length 3-D vias are listed, respectively, in Table 5.11 and Table 5.12. The error in all of these tables does not exceed 8% except in those cases where the pitch between vias is greater that 5·D and the aspect ratio is less than two, and even in these cases the error does not exceed 30%. The larger error in these cases is due to the small mutual inductance (less than 0.05 pH). The small mutual inductance exacerbates the per cent error between (5.11), (5.13), and (5.18) and the electromagnetic simulations.

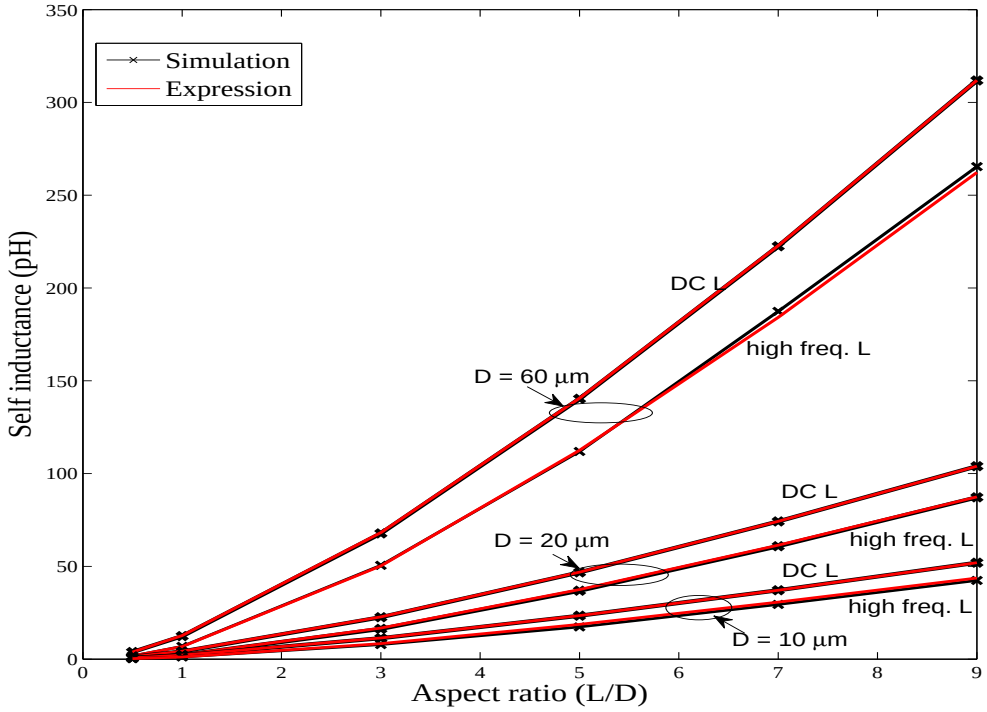


Figure 5.8: Self-inductance L_{11} of a cylindrical 3-D via.

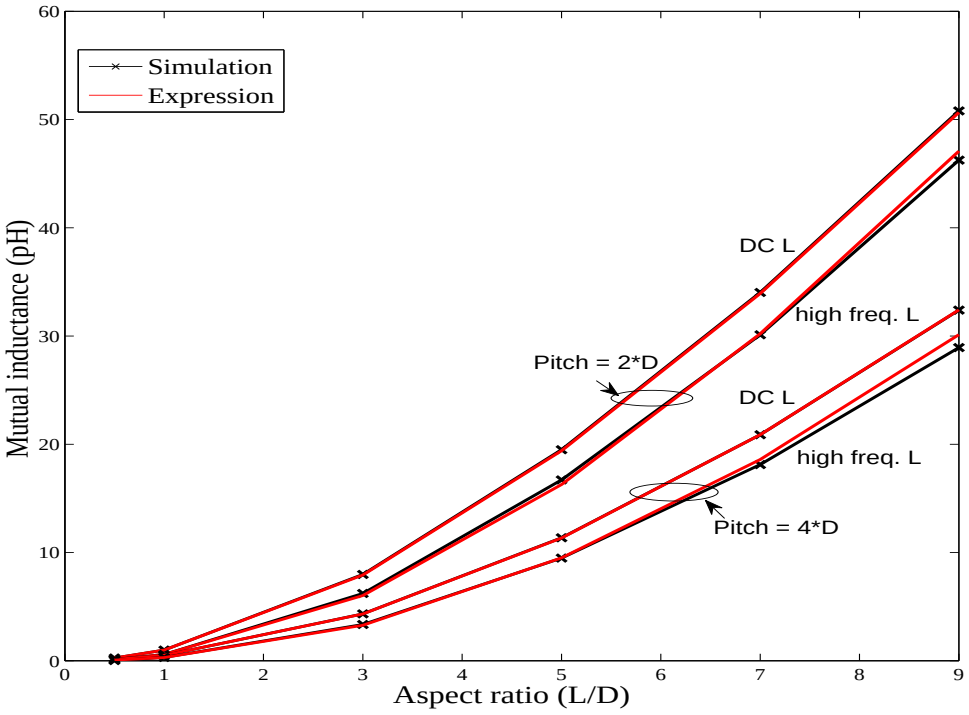


Figure 5.9: Mutual inductance L_{21} of a 20 μm diameter cylindrical 3-D via.

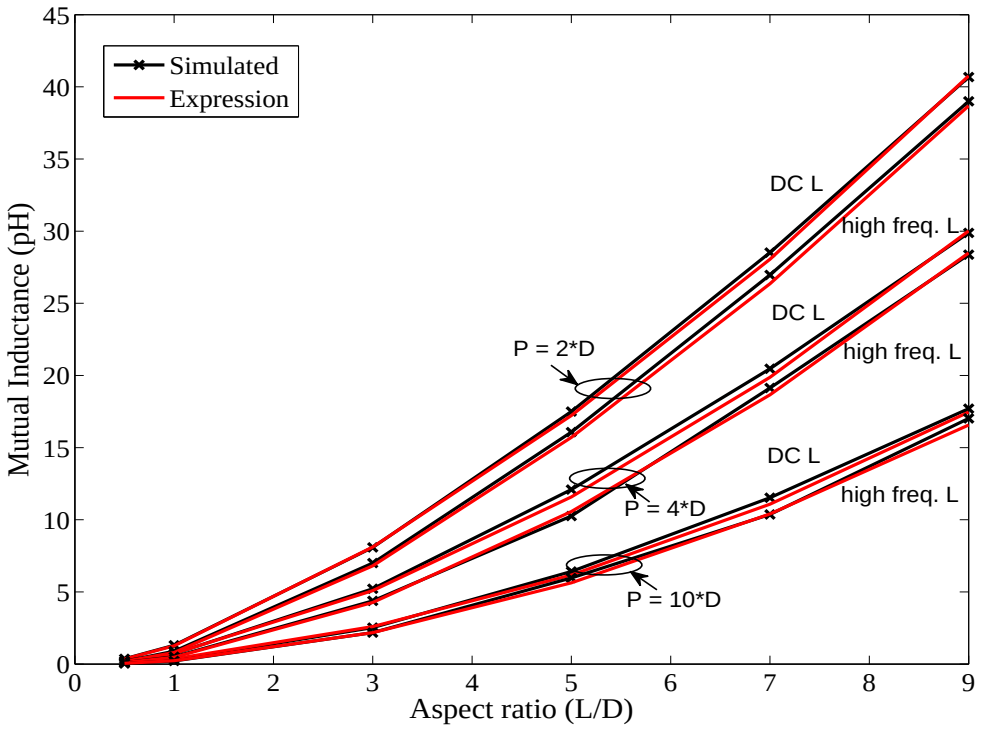


Figure 5.10: Mutual inductance L_{21} for two 3-D vias with different lengths (D = 10 μm, and $3\Omega_g = \Omega_y$).

5.10 Closed-form capacitance model of a 3-D via

Prior work, [126] and [150], examining the capacitance of bulk 3-D vias has neglected two important physical characteristics. The first issue is the formation of a depletion region in the bulk substrate surrounding the through silicon via, and the second issue is the assumption that the electrical field lines from the 3-D via terminate on a cylinder surrounding the via dielectric liner. Equations (5.24) and (5.25) from [126] and [150], respectively, overestimate the 3-D via capacitance. Equation (5.26) accounts for both the formation of a depletion region surrounding a p-type substrate in bulk processes, and the termination of the electrical field lines on a ground plane below the 3-D via. Termination of the field lines from the 3-D via to the ground plane forms a capacitance to the on-chip metal interconnect.

$$C = \frac{\epsilon_{Si}}{t_{diel}} 2\pi \Re H, \tag{5.24}$$

$$C = \frac{\epsilon_{Si}}{ln(\frac{\Re + t_{diel}}{\Re})} 2\pi H, \tag{5.25}$$

$$C = \alpha\beta \cdot \frac{\epsilon_{Si}}{t_{diel} + \frac{\epsilon_{SiO_2}}{\epsilon_{Si}} x_{dTp}} 2\pi \Re H. \tag{5.26}$$

Note that (5.26) is dependent on the depletion region depth x_{dT_p} in doped p-type silicon (the doped acceptor concentration N_A is, in this case, 10^{15} cm^{-3}). The depletion region is, in turn, dependent on the p-type silicon work function ϕ_{f_p}. The thermal voltage $\frac{kT}{q}$ at $T = 300$ K is 25.9 mV, where q is the electron charge (1.6×10^{-19} C) and k is the Boltzmann constant, 1.38×10^{-23} J/K.

$$x_{dTp} = \sqrt{\frac{4\epsilon_{Si}\phi_{fp}}{qN_A}}, \tag{5.27}$$

$$\phi_{fp} = V_{th} ln(\frac{N_A}{n_i}). \tag{5.28}$$

The fitting parameters, α and β, are used to adjust the capacitance for the two physical factors. The β parameter adjusts the capacitance of a 3-D via since a smaller component of the capacitance is contributed by the portion of the 3-D via farthest from the ground plane. A decrease in the growth of the capacitance therefore occurs as the aspect ratio increases. The α term is used to adjust the capacitance based on the distance to the ground plane S_{gnd}. As S_{gnd} increases, the capacitance of the 3-D via decreases. The α and β terms are

$$\alpha = (-0.0351\frac{\Omega}{D} + 1.5701)S_{gnd}^{0.0111\frac{\Omega}{D}-0.1997}, \tag{5.29}$$

$$\beta = 5.8934D^{-0.553}(\frac{\Omega}{D})^{-(0.0031D+0.43)}. \tag{5.30}$$

A plot of the 3-D via capacitance for diameters of 20, 40, and 60 μm are shown in Figure 5.11. The per cent variation between simulation and the model is included in Section 5.11.3 as Table 5.13. The error of the capacitance produced by (5.26) does not exceed 8%.

In addition to a closed-form expression for the capacitance of a single 3-D via over a ground plane, an expression for the coupling capacitance between two 3-D vias over a ground plane is presented. The expression for the coupling capacitance between two 3-D

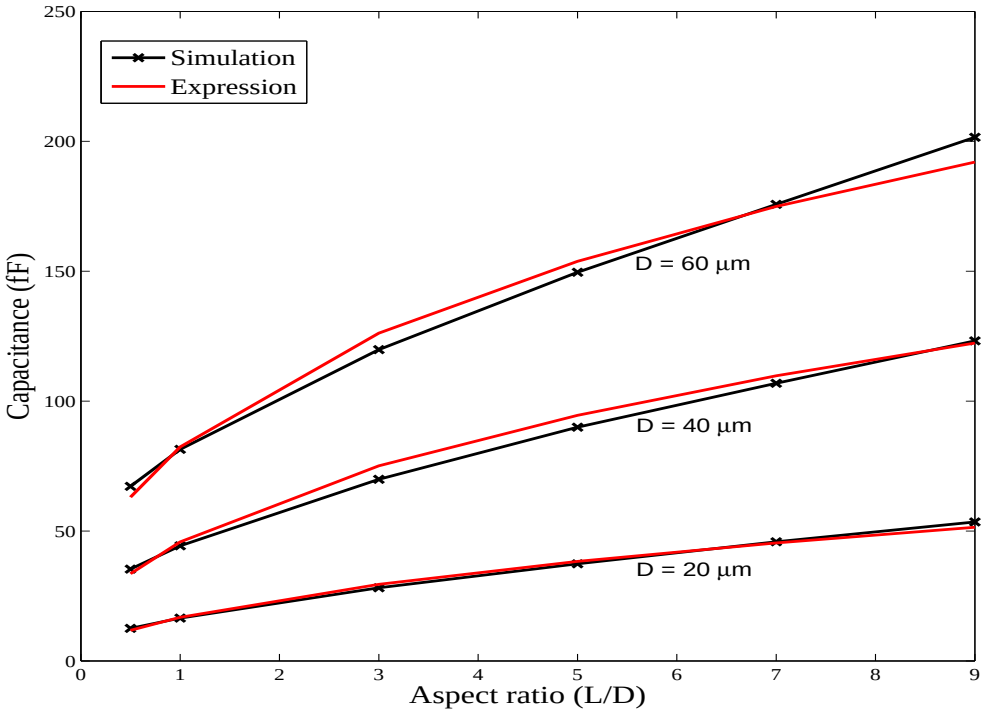


Figure 5.11: Capacitance of a cylindrical 3-D via over a ground plane.

vias is

$$C_c = 0.4\alpha\beta\gamma \cdot \frac{\epsilon_{Si}}{S}\pi DH. \tag{5.31}$$

The 0.4 multiplier in (5.31) adjusts the sheet capacitance between two TSVs when assuming that all electric field lines originating from half of the surface of one TSV terminate on the other TSV. Each fitting parameter (α, β, and γ) adjusts the coupling capacitance for a specific physical factor. The α term accounts for the nonlinearity of the coupling capacitance as a function of the aspect ratio $\frac{\Omega}{D}$. The effect of the separation between the TSVs to the ground plane S_{gnd} on the coupling capacitance is included in β. Note that the β term is dependent on the aspect ratio of the TSV. Finally, the γ parameter accounts for the nonlinearity of the coupling capacitance as a function of the distance S between the two TSVs. The γ parameter is also dependent on the TSV aspect ratio. The pitch P, which is the sum of the distance between the two vias and a single TSV diameter ($P = S + D$), is also included in γ. The three terms are

$$\alpha = 0.225ln(0.97\frac{\Omega}{D}) + 0.53, \tag{5.32}$$

$$\beta = 0.5711\frac{\Omega}{D}^{-0.988} ln(S_{gnd}) + (0.85 - e^{-\frac{\Omega}{D}+1.3}), \tag{5.33}$$

$$\gamma = \begin{cases} 1, & \text{if } \frac{S}{D} \leq 1 \quad (5.34) \\ \zeta[ln(\frac{\Omega}{D} + 4e^{-\frac{S}{9}} + 2.9) - 10.625S^{-0.51}], & \text{if } \frac{S}{D} > 1, \quad (5.35) \end{cases}$$

where ζ includes the dependence on the ratio of the pitch to the diameter and is

$$\zeta = (1 + e^{-(0.5+|\frac{P}{D}-4|)}). \tag{5.36}$$

A plot of the 3-D via coupling capacitance for diameters of 20, 40, and 60 μm are shown in Figure 5.12. The per cent variation between the simulations and the model is described in Section 5.11.4. The error of the capacitance produced by (5.31) does not exceed 15% for all aspect ratios greater than one. When the aspect ratio is between 0.5 and 1, the error between the simulations and the closed-form expression is much greater as the coupling capacitance between the 3-D vias drops below 1.5 fF in all cases.

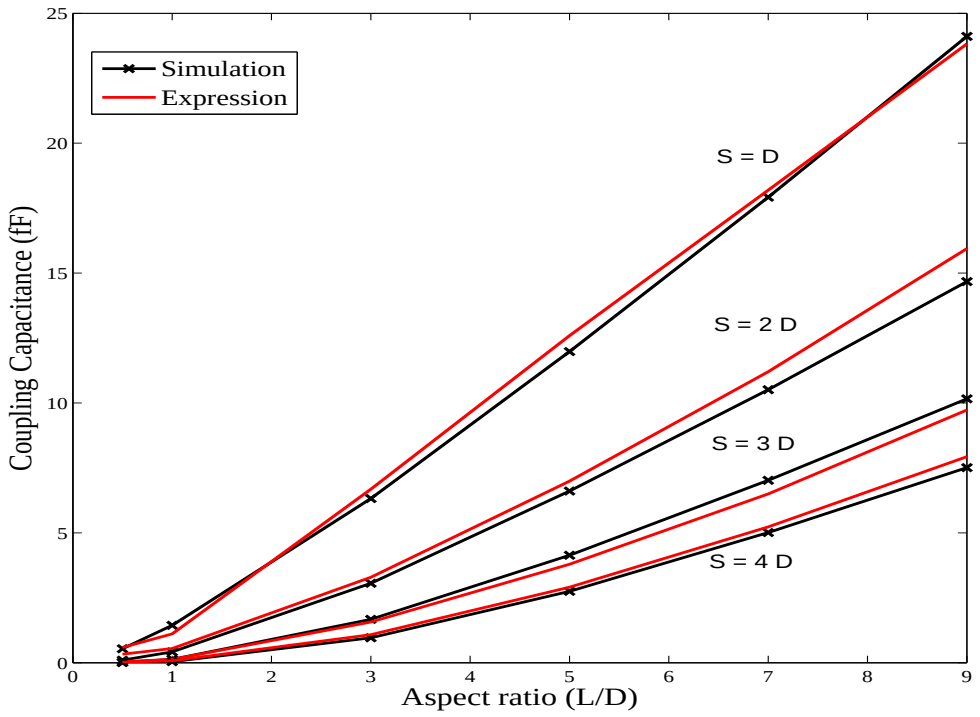


Figure 5.12: Coupling capacitance between two 3-D vias over a ground plane (D = 20 μm).

5.11 Comparison between closed-form expressions and electromagnetic simulations

A comparison between the closed-form expressions and the electromagnetic simulations of the resistance, inductance, and capacitance of a TSV, and the coupling capacitance between two TSVs is provided in this section. Examining the tabulated per cent difference for the electrical characteristics of the TSV reveals good agreement between the derived closed-form expressions and the electromagnetic simulations.

5.11.1 Per cent variation in resistance

The per cent error between the closed-form expressions and the electromagnetic simulation for the DC, 1 GHz, and 2 GHz resistances are listed, respectively, in Tables 5.6, 5.7, and 5.8. None of the errors listed in Tables 5.6, 5.7, and 5.8 exceeds 5.5% for all investigated diameters and aspect ratios.

Table 5.6: Per cent error between simulations and closed-form expressions of a 3-D via, DC resistance.

Diameter (μm)	Aspect ratio					
	0.5	1	3	5	7	9
1	-1.1	-1.1	-1.0	-1.0	-1.0	-1.0
5	-1.1	-1.1	-1.0	-1.0	-1.1	-0.9
10	-1.1	-1.0	-0.9	-0.9	-1.0	-0.9
20	-1.1	-0.9	-0.9	-0.7	-0.9	-0.9
40	-1.1	-1.1	-0.8	-0.9	-0.9	-0.9
60	-1.7	-0.9	-0.9	-0.9	-1.0	-0.9

Table 5.7: Per cent error between simulations and closed-form expressions of a 3-D via, 1 GHz resistance.

Diameter (μm)	Aspect ratio					
	0.5	1	3	5	7	9
1	0.1	-0.1	0.04	0.1	-0.1	0.1
5	0.5	-0.1	-0.9	-0.4	0.2	0.7
10	0.9	-0.1	-0.2	0.8	0.8	1.9
20	1.4	-0.5	-3.3	-2.7	-1.7	-0.7
40	4.4	0.9	-2.7	-1.5	-0.1	1.9
60	2.3	-0.2	-2.7	0.04	1.4	4.1

Table 5.8: Per cent error between simulations and closed-form expressions of a 3-D via, 2 GHz resistance.

Diameter (μm)	Aspect ratio					
	0.5	1	3	5	7	9
1	0.6	0.2	0.4	0.5	0.2	0.5
5	1.0	0.2	-0.8	-0.3	0.5	1.2
10	1.4	0.2	-0.1	1.2	1.3	2.6
20	1.5	-0.4	-3.7	-3.1	-1.8	-0.7
40	5.2	1.0	-2.9	-1.6	0	2.2
60	2.7	-0.4	-2.9	0.1	1.5	4.4

5.11.2 Per cent variation in inductance

The per cent error between the closed-form expressions and the electromagnetic simulation for the self- (L_{11}) and mutual (L_{21}) inductance of a 3-D via for several diameters and aspect ratios are listed, respectively, in Tables 5.9 and 5.10. The worst case DC and high frequency mutual inductance L_{21} between two non-equal length 3-D vias are listed, respectively, in Tables 5.11 and 5.12. The error does not exceed 8% in all cases except for those cases where the spacing between the 3-D vias is at least four times the diameter. In these cases, the mutual inductance is sufficiently small that minor deviations between the expression and the simulations produce a large per cent error.

Table 5.9: Per cent error between simulations and closed-form expressions of a 3-D via, self-inductance (L_{11}).

Diameter (μm)		Aspect ratio					
		0.5	1	3	5	7	9
1	DC	-1.6	0	0.9	0	0	0
	High f	3.6	-8.3	7.7	6.8	-1.6	0.5
5	DC	0	2.9	0.5	0.2	0.1	0
	High f	-6.7	-8.2	-0.2	-1.1	-1.9	-2.8
10	DC	-1.6	2.4	0.7	0.2	0.08	0.04
	High f	3.6	0.9	6.5	8.0	4.2	3.2
20	DC	0	2.7	0.7	0.2	0.1	0.04
	High f	3.6	1.8	2.8	1.8	0.9	0.3
40	DC	-0.4	2.8	0.6	0.2	0.1	0.04
	High f	2.7	-0.9	-0.6	-0.7	-2.1	-2.1
60	DC	-0.3	2.7	0.7	0.2	0.1	0.04
	High f	3.0	-1.2	-0.6	0.6	-1.7	-1.2

Table 5.10: Per cent error between simulations and closed-form expressions of a 3-D via, mutual inductance (L_{21}).

Diameter (μm)	Pitch		Aspect ratio					
			0.5	1	3	5	7	9
1	2·D	DC	-7.7	-2.0	0	0	0	-0.4
		High f	0	-6.5	0	1.3	4.1	0.9
5	2·D	DC	-1.6	0	-0.5	-0.4	-0.2	-0.3
		High f	0	0	-3.8	-3.1	-1.2	0.9
10	2·D	DC	-1.6	2.4	0.7	0.2	0.08	0.04
		High f	3.6	0.9	6.5	8.0	4.2	3.2
20	2·D	DC	0	-1.0	-0.6	-0.5	-0.3	-0.3
		High f	0	0	-3.1	-2.5	0.3	1.8
	3·D	DC	0	-1.5	-0.4	-0.2	-0.2	-0.2
		High f	0	-2.5	-4.1	-2.4	0.4	4.5
	4·D	DC	-7.7	0	-0.2	-0.2	0	0.1
		High f	1.7	-3.3	-2.1	0.3	2.6	4.1
	5·D	DC	0	0	-0.3	-0.1	-0.1	0.1
		High f	19.5	0	-4.3	-1.5	0.5	2.7
40	2·D	DC	-2.0	-1.5	-0.7	-0.4	-0.3	-0.3
		High f	0	-0.9	-3.8	-1.4	-0.1	2.5
60	2·D	DC	-1.3	-1.3	-0.7	-0.4	-0.3	-0.2
		High f	0	-1.7	-4.2	-2.6	-0.1	1.4

5.11.3 Per cent variation in capacitance

The per cent error between the closed-form expression and the electromagnetic simulations for the capacitance of a 3-D via over a ground plane is listed in Table 5.13. All errors listed in Table 5.13 exhibit a deviation of less than 8% between the expressions and simulations.

Table 5.11: Per cent error between simulations and closed-form expressions of two 3-D vias with non-equal length, DC mutual inductance (L_{21}).

Diameter (μm)	$\frac{\ell_{via2}}{\ell_{via1}}$	Pitch	Aspect ratio					
			0.5	1	3	5	7	9
1	1	2·D	-7.7	-2.0	0	0	0	-0.39
	2	2·D	-4.0	-3.2	-2.4	-4.7	-4.0	-1.7
	3	2·D	-5.6	-3.1	0	-1.7	-1.4	0.3
10	1	2·D	-4.8	-2.0	-0.8	-0.4	-0.3	-0.2
	2	2·D	-6.0	-3.2	-2.6	-4.5	-4.0	-1.6
	3	2·D	-5.8	-3.1	0.4	-1.6	-1.6	0.2
	3	4·D	5.4	1.4	-2.5	-4.2	-2.9	0.4
	3	6·D	16.7	8.2	-1.8	-5.6	-4.1	-0.6
	3	10·D	28.0	20	3.2	-3.7	-3.9	-1.4
40	1	2·D	-2.0	-1.5	-0.7	-0.4	-0.3	-0.3
	2	2·D	-5.1	-3.2	-2.5	-4.5	-4.0	-1.6
	3	2·D	-6.3	-3.3	0.4	-1.6	-1.6	0.2

Table 5.12: Per cent error between simulations and closed-form expressions of two 3-D vias with non-equal length, high frequency mutual inductance (L_{21}).

Diameter (μm)	$\frac{\ell_{via2}}{\ell_{via1}}$	Pitch	Aspect ratio					
			0.5	1	3	5	7	9
1	1	2·D	0	-6.5	0	1.3	4.1	0.9
	2	2·D	-7.1	-10.8	-0	-0.8	0.9	2.1
	3	2·D	-5.0	-15.8	0	1.3	3.1	4.6
10	1	2·D	0	0	0.3	3.0	4.1	5.7
	2	2·D	0	-1.7	-3.2	-3.9	1.3	-2.3
	3	2·D	-5.0	-10.1	-2.7	-2.2	-2.3	-0.8
	3	4·D	1.9	-7.8	-2.5	3.0	-2.5	0.4
	3	6·D	8.3	-2.9	-0.6	-6.7	-4.9	-2.1
	3	10·D	20.5	9.5	0	-5.7	0.3	-2.6
40	1	2·D	0	-0.9	-3.8	-1.4	-0.1	2.5
	2	2·D	0	-6.1	-2.9	-2.2	-2.1	1.0
	3	2·D	-9.8	-10.8	-1.4	-1.4	-0.2	1.3

Table 5.13: Per cent error between simulations and closed-form expressions of a 3-D via, capacitance.

Diameter (μm)	Space to Gnd (μm)	Aspect ratio					
		0.5	1	3	5	7	9
20	10	-5.8	2.0	4.7	2.4	-1.0	-3.9
	30	2.4	-8.0	-5.9	-2.1	0.7	3.1
	50	6.1	-5.6	-4.4	-0.9	2.3	3.3
40	10	-4.9	3.4	7.4	5.1	2.7	-0.7
60	10	-6.2	1.2	5.3	2.8	-0.5	-4.7

5.11.4 Per cent variation in coupling capacitance

The per cent error between the closed-form expression and the electromagnetic simulations for the coupling capacitance between two 3-D vias over a ground plane is listed in

Table 5.14. All errors listed in Table 5.14 exhibit a deviation of less than 15% between the expressions and simulation for all aspect ratios greater than one.

Table 5.14: Per cent error between simulations and closed-form expressions of a 3-D via, coupling capacitance.

Diameter (μm)	Space to Gnd (μm)	Space to Via (μm)	Aspect ratio					
			0.5	1	3	5	7	9
20	10	D	9.3	-22.9	5.7	5.1	1.5	-1.2
		$2{\cdot}D$	230	31.0	7.5	5.7	6.6	8.7
		$3{\cdot}D$	95	0	-6.0	-8.5	-7.4	-4.2
		$4{\cdot}D$	275	83.7	12.5	5.8	4.4	5.6
	30	D	-3.3	-16.2	4.3	4.6	1.4	-1.0
	50	D	-4.6	-14.5	3.4	3.7	0.9	-1.3
40	10	D	71	-10.1	2.3	-1.1	-3.4	-6.5
60	10	D	103	-5.9	1.4	-2.5	-5.8	-9.4

5.12 Summary

The electrical characterization of a single through silicon via for several 3-D via configurations is described. The *L/R* time constant is shown to dominate the *RC* time constant, implying that a 3-D via is inductively limited. Analysis of a two 3-D via system reveals that with increasing spacing between the two vias, the *RLC* impedances approach those of a single 3-D via. In addition, the analysis of a two 3-D via system verifies that inter-plane coupling must be considered when properly modeling 3-D integrated structures. The effects of placing a shield via between two signal vias on both the coupling capacitance and mutual inductance has also been investigated. A shield via decreases the capacitive coupling by as much as 77%, while increasing the inductive coupling by approximately 4.3%. Properly placing the return path reduces both the loop inductance and the mutual loop-to-loop inductance by as much as 4.5%.

Closed-formed expressions of the resistance, inductance, and capacitance of a 3-D via reveal good agreement with full-wave electromagnetic simulations. Errors of less than 2% between the closed-formed models and simulations are demonstrated for both the resistance and inductance of a 3-D via. Errors of less than 8% for the capacitance are also reported. The use of these closed-formed expressions rather than full-wave electromagnetic simulations to calculate the 3-D via impedances greatly enhances the system level design process by improving the computational efficiency, thereby permitting more performance efficient expressions for the TSV impedance to be applied during the design synthesis process. These models of the via impedance are accurate over a wide range of diameters, lengths, dielectric thicknesses, and spacings.

Chapter 6

Model of 3-D Clock Distribution Networks

An omnipresent and challenging issue for synchronous digital circuits is the reliable distribution of the clock signal to the many hundreds of thousands of sequential elements distributed throughout a synchronous circuit [286, 287]. The complexity of this task is further exacerbated in 3-D ICs as sequential elements belonging to the same clock domain (*i.e.*, synchronized by the same clock signal) can be located on multiple planes. Another fundamental issue in the design of clock distribution networks is low power consumption, since the clock network dissipates a significant portion of the total power consumed by a synchronous circuit [288, 289]. This constraint is stricter for 3-D ICs due to the higher power density and related thermal concerns.

In 2-D circuits, symmetric interconnect structures, such as H- and X-trees, are widely utilized to distribute the clock signal across a circuit [287]. The symmetry of these structures permits the clock signal to arrive at the leaves of the tree at approximately the same time, resulting in synchronous data processing. Maintaining this symmetry within a 3-D

circuit, however, is a difficult task.

An extension of an H-tree to three dimensions does not guarantee equidistant interconnect paths from the root to the leaves of the tree. The clock signal propagates through vertical interconnects, typically implemented by through silicon vias (TSVs) from the output of the clock driver to the center of the H-tree on the other planes. The impedance of the TSVs can increase the time for the clock signal to arrive at the leaves of the tree on these planes as compared to the time for the clock signal to arrive at the leaves of the tree located on the same plane as the clock driver. Furthermore, in a multi-plane 3-D circuit, three or four branches can emanate at each branch point, as depicted in Figure 6.1. The third and fourth branches propagate the clock signal to the other planes within the 3-D circuit. Similar to a design methodology for a 2-D H-tree topology, the width of each branch is reduced by a third (or more) of the segment width preceding the branch point to match the impedance at that branch point [287]. This requirement, however, is difficult to achieve as the third and fourth branches are connected through a TSV.

Global signaling issues in 3-D circuits, such as the clock signal distribution network, have only recently been explored [16, 291–293]. The latest studies consider thermal effects on buffered 3-D clock trees [294] and H-tree topologies [295, 296]. Experimental characterization of 3-D clock distribution networks has recently been presented [17, 290, 297]. Measurements from a 3-D test circuit employing several clock distribution architectures are presented in this chapter and are compared to electrical models described in Chapter 5. The test circuit was fabricated by the MIT Lincoln Laboratories (MITLL) [83, 89].

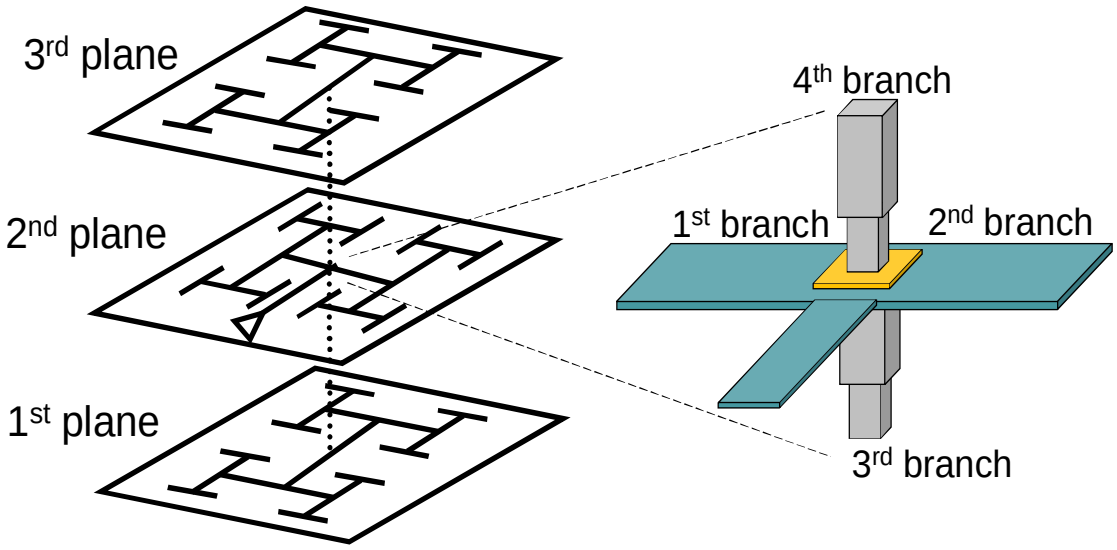


Figure 6.1: Schematic of a 3-D clock tree. The magnified detail illustrates the four branches that emanate from the output of the clock driver in the second plane. Two branches propagate the clock signal in the second plane, while the vertical TSVs transfer the clock signal to the first and third planes [290].

The objective of this chapter is to summarize the analysis of the temporal properties of different 3-D clock distribution topologies and compare these experimental results with circuit models. Clock distribution networks of increasing asymmetry in a 3-D stack are also investigated. Analysis of the delay provides enhanced understanding of the advantages and disadvantages of each topology, and aides in the design of synchronous circuitry in 3-D integrated systems. In addition, the effects of the TSVs on distributing the clock signals in a fabricated 3-D circuit are demonstrated and described.

In the following section, the design of the 3-D test circuit is reviewed. The experimental results and a discussion of the characteristics of the three clock distribution networks are presented in Section 6.2. Simulations of the clock distribution topologies including expressions modeling the 3-D via impedance are compared to the experimental results in Section 6.3. The circuit parameters used to model the clock delay within the 3-D clock topologies are also summarized in Section 6.3. Some conclusions are offered in Section 6.4. The closed-form expressions characterizing the impedance of the 3-D via are provided in Chapter 5, and the circuit parameters used to model the clock delay within the 3-D clock topologies are summarized in Section 6.3.2.

6.1 Design of the 3-D test circuit

The test circuit consists of three blocks. Each block includes the same logic circuit but implements a different clock distribution architecture. The total area of the test circuit is

3 mm $\times$ 3 mm, where each block occupies an approximate area of 1 mm^2. Each block contains about 30,000 transistors with a power supply voltage of 1.5 volts. The design kit used during the design process has been developed by North Carolina State University [298]. The different clock distribution architectures are reviewed in Section 6.1.1.

6.1.1 3-D clock topologies

Several clock network topologies for 3-D ICs are described in this section. These architectures combine different common 2-D topologies, such as H-trees, rings, and meshes [287]. Each of the three blocks includes a different clock distribution structure, which is schematically illustrated in Figure 6.2. The dashed lines depict vertical interconnects implemented by groups of through silicon vias. Multiple TSVs at the connection points between the clock networks lower the resistance of the vertical path while enhancing reliability.

As shown in Figure 6.2, these topologies range from purely symmetric to highly asymmetric networks to investigate the different features of these topologies. A primary objective is to determine the effect of the TSVs on the clock skew. The symmetry of the H-tree topology should be sufficient if the effect of the TSVs is small (for this specific technology). Alternatively, load balancing the global rings may reduce the delay of the clock signal caused by the TSVs. Local meshes may be preferable since the distribution of the clock signal to the sinks is primarily limited within a physical plane. Stacks of TSVs subsequently connect the sinks on other planes through local rings. This topology offers the

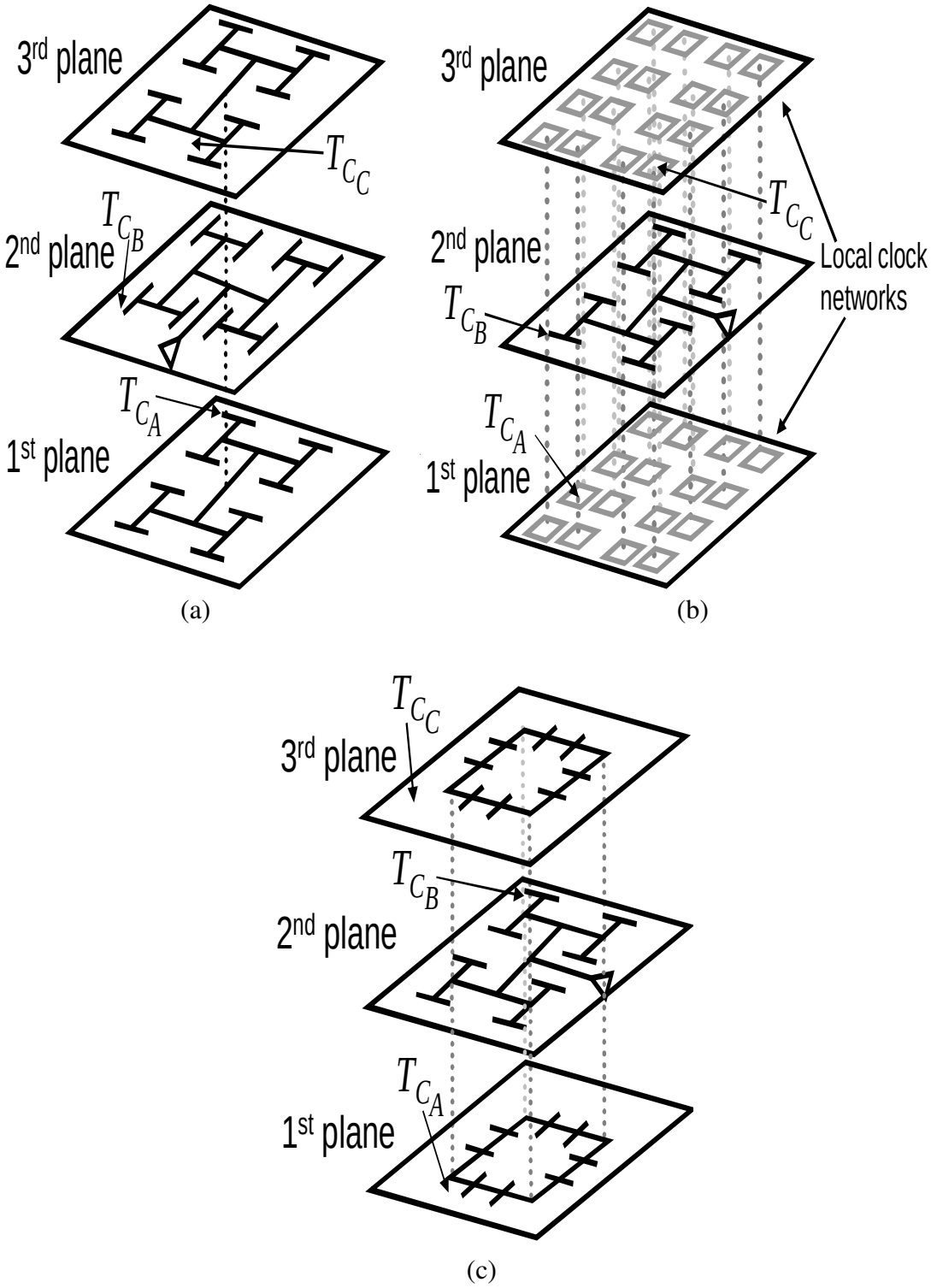


Figure 6.2: Three 3-D clock distribution networks within the test circuit, (a) H-trees, (b) H-tree and local rings, and (c) H-tree and global rings [290]

advantage of limiting most of the clock paths within one physical plane, while distributing the signal vertically to localized areas within neighboring planes.

The effect that these topological choices have on the clock skew, power dissipation, and signal slew are experimentally investigated. Since the clock signal is distributed in three dimensions, achieving equidistant signal propagation in a 3-D system is not straightforward. This task is further complicated by the different impedance characteristics of the vertical and horizontal interconnects. Consequently, the objective is to provide a global clock topology that produces sufficiently low skew (or predictable skew for delay compensation) within (intra-plane) and among (inter-plane) the planes of a 3-D circuit. The symmetry of an H-tree and the load balancing characteristics of rings and meshes are thereby exploited. The power consumed by each 3-D clock architecture is also considered due to the importance of thermal issues in 3-D circuits.

In each of the circuit blocks, the clock driver for the entire clock network is located on the second plane. The location of the clock driver is chosen to ensure that the clock signal propagates through identical vertical interconnect paths to the first and third planes, ideally resulting in the same delay. The clock driver is implemented with a traditional chain of tapered buffers [299, 300]. Additionally, buffers are inserted at the leaves of each H-tree in all three topologies. The width of the branches within the H-tree is halved at each branch point [301], with an initial width of 8 μm.

Note that in a 3-D circuit employing an even number of planes, the inherent symmetry of an H-tree topology in the vertical direction is not possible, increasing the inter-plane

clock skew between specific planes, as depicted in Figure 6.2. Therefore, for a 3-D technology supporting n physical planes, where n is even, symmetry along the vertical direction is not feasible. Placing the clock driver in plane $n/2 + 1$ or $n/2 - 1$ results in an increase in skew between the first and n^{th} plane equal to the effective delay of the group of TSVs connecting the two successive planes. This increase in delay can be compensated by placing fewer TSVs (with a higher effective impedance) in the vertical direction.

The architectures employed in the blocks are [290, 297, 302]:

Block A: All of the planes contain a four level H-tree (*i.e.*, equivalent to 16 leaves) with identical interconnect characteristics. All of the H-trees are connected through a group of TSVs at the output of the clock driver. Note that in Figure 6.2a the H-tree on the second plane is rotated by 90° with respect to the H-trees on the other two planes. This rotation eliminates inductive coupling between the H-trees. All of the H-trees are shielded with two parallel lines connected to ground.

Block B: A four level H-tree is included in the second plane. All of the leaves of this H-tree are connected by four TSVs to small local rings on the first and third planes, as illustrated in Figure 6.2b. As in Block A, the H-tree is shielded with two parallel lines connected to ground. Additional interconnect resources form local rings. Due to the limited interconnect resources, however, achieving a uniform mesh in each ring is difficult. Clock routing is constrained by the power and ground lines as only three metal layers are available on each plane [83, 89].

Block C: The clock distribution network for the second plane is a shielded four level H-tree.

Two global rings are utilized for the other two planes, as shown in Figure 6.2c. Buffers are inserted to drive each ring, which are connected by TSVs to the four branch points on the second level of the H-tree. The rings on planes A and C are connected to the second level of the H-tree for two reasons; first, to avoid an unnecessarily long ring that would result in a significant capacitive load, and second, to maintain a ring with sides of equal length. Additionally, connecting the ring to the leaves at the perimeter of the H-tree results in a considerable difference in the load among the sinks of the tree, since only the outer leaves are connected to the ring. The registers in each plane are connected either directly to the rings on the first and third planes or are driven by buffers at the leaves of the H-tree on the second plane. With this arrangement, the balancing properties of the rings results in low interplane skew for the first and third planes. In addition, since the interplane path to these planes is the same for both planes, the skew between these two planes is low as compared to the H-tree topology in the second plane.

A primary objective of this chapter is to evaluate the delay of different clock distribution architectures. A secondary and related objective is to analyze the characteristics of asymmetric topologies in 3-D systems. This objective poses several limitations on the power distribution network within each block. Power and ground rings at the periphery of each block are utilized. Although this architecture is not optimal, the structure is sufficiently small ($\sim$ 1 mm^2). The small size of the blocks does not cause a significant voltage drop across and among the planes. In addition, minimal *L di/dt* noise is observed during circuit operation. The power and ground rings on each plane are connected by a large number of

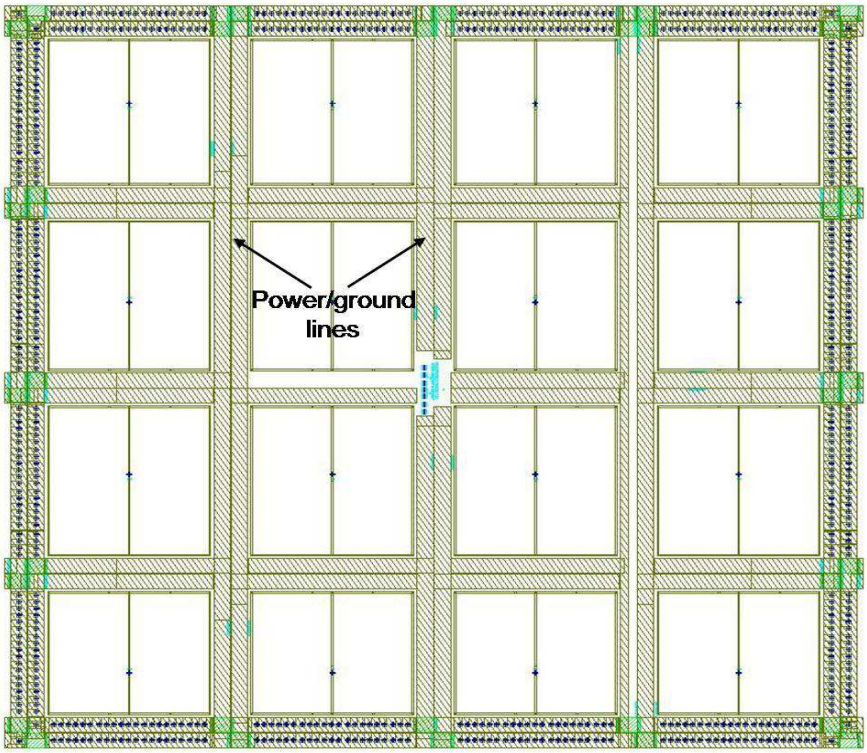
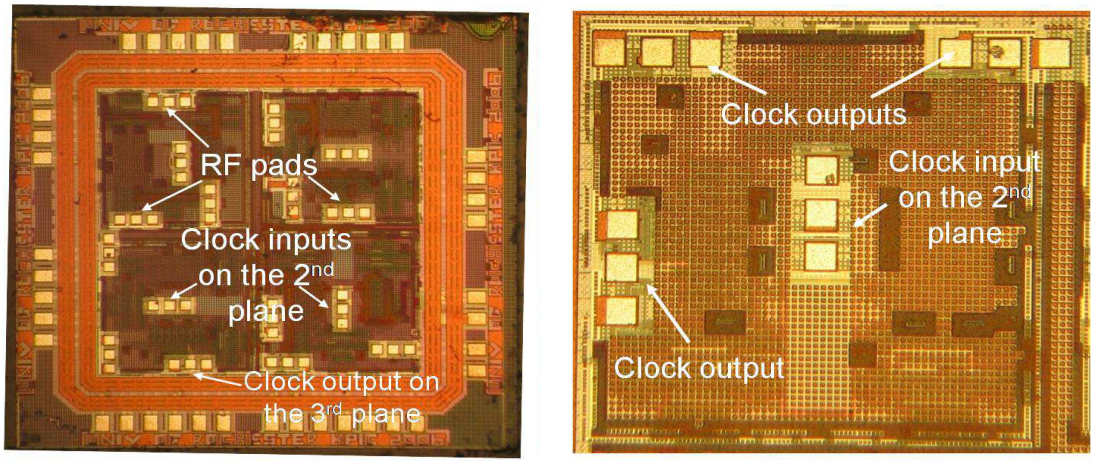


Figure 6.3: Power distribution network for the local ring topology.

TSVs to lower the impedance of the vertical interconnections. The local rings topology requires greater area for distributing power and ground in the first and third planes, where the local rings distribute the clock signal. The resulting clock and power networks for planes A and C are illustrated in Figure 6.3 for this specific block [17, 290]. In this topology, the power distribution network consists of a coarse mesh of power and ground lines [161].



(a) (b)

Figure 6.4: Fabricated 3-D circuit, (a) all of the blocks and DC pads, and (b) a magnified view of one block.

6.2 Experimental results

The clock distribution network topologies of the 3-D test circuit are evaluated in this section. The fabricated circuit is depicted in Figure 6.4a [17, 290], where the individual blocks can be distinguished. A magnified view of one block is shown in Figure 6.4b [17, 290, 297]. Each block includes four RF pads for measuring the delay of the clock signal. The pad located at the center of each block provides the input clock signal. The clock input waveform is a sinusoidal signal with a DC offset, which is converted to a square waveform at the output of the clock driver. The remaining three RF pads are used to measure the delay of the clock signal at specific points on the clock distribution network within each plane. A buffer is connected at each of these measurement points. The output of this buffer drives the gate of an open drain transistor connected to the RF pad.

The topology combining an H-tree and global rings demonstrated circuit operation at

1.4 GHz [17,290,297]. The delay of the clock signal from the RF input pad at the center of each block to the measurement point on plane i is denoted as t_i in Table 6.1. For example, t_A denotes the delay of the clock signal to the measurement point on plane A. The clock delay from the source node on the second plane to each leaf on the three separate planes is listed in Table 6.1. The clock delay is the average root to leaf delay from data collected at frequencies of 500 and 1000 MHz.

Table 6.1: Measured clock delay from the root to the leaves of each plane for each block.

Clock distribution network	Clock delay (ns)		
	t_A	t_B	t_C
H-trees (Figure 6.2a)	0.359	0.338	0.325
Local rings (Figure 6.2b)	0.445	0.350	0.300
Global rings (Figure 6.2c)	0.530	0.480	0.445

For the H-tree topology, the clock signal delay is measured from the root to a leaf of the tree on each plane, with no additional load connected to these leaves. A schematic of this topology including a path of the clock signal is shown in Figure 6.5. The delay of the clock signal to the sink of the H-tree on the second plane t_B is larger due to the additional capacitance coupled to that quadrant of the H-tree. This capacitance is intentional on-chip decoupling capacitance placed under the quadrant, increasing the measured delay. This topology produces, on average, comparable delay to the local rings topology, and less delay than the global ring clock structure.

Since three-dimensional integration greatly increases the complexity of designing an integrated system, a topology that offers low overhead during the design process of a 3-D clock distribution network is preferable. From this perspective, a potential advantage

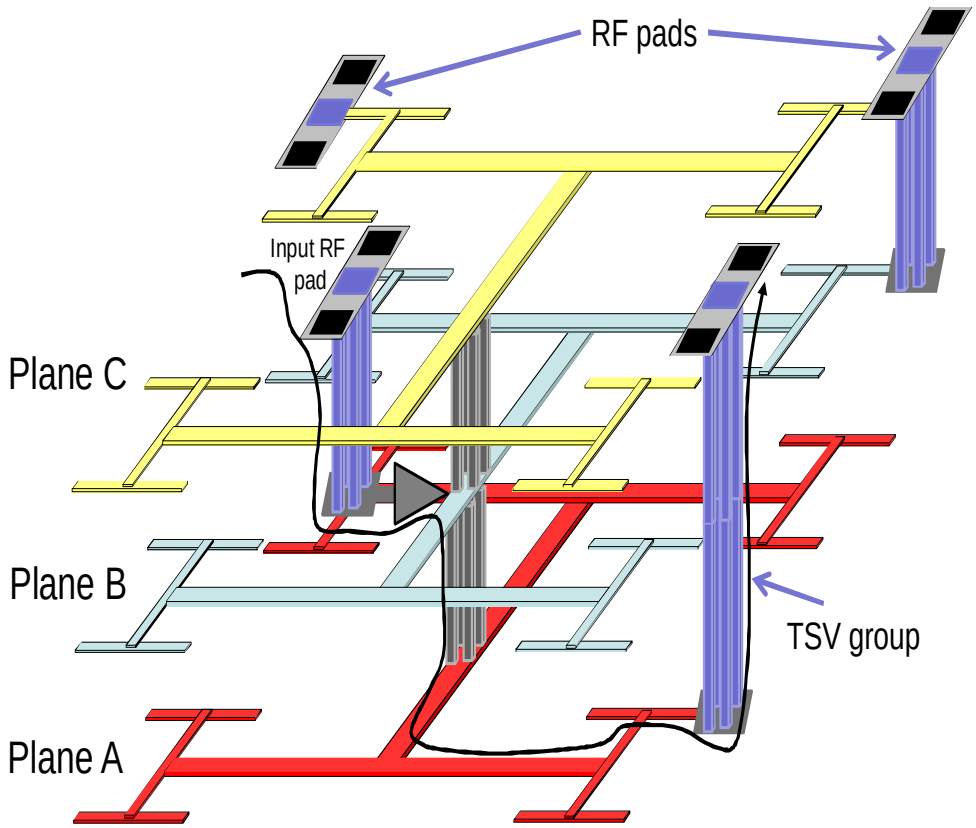


Figure 6.5: 3-D H-tree topology, where the path of the clock signal from the input to plane A and through a group of TSVs to an output RF pad is shown (not to scale) [290].

of the H-tree topology is that each plane can be individually analyzed. This capability is supported by the H-tree topology since the clock distribution network in each plane is exclusively connected to registers within the same plane. Alternatively, in the local rings topology, registers from all of the planes, which are connected to each sink of the tree on the second plane, all need to be simultaneously considered.

6.3 Models of the clock distribution network topologies incorporating the 3-D via impedance

Simulation of the fabricated clock distribution topologies incorporating the modeled electrical impedance of the interplane 3-D vias is described in this section. A comparison between the simulated and experimental results is also presented here. The electrical impedance of the 3-D vias is described for several diameters, lengths, dielectric thicknesses (bulk), and via-to-via spacings [134, 207]. The extracted parameters are used in the closed-form expressions characterizing the 3-D via impedance [134]. These equations are used here to model the contribution of the 3-D vias to the delay and skew characteristics of the clock distribution topologies and are described in detail in Chapter 5.

In addition to characterizing the electrical parameters of the TSVs, the electrical characteristics of the clock distribution network on each plane are determined through numerical simulation. This set of simulations has been performed for the three widths used in the fabricated test circuit, and for five different lengths. Trend lines for the capacitance, DC

resistance, 1 GHz resistance, DC self- and mutual inductance, and the asymptotic self- and mutual inductance f_{asym} approximate the electrical parameters of different length interconnect segments within the clock network. These simulations include two ground return paths spaced 2 μm from either side of the clock line. These return paths behave as ground for the electrical field lines emanating from the clock line, resulting in a more accurate estimate of the capacitance.

The electrical paths of the clock signal propagating from the root to the leaves of each plane for the H-tree clock topology (see Figure 6.2a) is depicted in Figure 6.6. The size of the source follower NMOS transistor and the dimensions of the clock buffers at the root, leaves, and output circuitry are provided in Section 6.3.2 B. The clock network on each plane is composed of 50 μm segments, where a π-network represents the electrical properties of each segment. These 50 μm segments model the distributive electrical properties of the interconnect. Similarly, when either meshes (Figure 6.2b) or rings (Figure 6.2c) are used on planes A and C (see Figure 6.6), each 50 μm segment is replaced with an equivalent π-network to more accurately represent the single mesh and ring structure within the test circuit. Note that for the mesh structures, the clock signal is distributed to planes A and C from the leaves of the H-tree in plane B while for the rings topology, the clock signal distributed to planes A and C is driven by buffers at the second level of the H-tree. The delay from the root to the leaves of each plane is included in Table V.

The clock delays listed in Table 6.2 are compared with the measured values listed in Table 6.1. Good agreement between the model and experimental data is shown. The per

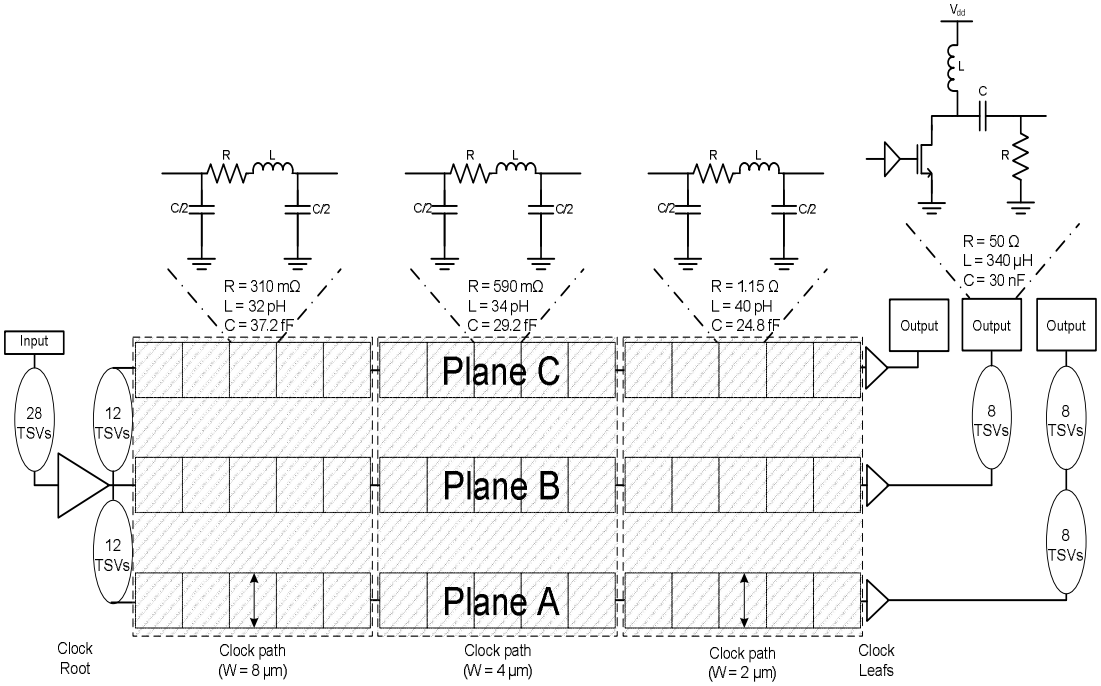


Figure 6.6: Structure of clock distribution network path from Figure 6.2a to model the clock skew. The number within each oval represents the number of parallel TSVs between device planes.

Table 6.2: Modeled clock delay from the root to the leaves of each plane for each block.

Clock distribution network	Clock delay (ns)		
	t_A	t_B	t_C
H-trees (Figure 6.2a)	0.359	0.355	0.351
Local rings (Figure 6.2b)	0.325	0.323	0.321
Global rings (Figure 6.2c)	0.510	0.465	0.442

cent error between the model and experimental clock delays is listed in Table 6.3. A maximum error of less than 10% is achieved for the clock paths within the H-tree topology. The larger errors listed in Table 6.3 are due to the small time scale being examined. All of the values listed in Table 6.1 and 6.2 are less than 550 picoseconds; therefore, any small deviation in delay produces a large error.

Table 6.3: Per cent error between modeled and experimental clock delay.

Clock distribution network	Clock delay % error		
	t_A	t_B	t_C
H-trees (Figure 6.2a)	0	-4.8	-7.4
Local rings (Figure 6.2b)	36.9	8.4	-6.5
Global rings (Figure 6.2c)	3.9	3.2	0.7

6.3.1 Equivalent electrical model of TSV for clock delay simulations

The resistance, inductance, and capacitance expressions are compared to numerical simulations for the TSV structures in the MITLL multi-project wafer (for 3-D via parameters, $\Re$ = 1 μm, $\wp$ = 8.5 μm, and P = 5 μm) in Table 6.4. The equivalent electrical model of a TSV is shown in Figure 6.7.

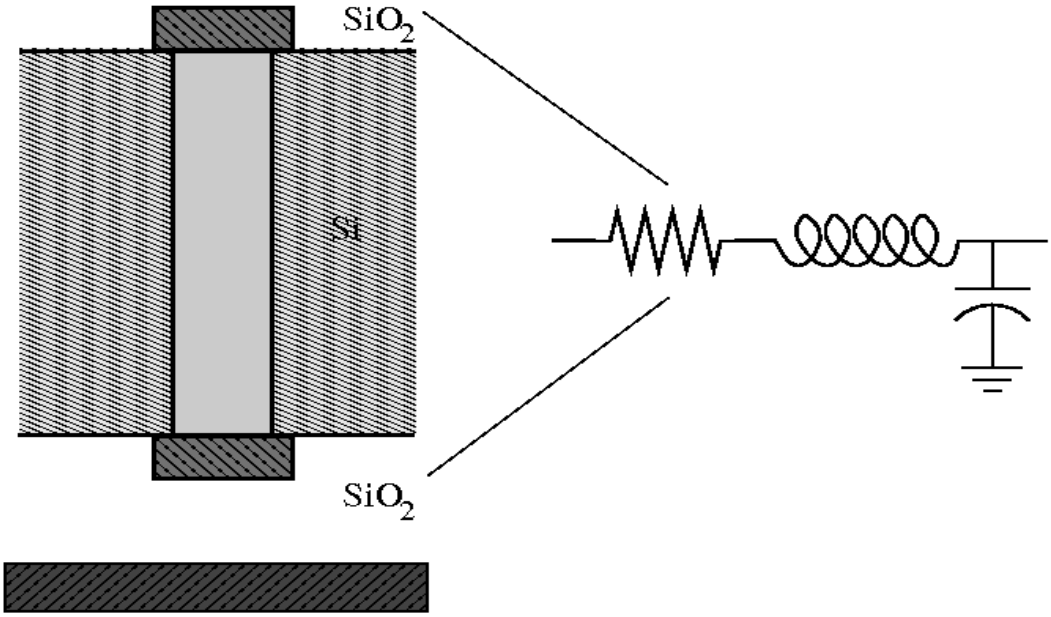


Figure 6.7: Equivalent electrical model of a TSV.

Table 6.4: Comparison of numerical simulations and analytic expressions of the TSV electrical parameters.

Electrical Parameters	Numerical Simulation	Analytic Expressions	% Error
DC resistance (mΩ)	148	154	4.1
1 GHz resistance (mΩ)	166	177	6.6
DC self inductance (pH)	3.9	3.9	0
f_{asym} self inductance (pH)	2.9	3.1	6.9
DC mutual inductance (pH)	1.40	1.32	-5.7
f_{asym} mutual inductance (pH)	1.10	1.08	-1.8
Capacitance (fF)	1.43	—	—

6.3.2 Circuit parameters to model the clock skew of the 3-D clock topologies

The circuit parameters used to model the skew within the clock network are provided below. The dimensions of the buffer circuits at the root, leaves, and output circuitry are listed in Table 6.5. Two sets of transistor widths are provided as each location is double buffered to maintain the same signal logic level. The dimensions of the ring and a single mesh are listed in Table 6.6. These segments are 50 μm long, and each segment is modeled by an equivalent π network for a line width of 4 μm. The source follower NMOS transistor located in the output circuitry has a length of 180 nm, and a width of 12 μm. The interconnect length connecting the output circuitry and pads to the leaves on each of the three device planes varies from 0 to 150 μm depending upon the clock topology (line width of 2 μm), and is also represented by an equivalent π network.

Table 6.5: Transistor width of the clock buffers at the root, leaves, and output circuitry (All lengths are 180 nm).

Buffer location		W_N (μm)	W_P (μm)
Root	Buffer 1	20	50
	Buffer 2	54	136
Leaf	Buffer 1	15	38
	Buffer 2	15	38
Output circuitry	Buffer 1	2.5	7
	Buffer 2	2.5	7

Table 6.6: Dimensions of local and global clock rings.

Clock topology	Length (μm)	Width (μm)
Global rings	500	500
Local rings	200	200

6.4 Summary

The design of a clock distribution network for application to 3-D circuits is considerably more complex than the design of a 2-D clock distribution network. Three topologies to globally distribute a clock signal within a 3-D circuit have been evaluated. A 3-D test circuit, based on the MITLL 3-D IC manufacturing process, has been designed, fabricated, and measured and is shown to operate at 1.4 GHz. Clock delay simulations incorporating both numerical simulations and analytic expressions produce comparable results to the experimentally extracted clock delay measurements. The clock delay measurements indicate that a topology combining the symmetry of an H-tree on the second plane and local rings on the remaining two planes results in a low clock delay in 3-D circuits. The topology with the H-tree on the second plane and global rings on the remaining two planes produced the largest root to leaf clock delay as compared to the other investigated topologies. The different performance characteristics of these topologies suggest that the target requirements should be considered when designing a 3-D clock distribution network.

Chapter 7

3-D Power Distribution Topologies and Models

A technological push, due to the proliferation of wireless devices, has accelerated the need to find alternative technologies that increase device density, merge disparate technologies, while reducing the power budget. 3-D integration has emerged as a potential solution that supports mixed-signal systems without the need for greater process complexity. A critical component of 3-D systems, which is experimentally explored in this paper, is delivering power to multiple device planes [177].

An important issue for 3-D integrated circuits (ICs) is the design of a robust power distribution network that can provide sufficient current to every load within a system. In planar ICs where flip-chip packaging is adopted as the packaging technique, an array of power and ground pads is allocated throughout the integrated circuit. Increasing current densities and faster current transients, however, complicate the power distribution design process. Three-dimensional integration provides additional metal layers for the power distribution networks through topologies unavailable in two-dimensional circuits. With 3-D technologies, individual planes can potentially be dedicated to delivering power.

The challenges of efficiently delivering power across a 2-D circuit while satisfying local current requirements have been explored for decades [141, 162, 303]. Two-dimensional power distribution networks are designed to achieve specific noise requirements. A variety of techniques have been developed to minimize both *IR* drops and *L·di/dt* noise [304, 305], such as multi-tiered decoupling placement schemes [306–308] and power gating [309]. These techniques have been effective with increasing current demands of each progressive technology node. 3-D integrated systems however are in its infancy, and much work is required to design efficient power distribution topologies for these vertically integrated systems.

Power delivery in 3-D integrated systems presents difficult new challenges for delivering sufficient current to each device plane. Stacking device planes in the vertical direction leads to higher power densities [17]. Specialized techniques are required to ensure that each device plane is operational, while not exceeding the target output impedance [141, 162]. The focus of this chapter is on a primary issue in 3-D power delivery, the power distribution network, and provides a quantitative experimental analysis of the noise measured on each plane within a three plane 3-D integrated stack.

The effects of the through silicon vias (TSVs) on the *IR* voltage drops and *L·di/dt* noise are significant, as the impedance of a 3-D power distribution network is greatly affected by the TSV density. In addition, the electrical characteristics of a TSV vary based on the 3-D via diameter, length, and dielectric thickness [129,134,207]. A comparison of two different via densities for identical power distribution networks is also provided in this chapter, and implications of the 3-D via density on the power network design process is discussed.

The proper placement of decoupling capacitors can potentially reduce noise within the power network, while enhancing performance. The effect of board level decoupling capacitors on *IR* and *L·di/dt* noise in 3-D circuits is also described here. Methods for placing decoupling capacitors at the interface between planes to minimize the effects of inter-plane noise coupling are also suggested.

The 3-D test circuit is described in the following section. A brief summary of the MITLL 3-D process is provided in Section 7.2. Experimental results of the noise characteristics of the power distribution networks are presented in Section 7.3. A discussion of the experimental results and the effect of the choice of power distribution topology on the noise characteristics of the power network is provided in Section 7.4. Some conclusions are offered in Section 7.5.

7.1 Design of 3-D power distribution network test circuit

The fabricated test circuit is 2 mm × 2 mm, and composed of four equal area quadrants. Three quadrants are used to evaluate the effects of the topology of the power distribution network on the noise propagation characteristics, and one quadrant is dedicated to DC-to-DC conversion. Each stacked power network is 530 μm x 500 μm, and includes three discrete two-dimensional power networks, one network within each of the three device planes. The total area occupied by each block is less than 0.3 mm^2, representing a portion of a power delivery network. Each block includes the same logic circuit but utilizes a different power distribution architecture. The power supply voltage is 1.5 volts for all of the blocks. The different power distribution architectures are reviewed in Section 7.1.1, an

overview of the various circuit layouts and schematics is provided in Section 7.1.2, and the logic circuitry common to each power module is described in Section 7.1.3.

7.1.1 3-D power topologies

Interdigitated power/ground lines are used in each of the power network topologies. There are five main objectives for the test circuit: i) determine the peak and average noise within the power and ground distribution networks, ii) determine the effect of the board level decoupling capacitors on reducing power noise, iii) explore the effects of a dedicated power/ground plane on the power noise, iv) investigate the effects of the TSV density on the noise characteristics of the power network, and v) evaluate the resonant characteristic frequency of each power network topology.

The three topologies are illustrated in Figure 7.1. The difference between the left (Block 1) and central (Block 2) topologies is the number of TSVs, where the latter topology contains two-thirds the number of TSVs. The third topology (Block 3) replaces the interdigitated power and ground lines on the second device plane with two metal planes to assess the benefit of allocating dedicated power and ground planes to deliver current to the loads within a 3-D system. The interdigitated power and ground lines are both 15 μm wide and separated by a 1 μm space for all three power distribution networks. The power and ground planes of Block 3 are separated by a 1 μm space of inter-layer dielectric composed of plasma etched, chemical vapor deposited tetra-ethyl-ortho-silicate (TEOS). Based on the experimental results, enhanced understanding of the noise propagation properties of these 3-D power network topologies has been achieved.

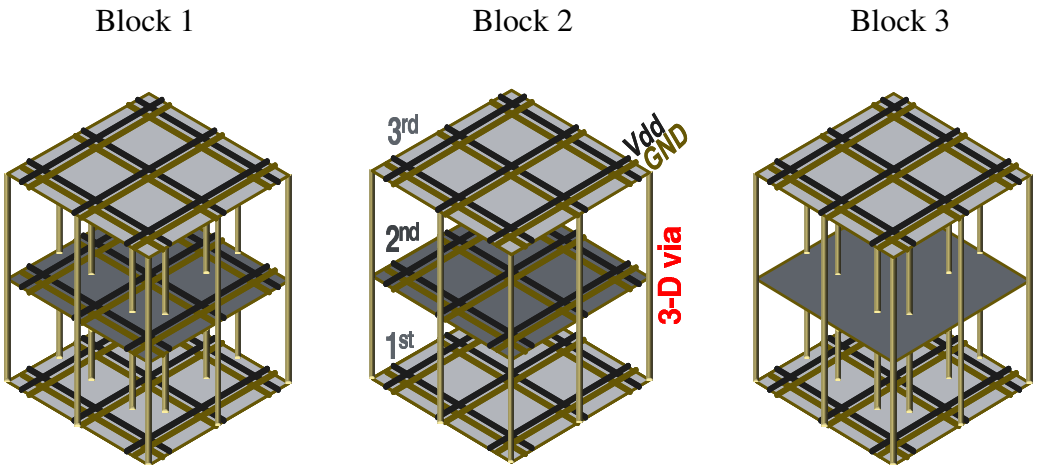


Figure 7.1: Power distribution network topologies. (a) interdigitated power network on all planes with the 3-D vias distributing current on the periphery and through the middle of the circuit, (b) interdigitated power network on all planes with the 3-D vias distributing current on the periphery, and (c) interdigitated power network on planes 1 and 3 and power/ground planes on plane 2 with the 3-D vias distributing current on the periphery and through the middle of the circuit.

7.1.2 Layouts and schematics of the 3-D test circuit

The components of both the noise generation and noise detection circuits are described in this section. Layout and schematic views of the various sections of the test circuit are provided. The transistor dimensions of the various circuits are also included in the circuit schematics. The minimum sized transistor length is 150 nm. Most transistors are designed, however, with a 200 nm channel length (matching the dimensions of the standard library). A select few circuits, such as the ring oscillators (RO), are designed with larger channel lengths (see Figure 7.5 and Table 7.1). A layout view of the overall test circuit is shown in Figure 7.2. Also shown in this figure are the three blocks discussed in this chapter, and a fourth block for evaluating a 3-D dimensional DC-to-DC buck converter [310].

As previously mentioned, each of the three blocks include two sets of circuits, noise

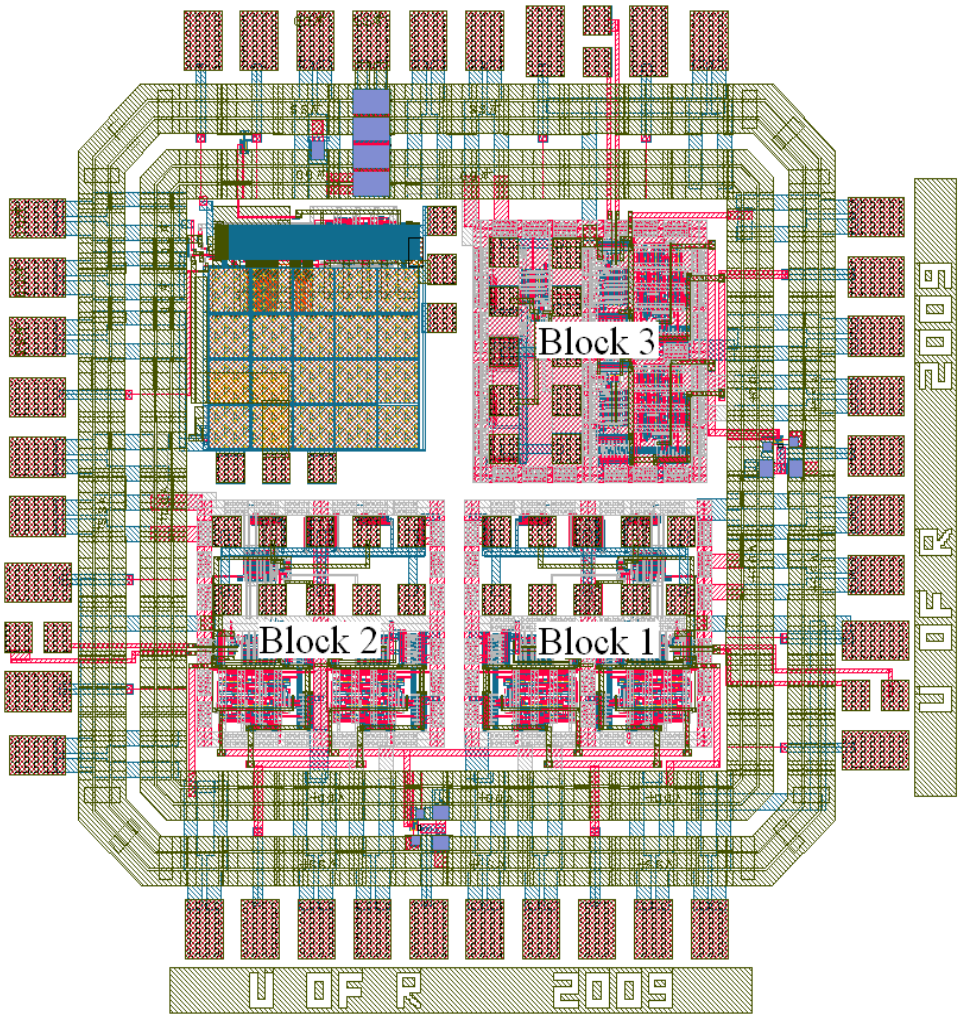
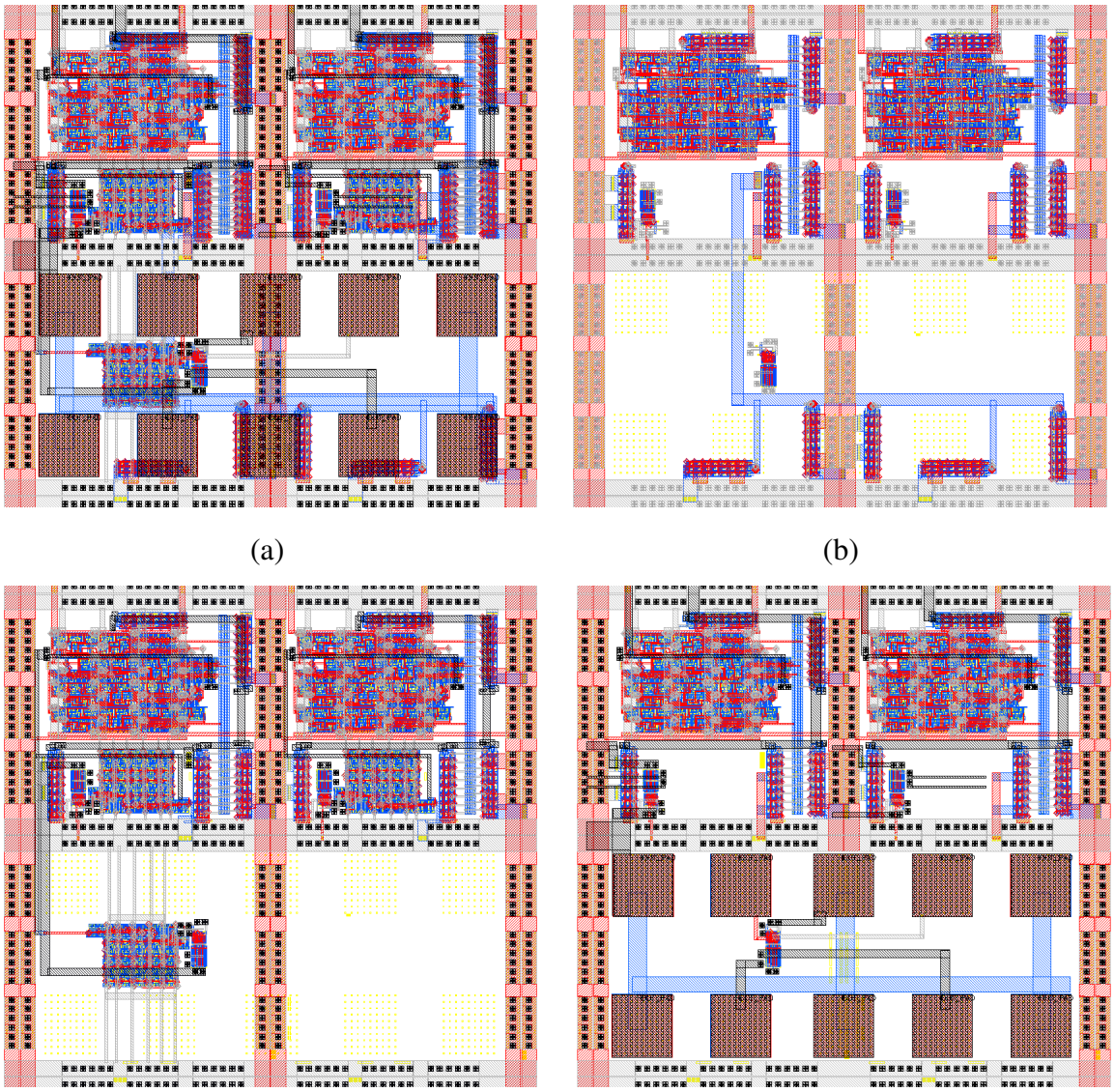


Figure 7.2: Layout of the power distribution network test circuit.

generation and noise detection circuits. In addition to these circuits, calibration circuits are also included to analyze the gain and bandwidth of the source follower noise sensing devices. Each block contains the calibration circuits, but one set is sufficient to adequately characterize the gain and bandwidth of the sense circuits for a given IC. The calibration circuits are identical to the detection circuits, other than the input to these circuits being directly supplied from an external source. The Block 1 layout is depicted in Figure 7.3. As previously noted, Block 1 includes interdigitated power and ground lines on all three device planes. The four layouts depict the three stacked device planes (Figure 7.3a), the bottom device plane (Figure 7.3b), the middle device plane (Figure 7.3c), and the top device plane (Figure 7.3d). The three different subcircuits are labeled in Figure 7.3a. The RF pads shown in the figure are associated with the calibration circuits.

The layout of the logic block used to generate sequence patterns for the current mirrors is shown in Figure 7.4. The components of the sequence generator are enclosed in boxes and labeled in Figure 7.4a. The components included in the test circuit are the three ring oscillators (RO), buffers for the ring oscillators, the four pseudorandom number generators (PRNG), and buffers for the number generators. The RO and PRNG buffers are identically sized except that the enable signal for the RO is connected to the external reset signal pin, whereas the PRNG enable signal is tied to the 1.5 volt supply. A schematic of each component and the corresponding transistor sizes are included in Figures 7.5 and 7.6.

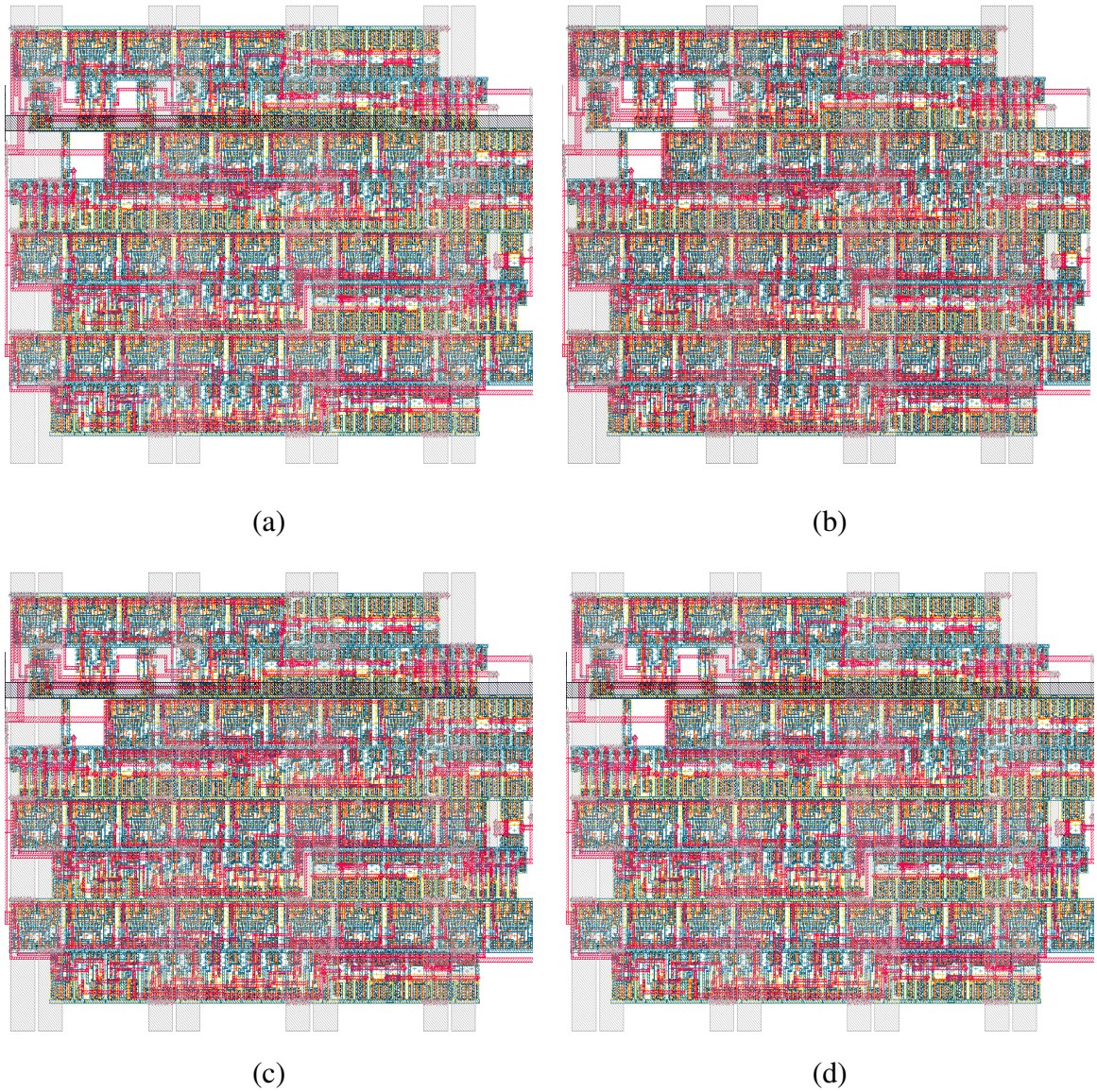
The layout and schematic of the current mirror and transmission gate switch used to modulate the current mirror are shown, respectively, in Figures 7.7 and 7.8. The four figures depict the three stacked device planes (Figure 7.7a), bottom device plane (Figure 7.7b), middle device plane (Figure 7.7c), and top device plane (Figure 7.7d). The output from the



(a) (b)

(c) (d)

Figure 7.3: Layout of the test circuit containing three interdigitated power and ground networks and the test circuits for generating and measuring noise. (a) Overlay of all three device planes, (b) power and ground networks of the bottom plane (tier 1), (c) power and ground networks of the middle plane (tier 2), and (d) power and ground networks of the top plane (tier 3).



(a) (b)

(c) (d)

Figure 7.4: Layout of the pattern sequence source for the noise generation circuits. (a) all three device planes, (b) noise generation circuits on the bottom plane (tier 1), (c) noise generation circuits on the middle plane (tier 2), and (d) noise generation circuits on the top plane (tier 3)

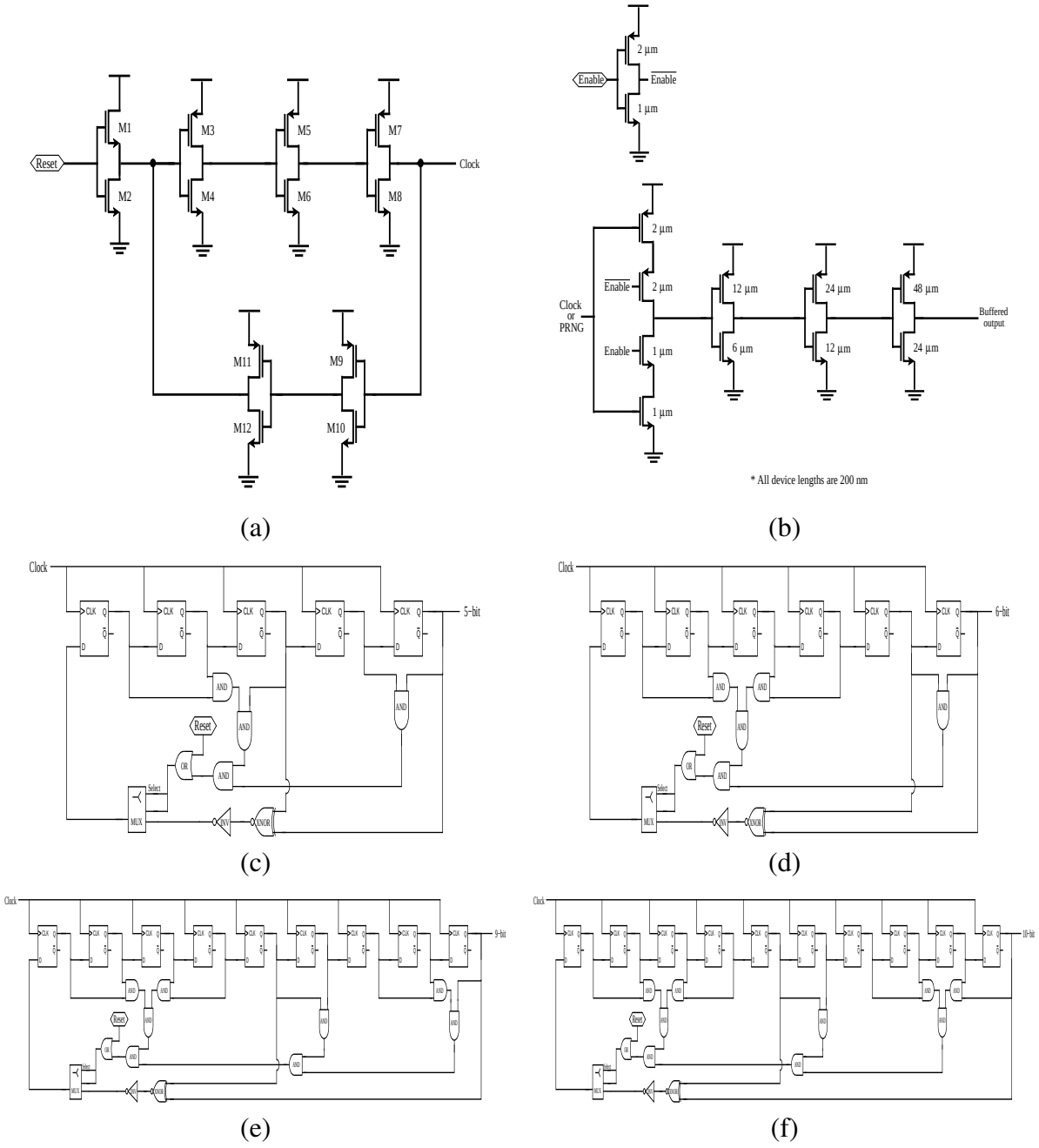


Figure 7.5: Schematic of the pattern sequence source for the noise generation circuits. (a) Ring oscillator, (b) buffer used for the RO and PRNG, (c) 5-bit PRNG, (d) 6-bit PRNG, (e) 9-bit PRNG, and (f) 10-bit PRNG.

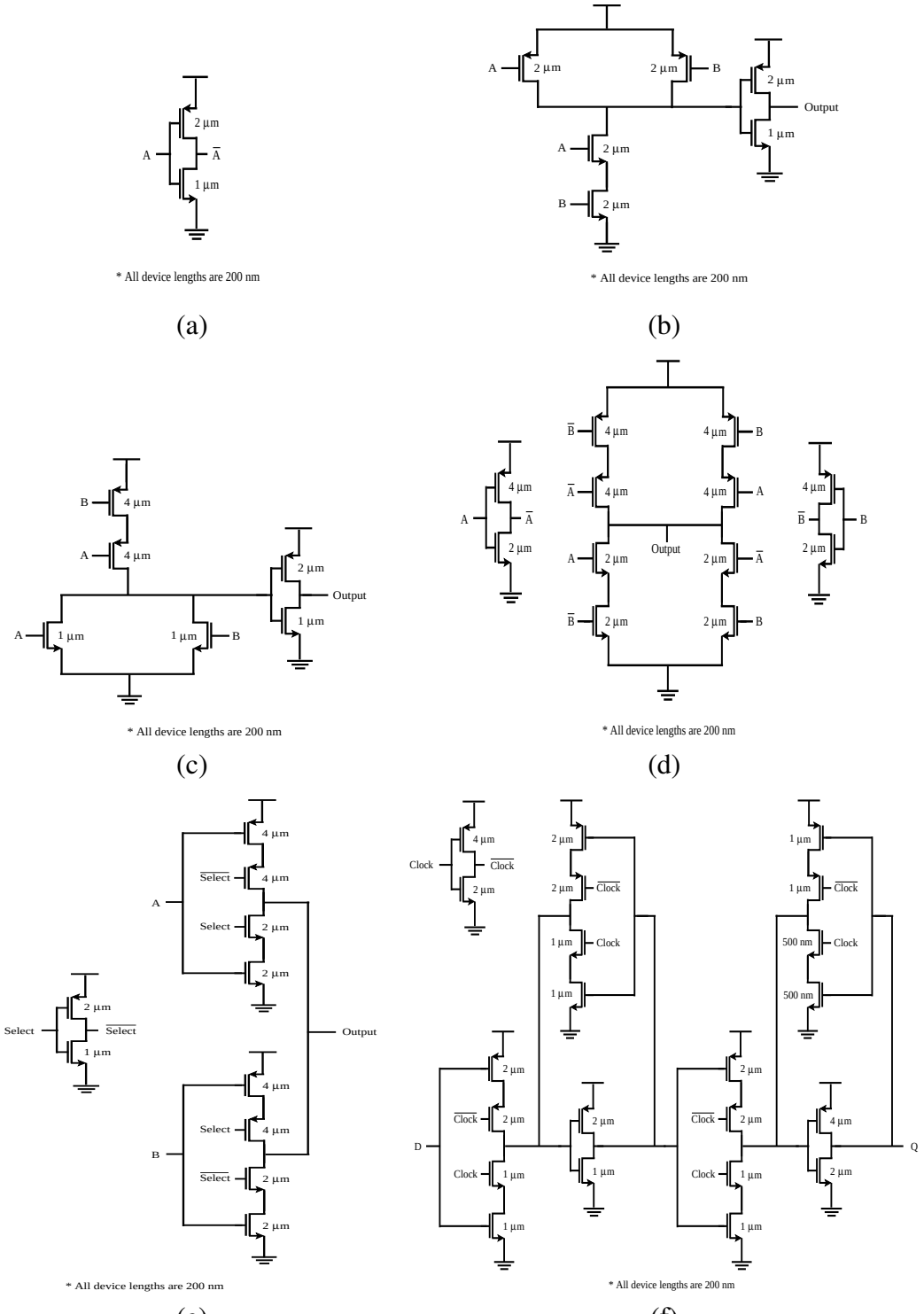
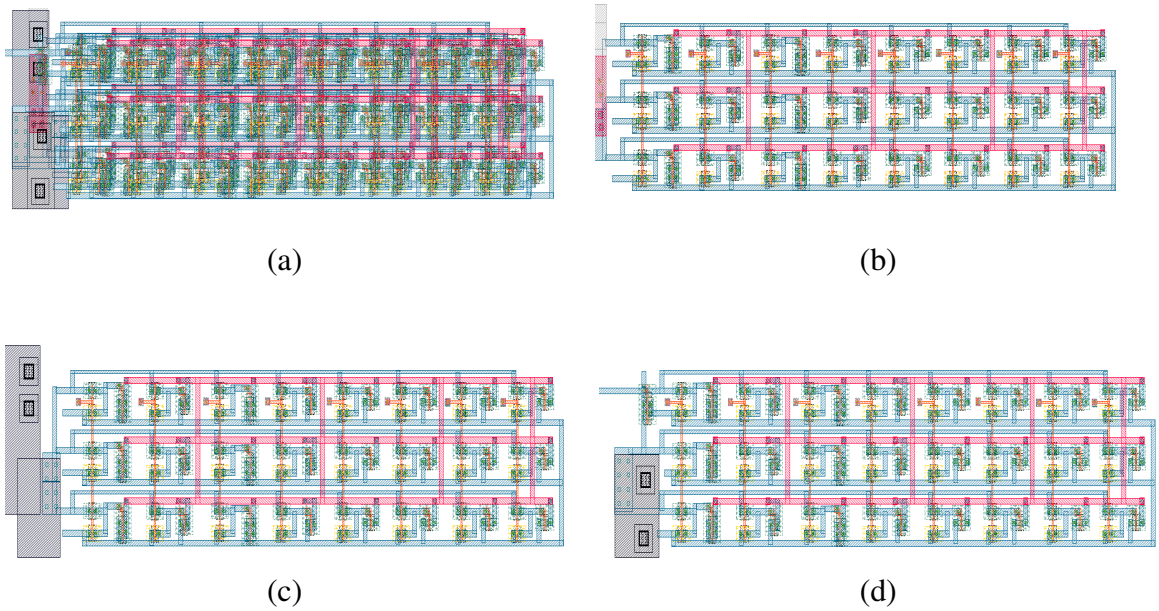


Figure 7.6: Schematic of the individual components in Figure 7.5c-f with the corresponding transistor sizes. (a) Inverter, (b) AND gate, (c) OR gate, (d) XNOR, (e) 2-to-1 MUX, and (f) D flip-flop.

Table 7.1: Length and width of the transistors within the ring oscillator shown in Figure 7.5a.

Ring Oscillator	Device	Type	L (nm)	W (nm)
250 MHz	M1, M2	NMOS	300	1500
	M3, M5, M9, M11	PMOS	1225	1000
	M4, M6, M10, M12	NMOS	1225	600
	M7	PMOS	300	3000
	M8	NMOS	300	1500
500 MHz	M1, M2	NMOS	300	1500
	M3, M5, M9, M11	PMOS	800	1000
	M4, M6, M10, M12	NMOS	800	600
	M7	PMOS	300	3000
	M8	NMOS	300	1500
1 GHz	M1, M2	NMOS	300	1500
	M3, M5, M9, M11	PMOS	500	1000
	M4, M6, M10, M12	NMOS	500	600
	M7	PMOS	300	3000
	M8	NMOS	300	1500



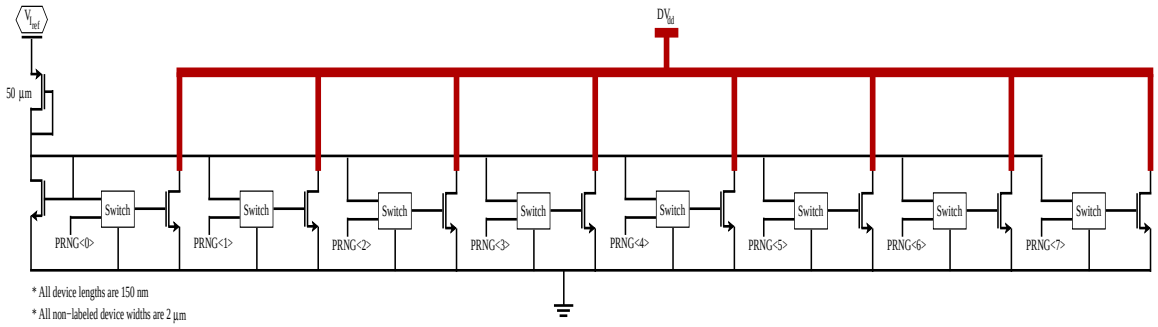
(a) (b)

(c) (d)

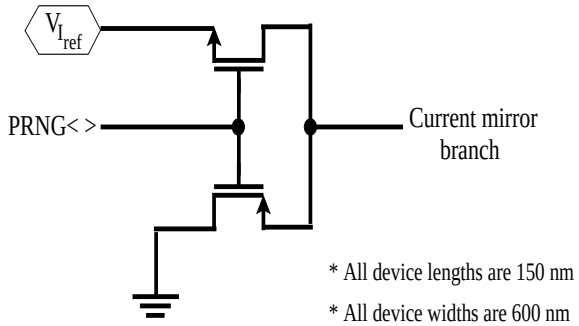
Figure 7.7: Layout of the noise generation circuits. (a) All three device planes, (b) noise generation circuits on the bottom plane (tier 1), (c) noise generation circuits on the middle plane (tier 2), and (d) noise generation circuits on the top plane (tier 3)

PRNGs is the input to the switches of the current mirror, as shown in Figure 7.8.

The layout of the power and ground noise detection circuits with the corresponding



(a)



(b)

Figure 7.8: Schematic view of the (a) current mirror, and (b) switches that vary the total current through the current mirror.

control logic is shown in Figure 7.9. The four figures depict the circuits on each of the three device planes (Figure 7.9a), bottom device plane (Figure 7.9b), middle device plane (Figure 7.9c), and top device plane (Figure 7.9d). As shown in the figure, the control logic is only present on the second device plane. The power and ground detection circuits and the control logic are both shown in Figure 7.9a.

A schematic of the logic that controls the RF output pads among the three device planes is shown in Figure 7.10. The signal that controls the RF pads is provided for both the power and ground detection signals on each device plane, and is rotated every 2^{18} cycles. The control of the output pads has four phases, three phases rotate control amongst the three device planes, and a fourth phase used to tie the RF output pads for both the power and ground detection circuits to ground. The fourth phase provides a delineating time during measurement to distinguish data from each of the three device planes as the four phases are cycled.

The test circuit occupies an area of 2 mm by 2 mm. Each power distribution network is 530 μm x 500 μm. DC and RF pads surround the four test blocks. A block diagram of the DC and RF pads is shown in Figure 7.11. The DC pads corresponding to the numbered pads shown in the figure are listed in Table 7.2. These DC pads provide the various bias voltages, reset signals, and power and ground signals. The RF pads surrounding the circuit blocks provide a pathway to probe and detect noise on the power and ground distribution networks. As shown in Figure 7.11, there are three sets of two RF pads surrounding the blocks, one set for each of the three different 3-D power distribution networks. The RF pads internal to each block are also shown in Figure 7.11. These RF pads are used to calibrate

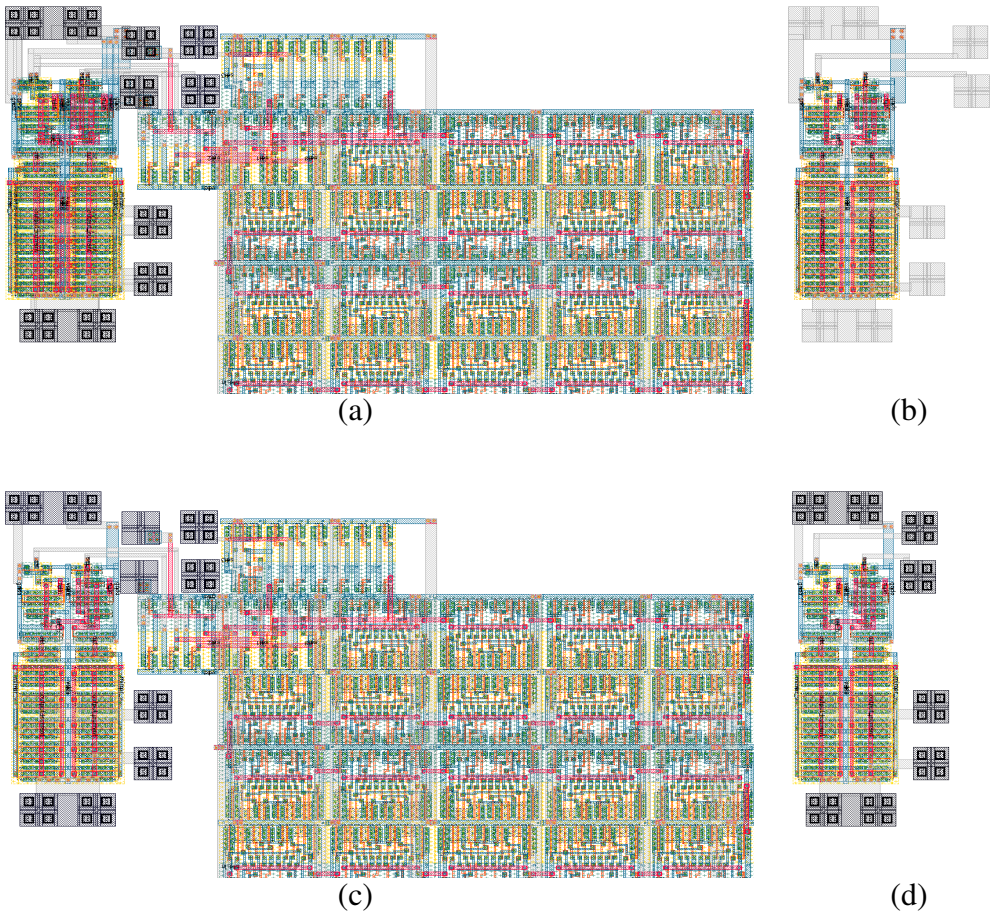


Figure 7.9: Layout of the power and ground noise detection circuits including the control circuit. (a) All three device planes, (b) power and ground sense circuits for the bottom plane (tier 1), (c) power and ground sense circuits for the middle plane (tier 2), and control circuit for all three planes, and (d) power and ground sense circuits for the top plane (tier 3).

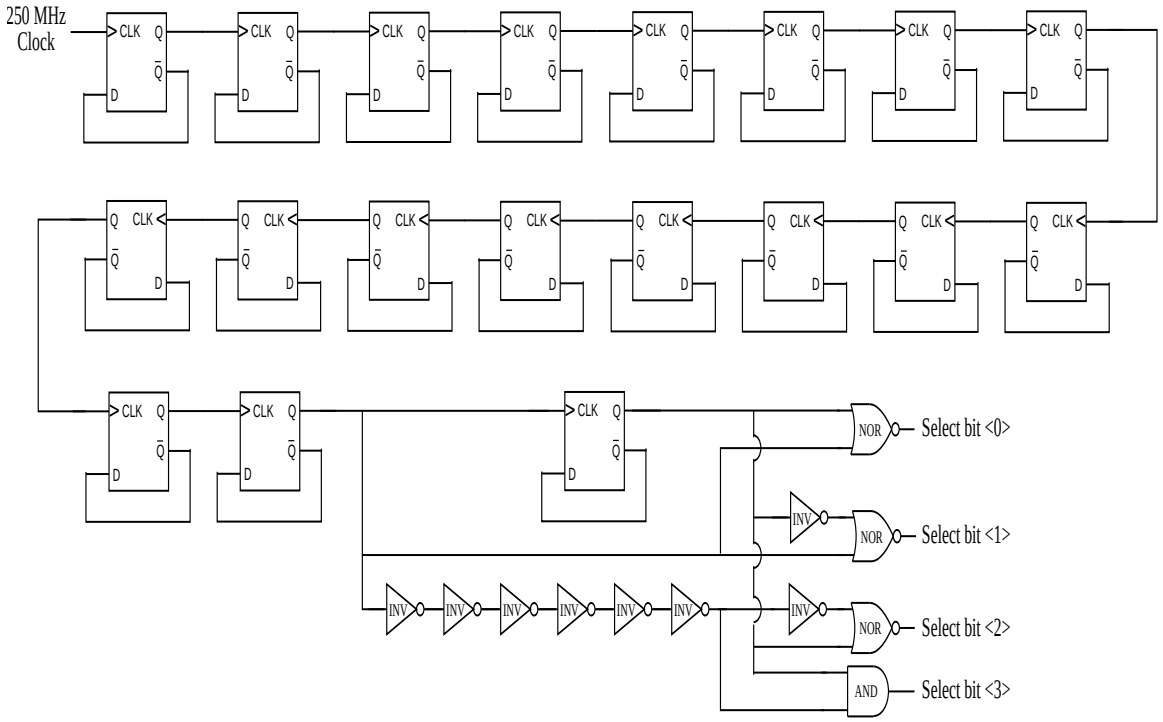


Figure 7.10: Schematic of the rotating control logic to manage the RF output pads among the three device planes. The control signals to the RF pads are provided for both the power and ground detected signals for each device plane.

Table 7.2: DC pad assignment of the 3-D test circuit. The pad index is shown in Figure 7.11.

Index	Pad Connectivity	Index	Pad Connectivity
1	3.3V V_{dd}	17	Analog V_{gnd}
2	1.5V V_{dd}	18	1.5V V_{dd}
3	V_{gnd}	19	V_{gnd}
4	$V_{I_{ref}}$	20	V_{reset}
5	V_{gnd}	21	1.5V V_{dd}
6	1.5V V_{dd}	22	V$_{g}nd$
7	Analog V_{gnd}	23	$V_{I_{ref}}$
8	Analog V_{dd}	24	Analog V_{dd}
9	$V_{I_{ref}}$	25	Analog V_{gnd}
10	V_{gnd}	26	1.5V V_{dd}
11	1.5V V_{dd}	27	V_{gnd}
12	V_{reset}	28	ESD V_{gnd}
13	V_{gnd}	29	ESD V_{dd}
14	1.5V V_{dd}	30	V_{reset}
15	$V_{I_{ref}}$	31	V_{gp}
16	Analog V_{dd}		

the source follower power and ground noise detection circuits.

There are two sets of five RF pads to calibrate the noise detection circuits, as shown in Figure 7.11. Each set of five pads is allocated as a ground-signal-ground-signal-ground (GSGSG) pattern. Two sets of GSG RF pads are within the set of five pads with the middle ground pad shared between the two sets of pads. One set of GSG pads is dedicated to the sense circuit detecting noise on the power lines, and the other GSG is dedicated to the sense circuit detecting noise on the ground line. In addition, the set of five RF pads closest to the center of the block is used to pass a sinuisodial signal from an Agilent E8364A general purpose network analyzer (PNA) to the noise sensing circuits. The other set of five RF pads is used as an output from the sense circuits to the PNA. The input and output RF pads for calibrating the power line noise detection circuit and the ground line noise detection circuit are enclosed in boxes to emphasize that the middle pad is a shared ground pad between the two circuits, as shown in Figure 7.11.

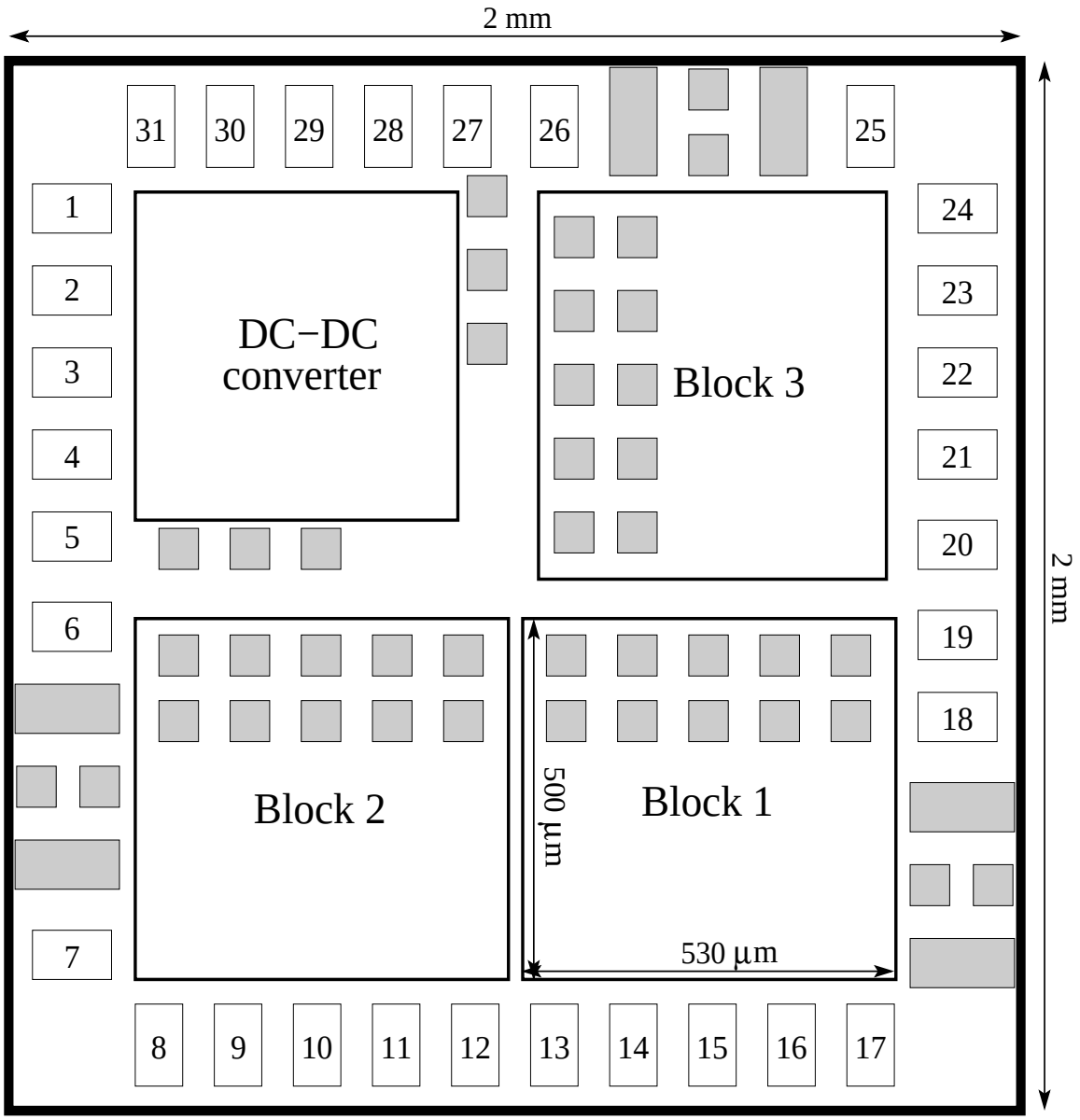


Figure 7.11: Block and I/O pin diagram of the DC and RF pad layout. The numbered rectangles are DC pads providing power and ground, DC bias points for the current mirrors, reset signals, and electrostatic discharge protection. The light colored squares and rectangles are RF pads used to calibrate the sense circuits (internal to the labeled blocks) and measure noise on the power/ground networks (external to the labeled blocks).

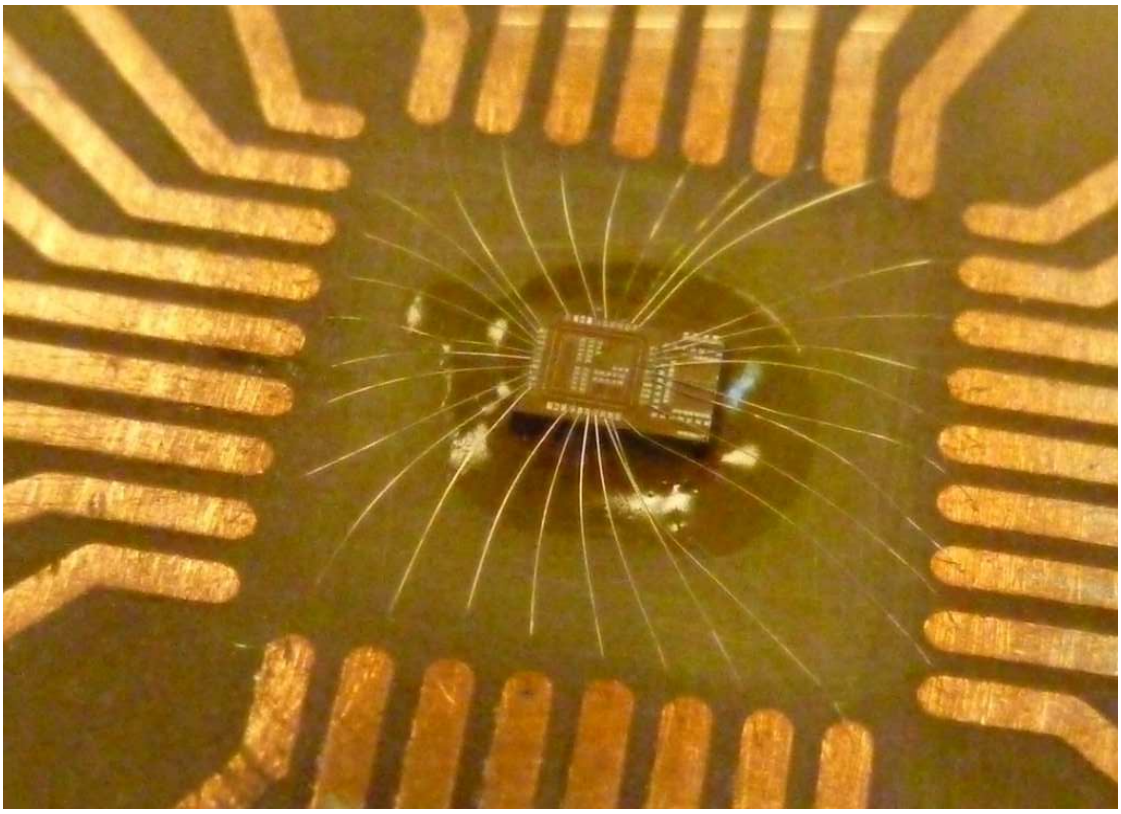


Figure 7.12: Wire bonded test circuit.

A microphotograph of the post wirebonded test circuit is shown in Figure 7.12. Wirebonding is performed on a West Bond 7400 bonder with ultrasonic bonding. The aluminum wirebonds range in length from approximately 0.7 cm to 1 cm.

7.1.3 3-D circuit architecture

The three power networks utilize identical on-chip circuitry, with two of the topologies only differing by the total number of TSVs required to distribute power to the lower two device planes. With one network, 1,728 TSVs are placed along both the inner and outer interdigitated power/ground lines, as shown in Figure 7.1(a), while the second network includes 1,152 TSVs located along the outer peripheral interdigitated power/ground lines,

as depicted in Figure 7.1(b). The third power network topology also includes 1,728 TSVs, but the interdigitated power network on the second stacked die is replaced with two metal planes, one for ground and one for power.

The basic logic blocks are illustrated in Figure 7.13. The power supply noise generators deliver different current to the power lines. These noise generators are placed on each plane of each power network topology. Voltage sense circuitry is included on each of the planes of each test block to measure the noise on both the power and ground lines. The second plane of the power delivery network, illustrated in Figure 7.1(c), does not include noise sense circuitry or noise generation circuitry as this plane is dedicated to power and ground. The voltage range and average voltage of the sense circuitry on each plane for each test block are compared for the three topologies illustrated in Figure 7.1.

A schematic of the on-chip circuitry on each plane of each power network is shown in Figure 7.13. The circuit is designed to emulate a power load drawing current, generating noise within the power and ground networks. Two sets of circuits are present, noise generating circuits and noise detection circuits. The noise generating circuits include ten current mirrors per device plane (fifteen for the second device plane), 250 MHz, 500 MHz, and 1 GHz ring oscillators, and five, eight, nine, and ten bit pseudorandom number generators (PRNG). The clock pulses generated by the ring oscillators are buffered and fed into the PRNGs, which are again buffered before randomly driving the current mirrors as enable bits. Each current mirror requires eight enable bits, with each enable bit turning on one branch of the current mirror which draws a maximum 4 mA current at increments of 0.5 mA (0.5 mA per branch). With ten current mirrors on each of the first and third device

planes and fifteen current mirrors on the second plane, the maximum current drawn by each power network is 140 mA at 1.5 volts, producing a maximum power density of 79 W/cm^2 in each 530 μm x 500 μm topology.

The noise detection circuits include separate circuits for evaluating power and ground noise. A schematic of the power and ground detection circuits is shown in Figure 7.14. These single stage amplifier circuits utilize an array structured noise detection circuit, as suggested in [311]. Each device plane contains both power and ground noise detection circuits. An off-chip resistor, labeled R_{bias} in Figure 7.14, biases the output node of the PMOS current mirror (with a PMOS threshold voltage of -0.56 V), and ensures that the current mirror operates in the saturation region. The resistor is chosen to ensure that the maximum voltage at the output node of the current mirror does not exceed 400 mV, the voltage at which the current mirror enters the triode region. The measured current at the output of the current mirror ranges between 200 milliamperes and 300 milliamperes. Since R_{bias} is chosen as 100 Ω, the maximum voltage at the output node is 300 mV. In addition, the intrinsic impedance Z_o, shown in Figure 7.14, is 50 Ω.

Two sets of ground-signal-ground (GSG) output pads source current from the final stage of the noise detection circuits of all three planes, as shown in Figure 7.13. One set of GSG output pads is dedicated to the three noise detection circuits that monitor the power network, while the other set of GSG pads supports the three noise detection circuits that monitor the ground network. Counter logic rotates the control of the output pads among the three device planes every 2^{18} cycles (at a 250 MHz clock) for both the power and ground noise detection circuits. A second isolated power and ground network with a 2 μF board

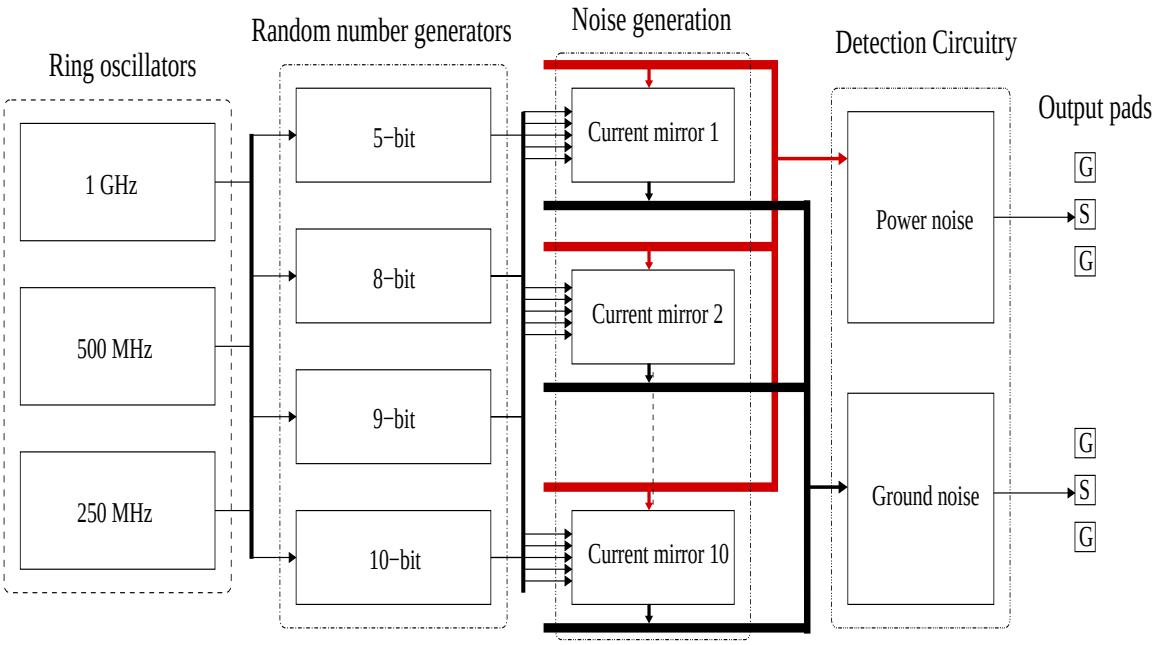
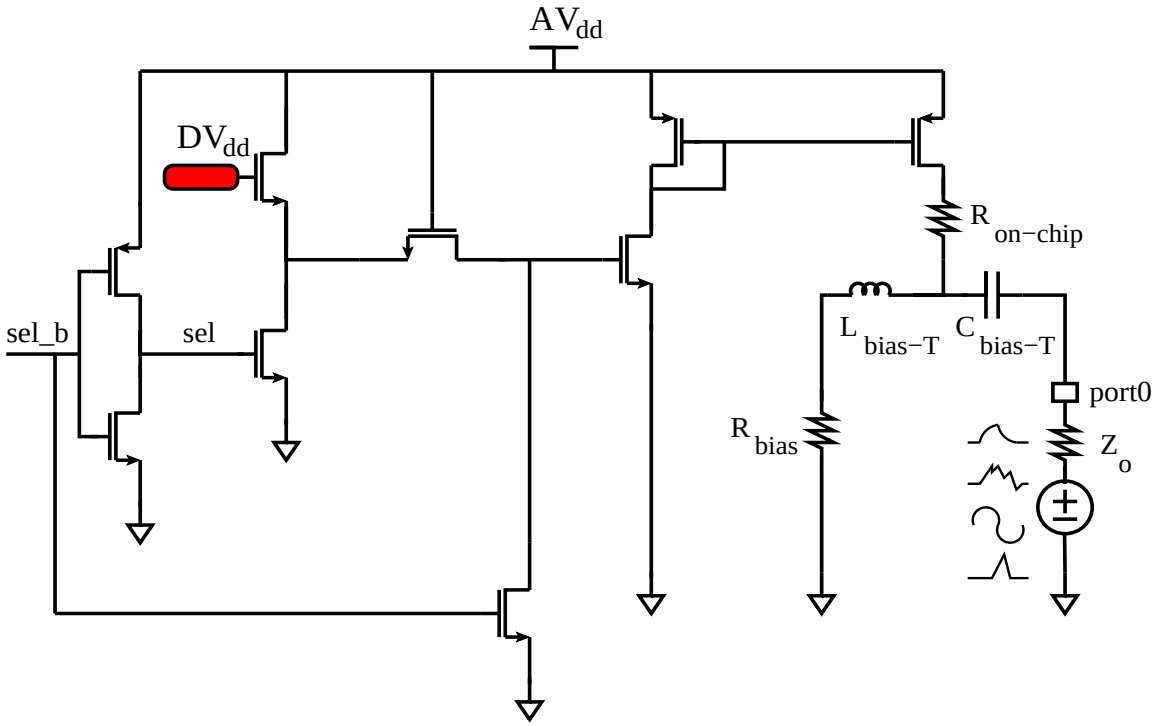


Figure 7.13: Block level schematic of noise generation and detection circuits.

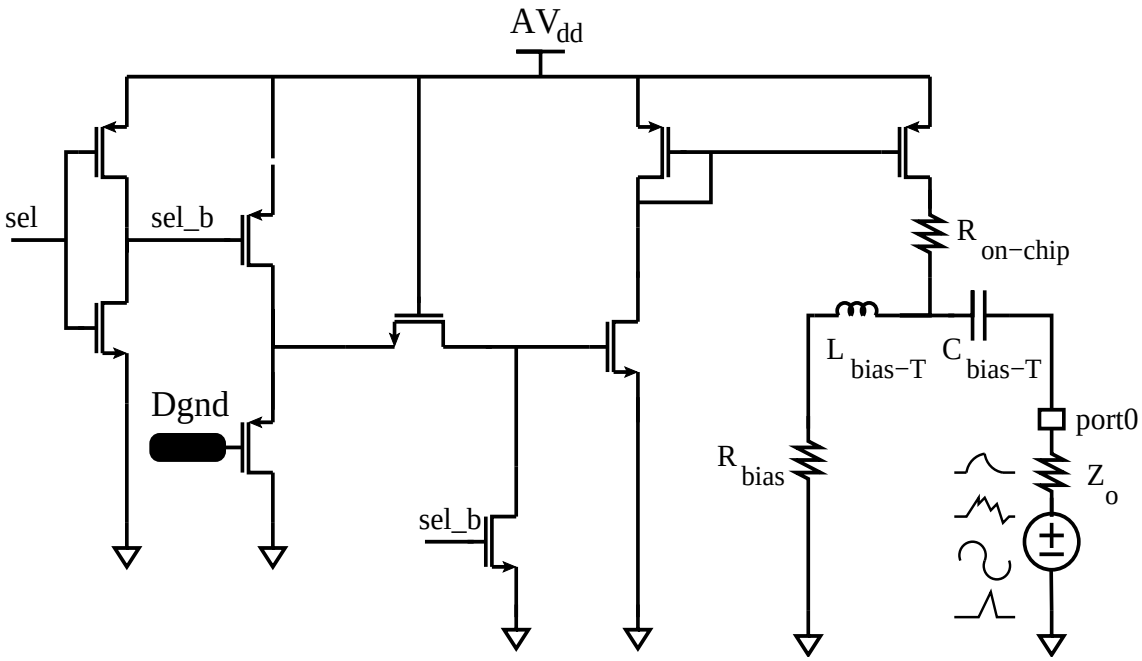
level decoupling capacitor ensures that the generated noise is not injected into the detection circuits.

7.2 3-D IC fabrication technology

The manufacturing process developed by MITLL for fully depleted silicon-on-insulator (FDSOI) 3-D circuits is described in [82, 83]. The MITLL process is a wafer level 3-D 150 nm integration technology with up to three FDSOI 150 mm wafers vertically bonded with TSVs to form a 3-D circuit. The operating voltage of the SOI transistors is 1.5 volts. The technology includes one polysilicon layer and three metal layers interconnecting the devices on each wafer. A backside metal layer also exists on the upper two planes, providing the starting and landing pads for the TSVs, and the I/O, power supply, and ground pads



(a)



(b)

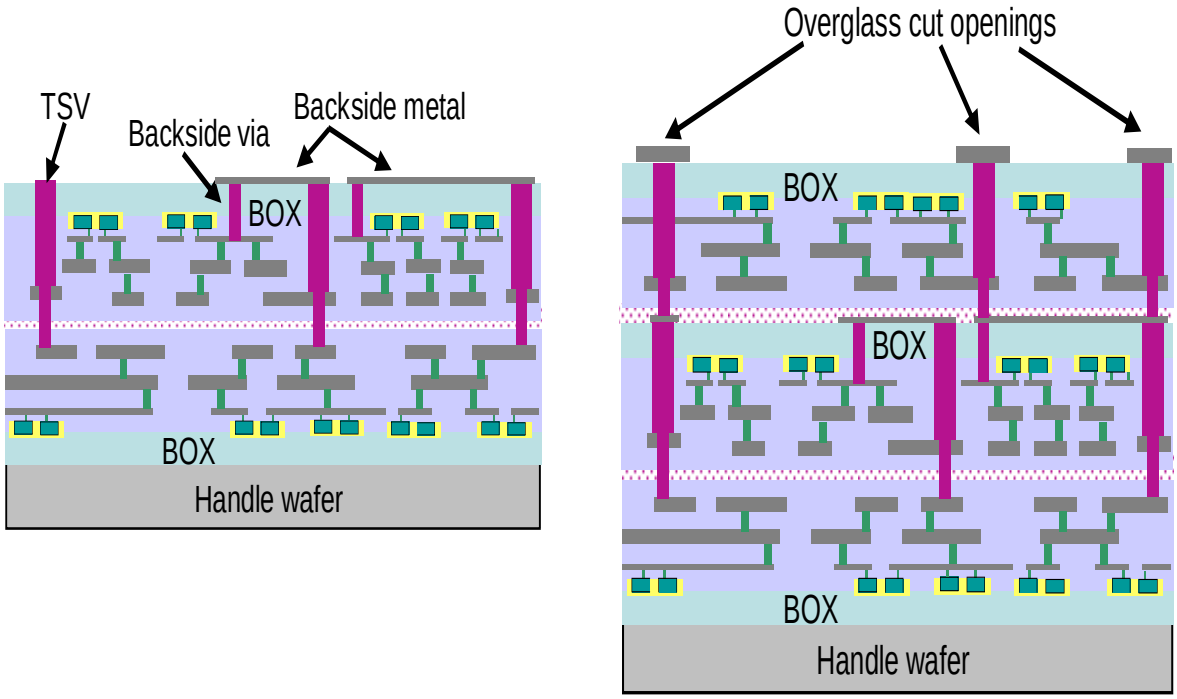
Figure 7.14: Source-follower noise detection circuits that detect noise on both the digital (a) power lines, and (b) ground lines.

for the overall 3-D circuit. An attractive feature of this process is the high density TSVs. The dimensions of the TSVs are 1.25 μm × 1.25 μm, and 10 μm deep, much smaller than many existing 3-D technologies [103, 312]. Two of the SOI wafers are thinned to less than 15 μm, while the third layer is used as a handle wafer with a substrate thickness of around 675 μm [82]. An intermediate step of the fabrication process is illustrated in Figure 7.15. As depicted in this figure, this process includes both face-to-face and face-to-back plane bonding. The SOI device layers are used for both monolithic [313] and wafer level 3-D integrated systems. SOI is an effective technology for 3-D circuits since the wafers can be aggressively thinned as compared to standard bulk CMOS technologies [314]. This capability results in significantly shorter TSVs, a critical issue in 3-D systems. The primary obstacle for 3-D SOI technologies is the high thermal resistance of the oxide which impedes the heat removal process [315]. A microphotograph of the fabricated die is shown in Figure 7.16.

7.3 Experimental results

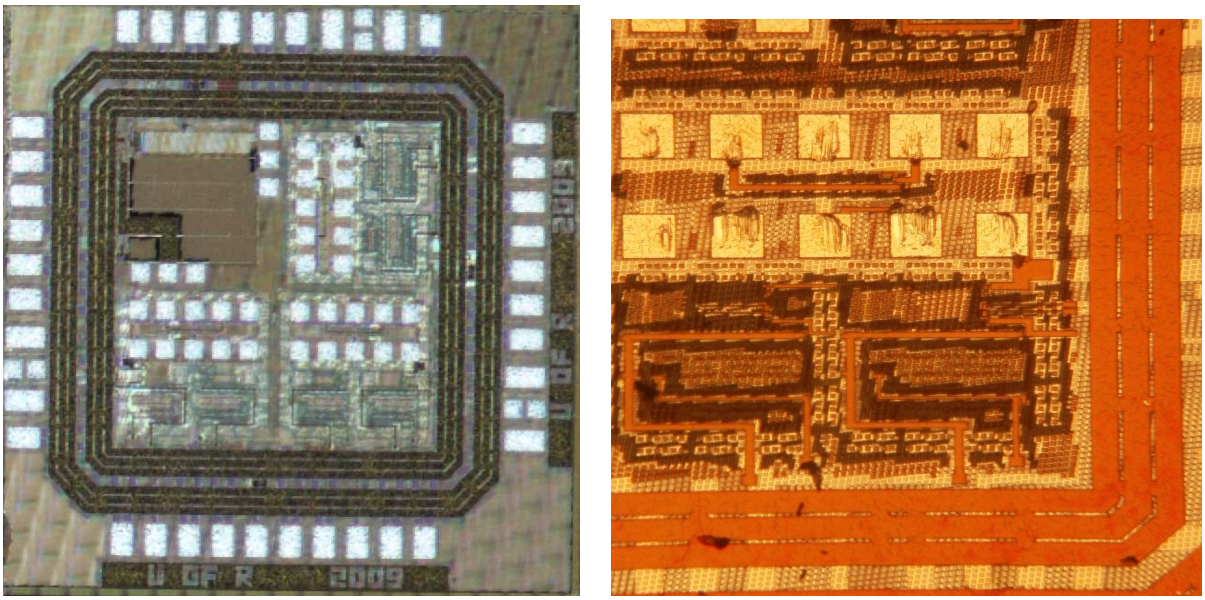
The noise generated within the power distribution network is detected by a source follower amplifier circuit. A schematic of the amplifier circuit is depicted in Figure 7.14. The sense circuits can detect a minimum voltage of 165 μV (from simulation). Noise from the digital circuit blocks is coupled into the sense circuit through the node labeled DV_{dd}. The gain of the circuit is controlled by adjusting the analog voltage, labeled AV_{dd} in Figure 7.14.

The gain and bandwidth of both the power and ground network noise detection circuits are calibrated by S-parameter extraction. The measured results are shown in Figure 7.17.



(a) (b)

Figure 7.15: Cross-sectional view of a 3-D circuit based on the MITTLL process, (a) intermediate step, and (b) fully fabricated 3-D stack [82]. The second plane is flipped and bonded with the first plane, while the third plane is bonded face-to-back with the second plane. The backside metal layer, vias, and through silicon vias are also shown [82].



(a) (b)

Figure 7.16: Fabricated test circuit examining noise propagation for three different power distribution networks and a distributed DC-to-DC rectifier, (a) die microphotograph of the 3-D test circuit, and (b) an enlarged image of Block 1.

The simulated DC gain and 3 dB bandwidth of the power network detection circuit are, respectively, -3.8 dB and 1.4 GHz. The measured DC gain and 3 dB bandwidth are, respectively, -4.1 dB and 1.3 GHz. Similarly, the simulated DC gain and 3 dB bandwidth of the ground network detection circuit are, respectively, -4.0 dB and 1.35 GHz, and the measured DC gain and 3 dB bandwidth are, respectively, -4.25 dB and 1.15 GHz. For both the power and ground detection circuits, the measured gain is within 3.4% of simulations. The models include the on-chip interconnect (15 Ω), the bias-T inductance (340 μH) and capacitance (3 μF), the resistance to bias the output node, and the 50 Ω intrinsic impedance Z_o of the network analyzer, as shown in Figure 7.14

The power spectral density of the generated noise within the power network with the voltage bias on the current mirrors set to 0.75 volts is shown in Figure 7.18. The noise data

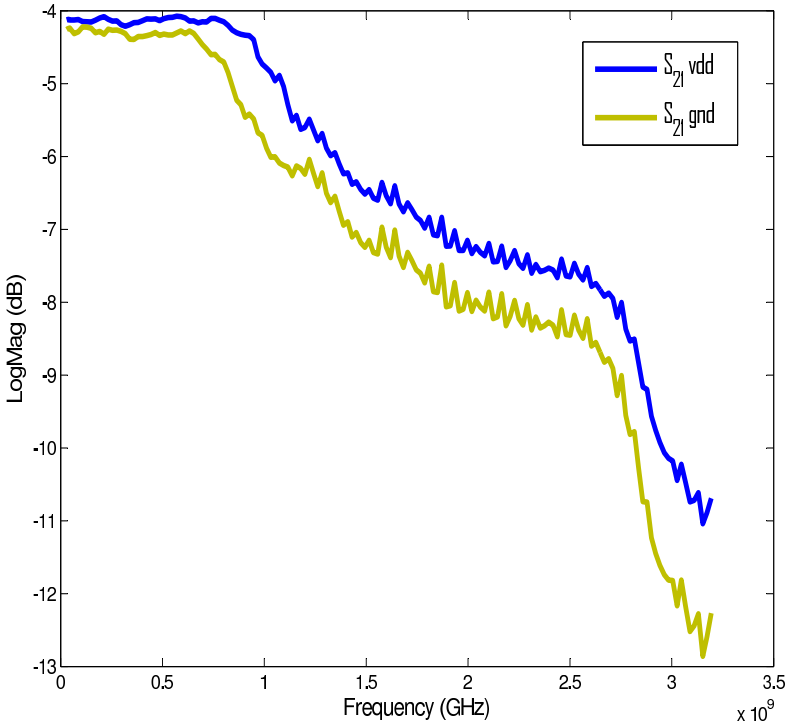
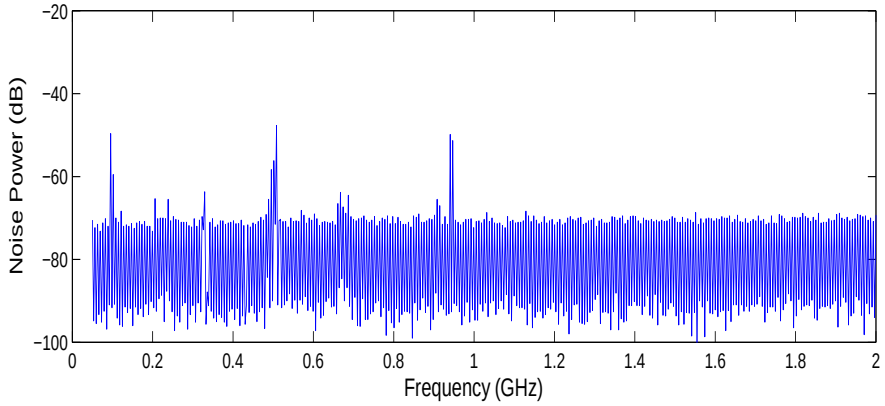
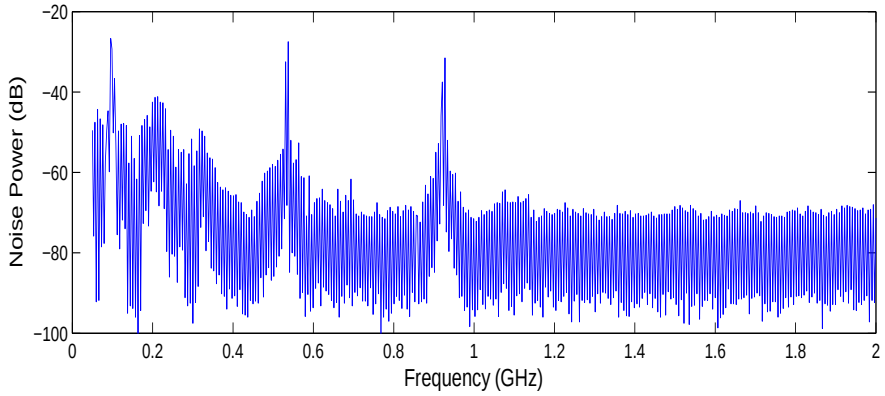


Figure 7.17: S-parameter characterization of the power and ground noise detection circuits.

shown on the top half of Figure 7.18 includes the effect of a 4 μF board level decoupling capacitance, whereas the noise data on the bottom half of Figure 7.18 is without any decoupling capacitance between the power and ground networks. Both plots illustrate the three noise components produced by the 250 MHz, 500 MHz, and 1 GHz ring oscillators. No on-chip decoupling capacitance is added to the three power distribution topologies other than the intrinsic capacitance of the power and ground networks. The peak noise power does not precisely match the ring oscillator frequencies as the ring oscillators are not tuned to the targeted 250, 500, and 1000 MHz frequencies. The peak noise therefore occurs at 97 MHz (-49 dB), 480 MHz (-47 dB), and 960 MHz (-50 dB) with a board level decoupling capacitor, and at 96 MHz (-27 dB), 520 MHz (-29 dB), and 955 MHz (-33 dB) without a decoupling capacitor. The inclusion of a board level decoupling capacitor reduces the peak



(a)



(b)

Figure 7.18: Spectral analysis of the noise generated on the power line of Block2 with (a) board level decoupling capacitance, and (b) without board level decoupling capacitance.

noise within the power networks by approximately 20 dB.

A time domain analysis of the generated noise is used to compare the three different power networks. The detected noise voltage is measured after the intrinsic capacitance of the bias T junction that couples the noise into the oscilloscope and spectrum analyzer (see node labeled port0 in Figure 7.14). The noise amplitude with increasing current mirror bias voltage (from 0 to 1 volt) is depicted in Figure 7.19. These results indicate that the current mirrors function properly. The 4,096 data points used to generate each subfigure shown in Figure 7.19 are centered around 0 volts for both the power and ground network, as only the

RF component is passed to the oscilloscope.

The average noise for each topology, with or without a board level decoupling capacitance, and for both the power and ground networks as a function of the applied bias voltage to the current mirrors, is shown in Figure 7.20. 4,096 data points are used to generate each average noise value for each topology at each of the six bias voltages. Since the current mirrors are activated using a random bit sequence generated from the random number generators, the 4,096 data points include voltages as low as the nominal off state (all zeroes bit sequence) to a peak voltage when all current mirrors are in the on state. Therefore, the lowest voltage (and the highest voltage as shown in Figure 7.21) in most cases is similar, however, the average voltage amplitude of the data set increases with each increase in the current mirror bias voltage. In all cases, with or without decoupling capacitors, the network topology that includes the metal planes reduces the average noise as compared to the other two topologies, as the metal planes behave as an additional decoupling capacitor. Results from a statistical analysis of the noise generated on the power network including the mean, median, and 25^{th} and 75^{th} quartiles are provided in Table 7.3. The number of TSVs also affects the magnitude of the noise, as shown in Figure 7.20. Reducing the number of TSVs by half increases the average amplitude of the noise by 2% to 14.2%, as the parasitic impedance of the three-dimensional power network is larger. The amplitude of the noise is less than twice as great since the portion of the impedance contributed by the TSVs along the path from the power supply, through the cables and wirebonds, into the 3-D power network, and back is a small fraction of the total impedance.

The peak noise for each topology, with or without a board level decoupling capacitance

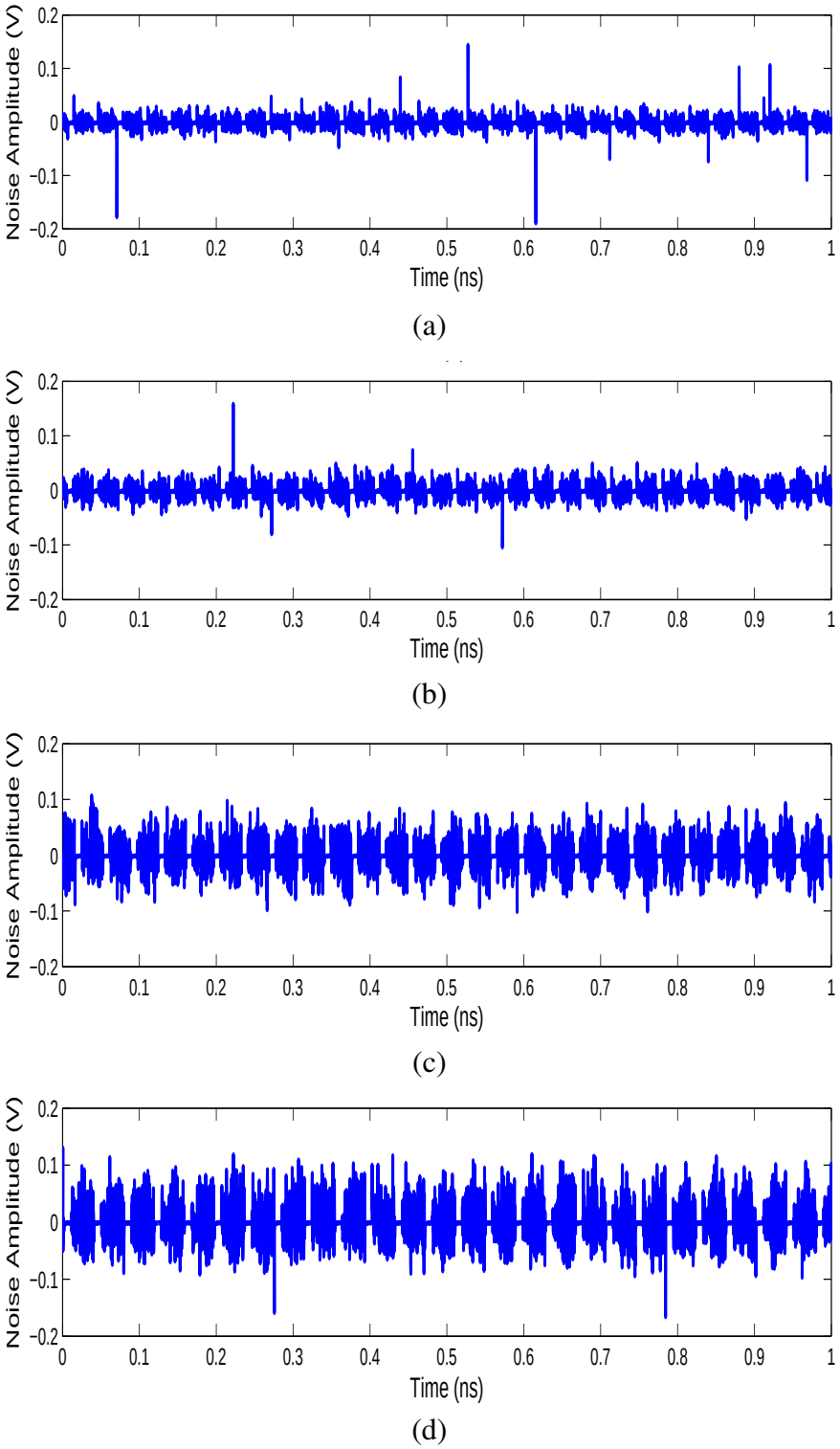


Figure 7.19: Time domain measurement of the generated noise on the power line of Block 2 without board level decoupling capacitance for (a) 0 volt, (b) 0.5 volt, (c) 0.75 volt, and (d) 1 volt voltage bias on the current mirrors.

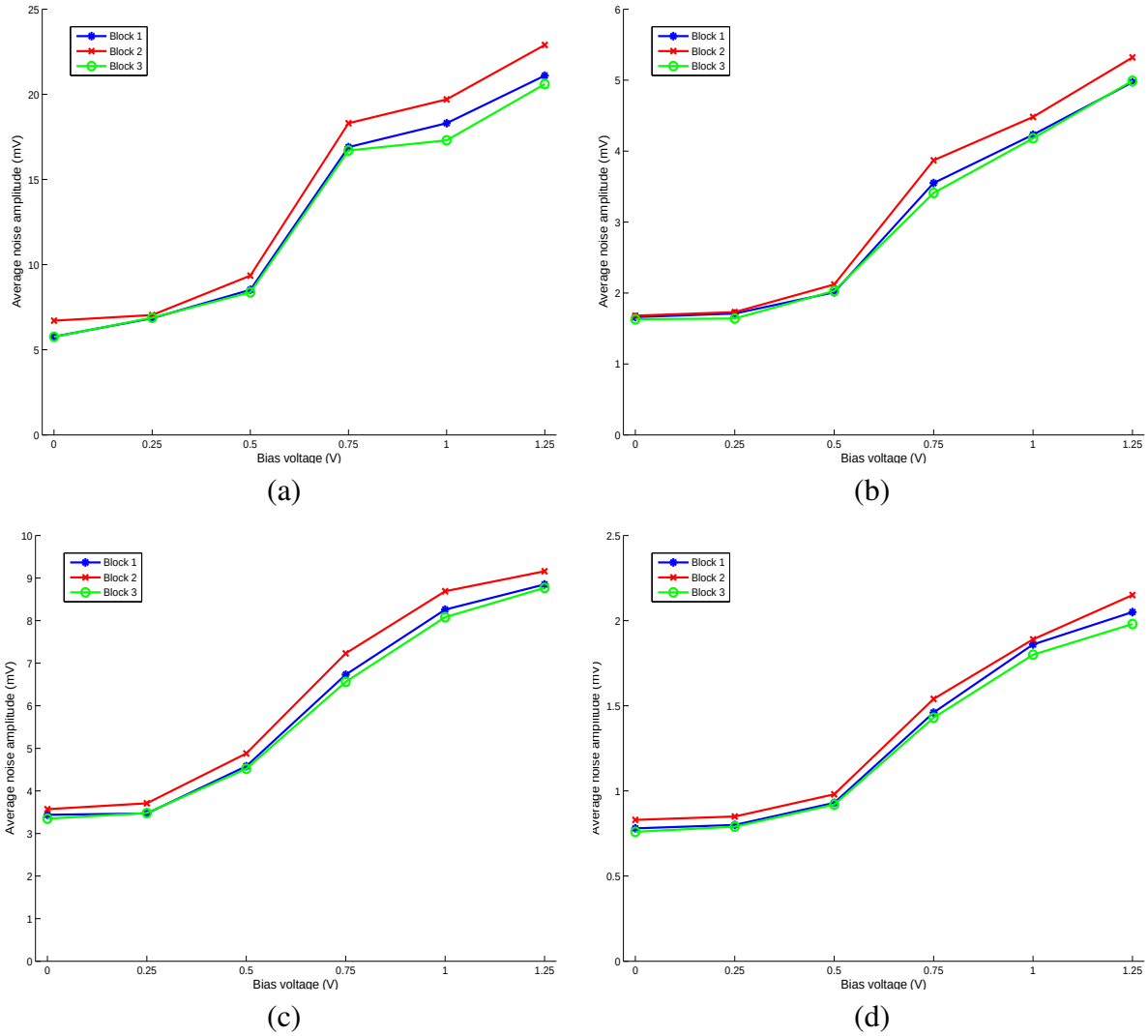


Figure 7.20: Average noise voltage on the power and ground distribution networks with and without board level decoupling capacitance. (a) Average noise of power network without decoupling capacitance, (b) average noise of power network with decoupling capacitance, (c) average noise of ground network without decoupling capacitance, and (d) average noise of ground network with decoupling capacitance. 4,096 data points are used to calculate the average noise value for each topology at each current mirror bias voltage.

Table 7.3: Statistical analysis of the noise generated on the power network. Sample size of 4096 points at each bias voltage.

Power network block	On-chip capacitor	Current mirror bias voltage (V)	Mean	25^{th} quartile	Median	75^{th} quartile
Block 1	No	0	5.77	1.65	4.50	8.84
		0.25	6.85	1.67	4.64	9.89
		0.5	8.52	1.93	6.64	13.17
		0.75	16.90	2.44	12.30	24.78
		1.0	18.30	2.18	12.95	26.50
		1.25	21.10	2.57	15.89	30.28
	Yes	0	1.66	0.80	1.10	2.93
		0.25	1.71	0.82	1.20	2.56
		0.5	2.01	0.99	1.74	3.18
		0.75	3.55	1.34	2.51	5.06
		1.0	4.23	1.59	3.09	6.74
		1.25	4.97	1.82	3.77	7.98
Block 2	No	0	6.66	1.74	4.56	9.54
		0.25	7.05	1.74	4.76	10.21
		0.5	9.31	2.03	6.72	13.87
		0.75	18.30	2.45	12.77	28.38
		1.0	19.01	2.22	13.03	28.94
		1.25	22.86	2.64	16.55	35.56
	Yes	0	1.68	0.87	1.24	2.99
		0.25	1.73	0.88	1.31	2.85
		0.5	2.12	1.14	1.88	3.55
		0.75	3.87	1.56	2.58	5.37
		1.0	4.48	1.77	3.23	7.05
		1.25	5.32	1.85	3.89	8.37
Block 3	No	0	5.76	1.65	4.48	8.80
		0.25	6.88	1.63	4.59	9.87
		0.5	8.37	1.94	6.60	12.89
		0.75	16.70	2.44	12.09	24.18
		1.0	17.30	2.06	12.89	26.22
		1.25	20.60	2.50	15.58	29.94
	Yes	0	1.63	0.77	1.10	2.88
		0.25	1.64	0.81	1.14	2.52
		0.5	2.03	0.99	1.70	3.11
		0.75	3.41	1.31	2.50	5.04
		1.0	4.18	1.60	3.02	6.69
		1.25	4.99	1.81	3.77	8.01

and for both the power and ground networks as a function of the applied bias voltage to the current mirrors, is shown in Figure 7.21. Unlike the average noise shown in Figure 7.20, the peak voltage detected from each topology does not follow a distinct pattern. No single topology contributes the largest noise voltage at any specific current mirror bias voltage. These voltages represent single data points indicative of the maximum peak-to-peak noise voltage for each topology at each bias voltage. For the average noise voltage, Block 3 produces a lower average noise than Block 1, which produces a lower average noise than Block 2. Interestingly, the average noise for each topology is approximately 75% to 90% lower than the peak-to-peak noise voltage, indicating that a majority of the noise data are located within close proximity of the nominal power and ground voltages. In addition, the saturation voltage of the detection circuitry at the output node (port0) is approximately 230 mV when the gain is -4.2 dB. The noise detection range is approximately 600 mV centered around 1.5 volts and 0 volts, respectively, for the power and ground lines. The detection circuits for the power network, therefore, detect noise that ranges from 1.2 volts to 1.8 volts, and for the ground networks, from -0.3 volts to 0.3 volts.

In addition to the peak and average noise voltage measurements for each power network topology, an analysis of the current drawn by the digital power network is presented. The measured current for each block within the test circuit as a function of the bias voltage applied to the current mirrors is listed in Table 7.4. Each topology draws a similar amount of current as the total number of current mirrors for each power network is identical. The data listed in Table 7.4 reveal that Blocks 1 and 2 sink similar currents. Block 3 draws about 30% to 45% more current as compared to the other two topologies. The Block 3

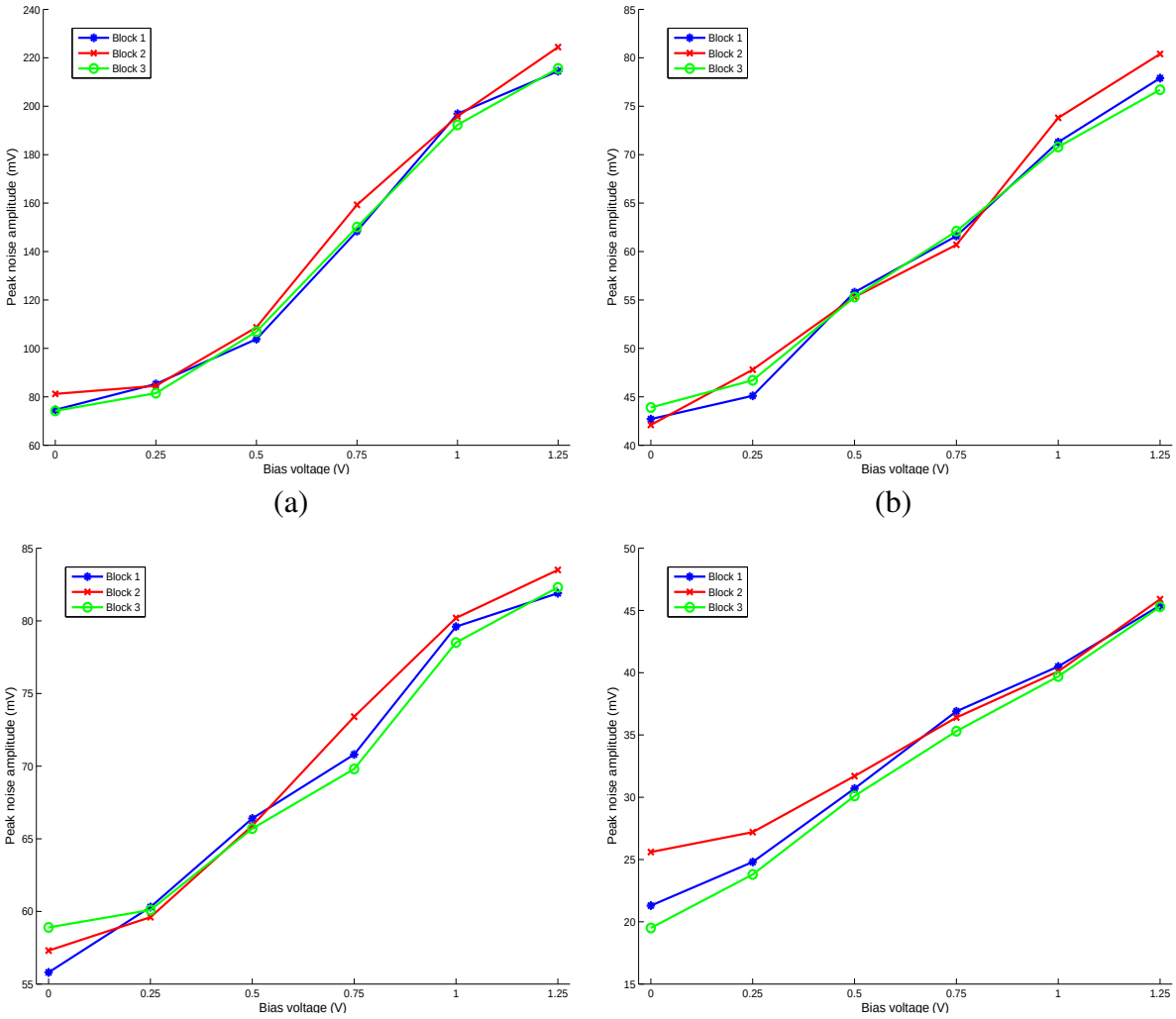


Figure 7.21: Peak noise voltage on the power and ground distribution networks with and without board level decoupling capacitance. (a) Peak noise of power network without decoupling capacitance, (b) peak noise of power network with decoupling capacitance, (c) peak noise of ground network without decoupling capacitance, and (d) peak noise of ground network with decoupling capacitance. A single peak data point (from 4,096 points) is determined for each topology at each current mirror bias voltage.

Table 7.4: Power supply current within the different power distribution topologies as a function of bias voltage on the current mirrors

Power network block	Current as function of bias voltage (mA)					
	0 volts	0.25 volts	0.5 volts	0.75 volts	1 volts	1.25 volts
Block 1	29.2 to 34.6	31.4 to 38.7	40.1 to 45.3	56.2 to 59.8	73.6 to 79.5	86.4 to 91.3
Block 2	32.6 to 38.5	33.6 to 40.1	43.7 to 47.8	59.2 to 66.3	75.5 to 82.5	88.6 to 93.8
Block 3	43.6 to 55.3	44.1 to 56.3	52.9 to 65.1	68.7 to 78.7	83.4 to 91.9	97.5 to 105.7

topology includes power and ground planes on the second device plane; therefore, in this topology, both leakage between these two large planes as well as leakage from the TSVs distributing power to the bottom plane contribute additional current. Based on the currents listed in Table 7.4 and a 1.5 volt power supply voltage, a peak power density of 59.8 W/cm^2 is achieved for Block 3 when the current mirrors are biased with a voltage of 1.25 volts.

7.4 Effect on 3-D power distribution topologies

The choice of power distribution topology affects the noise propagation characteristics of the power network, as indicated by the results described in Section 7.3. Additional insight into the design of the power networks (see Section 7.4.1), and a discussion on the effect of the design of the 3-D power distribution network based on the experimental results (see Section 7.4.2) is provided in this section.

7.4.1 Pre-layout design considerations

Several considerations are accounted for in the design of the power distribution topologies. One issue is the placement of a sufficient number of TSVs to ensure the mitigation of electromigration between planes. The total number of TSVs required to satisfy a target current density of 1×10^6 A/cm^2, a current load of 0.14 amperes, a TSV diameter of 1.25

μm, and a footprint of 530 μm by 500 μm (the footprint area of each power distribution network) verifies that the 576 and 864 TSVs per plane, respectively, for Block 2 and Blocks 1 and 3, far exceed the necessary twelve TSVs to mitigate electromigration. The number of TSVs to satisfy electromigration constraints is determined as follows,

$$Area_{target\ current\ density} = \frac{I_{load}}{J_{electromigration}}, \tag{7.1}$$

$$Number\ of\ TSVs = \frac{Area_{target\ current\ density}}{Area_{TSV}}. \tag{7.2}$$

The diameter of the TSV determines the "keep out zone" surrounding the TSV, an area where no devices can be fabricated. For the MIT Lincoln Laboratories 3-D technology, the keep out zone is approximately two times the diameter D. A square area with a width of 2 $\times$ D and depth of 2 $\times$ D is therefore used to determine the area penalty per TSV. An area of 1.36% and 2.04% of the 530 μm by 500 μm area is occupied by each power distribution topology on each device plane for, respectively, 576 and 864 TSVs. The area is the same for planes 2 and 3, where all I/Os originate from plane 3. No area penalty on device plane 1 exists as no TSVs are necessary in this front-to-back bonded plane (the TSVs land on metal 3 on plane 1). For a keep out zone equivalent to 3 $\times$ D, the area penalty is 3.06% and 4.58%, respectively, for 576 and 864 TSVs.

To determine the resonant frequency of each topology, a model of the 3-D power distribution networks includes the impedance of the cables, board, wirebonds, on-chip DC pads, TSVs, and the power distribution network on each device plane. The equivalent electrical

parameters are listed in Table 7.5, and the capacitance of the 3-D power distribution topologies is listed in Table 7.6. The impedances listed in Table 7.5 for the cables, board, and wirebonds are divided into three equivalent π-models to characterize the distributed nature of the lines, as shown in Figure 7.22.

The interdigitated power and ground lines in metal 3 are 530 μm long, 15 μm wide, and exhibit a sheet resistance of 0.08 Ω per sheet, producing a 2.83 Ω resistance. Similarly, the metal 2 lines are 500 μm long, 15 μm wide, with a sheet resistance of 0.12 Ω per sheet, producing a 4 Ω resistance. Each power and ground line is divided into eight equal *RLC* π-model segments to represent the distributed nature of the on-chip metal lines. An inductance of 1 pH is added to each π-model segment. There are three horizontal pairs (metal 3) of power and ground lines as well as three vertical pairs (metal 2) of power and ground lines for each interdigitated power distribution network on each device plane. The *RLC* parameters for a single π-segment of the interdigitated topology models the plane topology, but the connections between segments is modified to better characterize the power and ground planes, as shown in Figure 7.22.

The resonant frequency for the three different power distribution networks is determined with and without board level decoupling capacitors, and with and without an on-chip load capacitance (in addition to the intrinsic capacitance of the power distribution network). The resulting resonant frequency for the three power distribution topologies for the different capacitive configurations is listed in Table 7.7. The resonant frequencies reported in Table 7.7 account for the distributed inductance of the cables, board, wirebonds, and on-chip power distribution network. The on-chip inductance produces the highest resonant

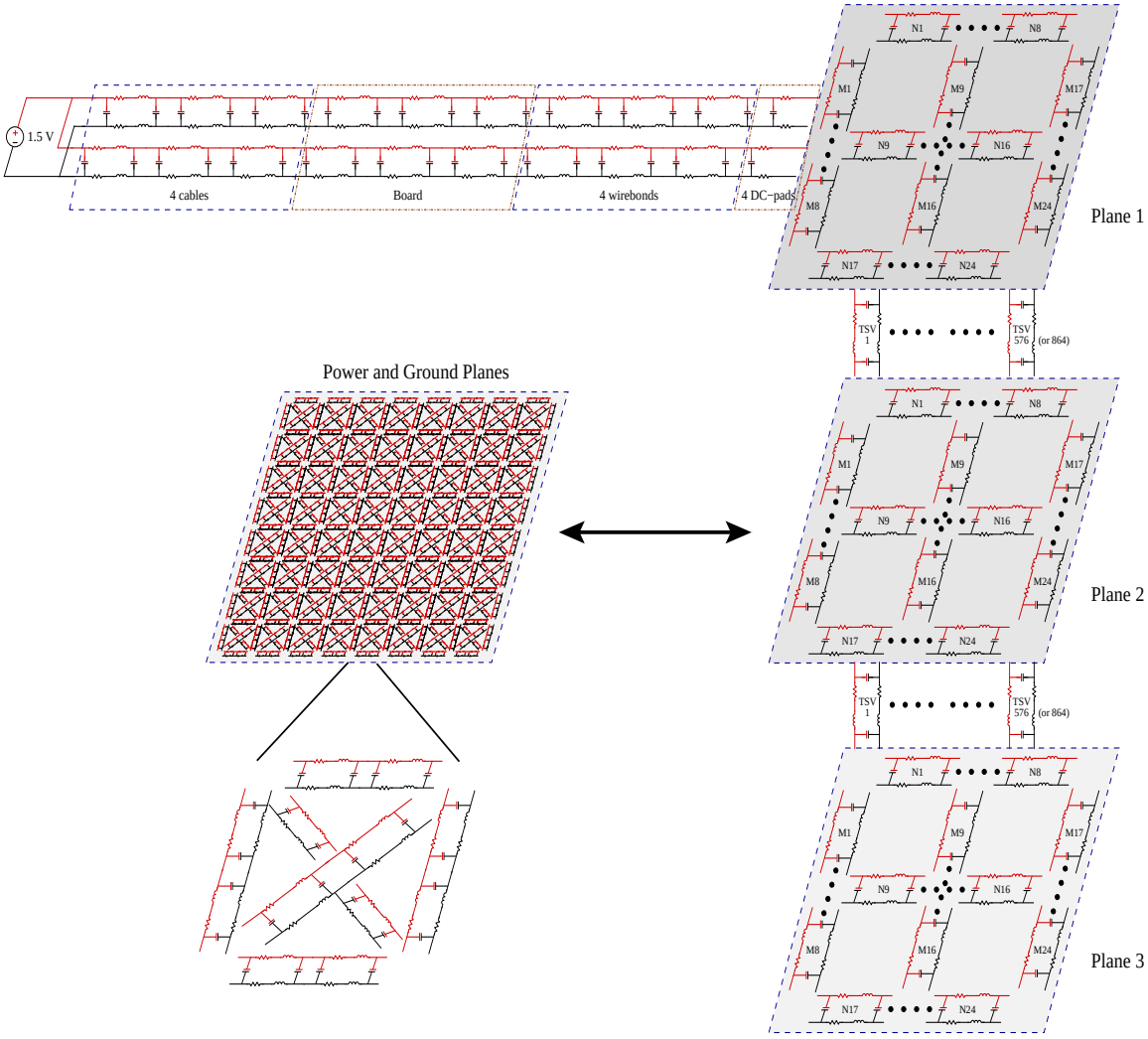


Figure 7.22: Equivalent electrical model of the cables, board, wirebonds, on-chip DC pads, power distribution networks, and TSVs.

frequency reported in Table 7.7 for any given block. The other three resonant frequencies, from the smallest to the largest, are dependent, respectively, on the cable, board, and wirebond inductances. The presence of a board-level decoupling capacitor eliminates all but the highest resonant frequency, the resonance caused by the on-chip inductance.

Table 7.5: Physical and electrical parameters of the cables, board, wirebonds, on-chip DC pads, power distribution networks, and TSVs.

Component	Width (μm)	Diameter (μm)	Length (mm)	Thickness (μm)	Resistance (Ω)	Capacitance (pF)	Inductance (nH)
Cables	—	1020	965.2	—	0.072	10.59	1445.75
Board	760	—	40	43.2	0.021	1.88 *	26.48
Wirebonds	—	25.4	4 to 5	—	0.278	0.0233	5.92
DC pads	80	—	0.12	2	—	0.0717	—
Pad to power grid	60	—	0.260	0.63	0.52	—	—
TSV	—	1.25	0.009	—	0.4	0.124×10^{-3} **	5.55×10^{-3}

* An additional 4 μF when board level decoupling capacitors are added. ** Coupling capacitance between two TSVs [134].

Table 7.6: Capacitance of the three different power distribution blocks, and interdigitated and power/ground planes.

Block or topology	Power or ground	Capacitance (fF)	Capacitance with load (fF)
Interdigitated	V_{dd}	330.09	—
	V_{gnd}	330.08	—
Planes	V_{dd}	1000.35	—
	V_{gnd}	1158.24	—
Block 1	V_{dd}	993.38	1631.68
	V_{gnd}	993.39	1596.38
Block 2	V_{dd}	964.37	1610.37
	V_{gnd}	964.38	1574.18
Block 3	V_{dd}	1862.74	2487.10
	V_{gnd}	1949.60	2590.36

Table 7.7: Resonant frequency of the three different power distribution networks with and without board level decoupling capacitors, and with and without an on-chip load capacitance.

Block	Decap (μF)	Extracted load (fF)	Resonant frequencies (GHz)
Block 1	none	none	0.224, 0.447, 0.794, 1.25
	4	none	1.25
	none	638	0.199, 0.447, 0.708, 1.25
	4	638	1.25
Block 2	none	none	0.223, 0.447, 0.794, 1.78
	4	none	1.78
	none	646	0.199, 0.447, 0.708, 1.78
	4	646	1.78
Block 3	none	none	0.224, 0.447, 0.794, 1.67
	4	none	1.67
	none	624	0.199, 0.447, 0.708, 1.67
	4	624	1.67

7.4.2 Design considerations based on experimental results

Based on the experimental results, both doubling the number of TSVs and utilizing a dedicated power and ground plane lowers the power noise. Although the noise is lower with both a greater number of TSVs and dedicated power network planes (2% to 14.2% lower noise, as listed in Table 7.8), the power noise is limited by the small series resistance of the 3-D power network as compared to the larger series resistance of the cables and wirebonds, as described in Section 7.4.1.

Two issues in 3-D power distribution networks are considered based on the experimental results: 1) the benefit of the power and ground planes to justify the use of two metallization levels, and 2) the benefits and drawbacks of increased TSV density on area and reducing noise. Addressing 1), the power planes provide an additional reduction of 0.2% to 5.5% in average noise as compared to the fully interdigitated topology. The reduction in average noise is primarily due to the increased capacitance of the power and ground networks, as indicated in Table 7.6. The average noise characteristics of the planes topology can be further improved by increasing the size of the power/ground planes or using multiple metal layers for both power and ground. The use of full planes to deliver current to the load does require significant metal resources, and therefore this cost must be considered. In addition, the power planes complicate the design of the signal interconnects, where holes are required in the power/ground planes to pass signals between device planes. For those applications where additional noise reduction is required, additional metallization and design complexity may be justified.

The TSV density provides the greatest reduction in average noise, ranging from 2.7%

Table 7.8: Per cent reduction in average noise from the noisiest topology (Block 2) to the least noisiest topology (Block 3) as a function of bias voltage on the current mirrors.

Power or ground	Decap present	Noisier Block	Quieter Block	Per cent noise reduction as function of bias voltage 0 V	0.25 V	0.5 V	0.75 V	1 V	1.25 V
Power	No	2	1	14.0	2.7	8.9	7.7	7.1	7.9
		2	3	14.2	2.3	10.5	8.7	12.2	10.0
		1	3	0.2	-0.4	1.8	1.2	5.5	2.4
Power	Yes	2	1	1.2	1.2	5.2	8.3	5.6	6.6
		2	3	3.0	5.2	4.2	11.9	6.7	6.2
		1	3	1.8	4.1	-1.0	3.9	1.2	-0.4
Ground	No	2	1	3.6	6.5	6.1	6.9	4.9	3.4
		2	3	6.2	6.2	7.4	9.3	7.0	4.3
		1	3	2.6	-0.3	1.3	2.5	2.2	0.9
Ground	Yes	2	1	6.0	5.9	5.1	5.2	1.6	4.7
		2	3	8.4	7.1	6.1	7.1	4.8	7.9
		1	3	2.6	1.3	1.1	2.1	3.3	3.4

to 14.0%, as listed in Table 7.8. The additional area penalty when increasing the number of TSVs by 50% from 576 to 864 is 0.68% (1.36% as compared to 2.04%), noting that the area penalty is also dependent on the keep out region, as described in Section 7.4.1. The increased area due to the higher TSV density is small as compared to the large reduction in average noise.

7.5 Summary

The design of a power distribution network for application to 3-D circuits is considerably more complex than the design of a two-dimensional power network. The preferable topology of a power distribution network is not dictated by a single design objective but rather by the overall 3-D system level requirements including the availability of I/O pins, and number of bonded planes. This test circuit provides enhanced understanding of topology dependent noise generation and propagation in 3-D power delivery systems.

A three-dimensional (3-D) test circuit examining power grid noise in a 3-D integrated stack has been designed, fabricated, and tested. Three topologies to distribute power within

a 3-D circuit have been evaluated, and an analysis of the peak noise voltage, voltage range, average noise voltage, and resonant frequency characteristics for both power and ground is described. Fabrication and vertical bonding were performed by MIT Lincoln Laboratory for a 150 nm, three metal layer SOI process. Three wafers are vertically bonded to form a 3-D stack. A noise analysis of three power delivery topologies is described. Calibration circuits for a source follower sense circuit compare the different power delivery topologies as well as the individual 3-D circuits. The effect of the TSV density on the noise profile of a 3-D power delivery network is experimentally described. A comparison of the peak noise and resonant behavior for each topology with and without board level decoupling capacitors is provided and suggestions for enhancing the design of the power delivery network are offered. This test circuit provides enhanced understanding of topology dependent noise generation and propagation in 3-D power delivery systems.

Chapter 8

Power Delivery With Decoupling Capacitors in 3-D Integrated Circuits

Power delivery, in the simplest representation, consists of a power supply, a power load, and interconnect connecting the power supply to the load [141, 161, 162]. The power distribution network, the interconnect lines that connect the supply to the load, must be designed to ensure that all of the circuit blocks on an IC are powered properly while satisfying the required noise specifications of each block. Power supply noise deleteriously affects the operation of an integrated circuit through: 1) producing variations in the delay of clock and data signals, 2) increasing the uncertainty of the on-chip timing reference signal, thereby lowering the clock frequency, 3) reducing noise margins, and 4) lowering the maximum supply voltage, again decreasing the maximum clock frequency [141, 161, 162].

Established techniques such as isolated power networks or decoupling capacitor placement minimize the noise generated on the power network and are regularly applied to 2-D ICs. The strategic placement of decoupling capacitors within the power network quickly supplies (or dampens) current when noise is present. Placing a hierarchy of decoupling capacitors maintains the impedance of the power network below a target specification over

the entire range of operating frequencies [141, 162].

Although decoupling capacitor placement is standard practice in modern 2-D ICs, this approach is not yet used in 3-D ICs. Much work is necessary to better understand which decoupling capacitor placement practices are best suited for 3-D integrated circuits. There is potential to exploit the vertical direction when placing localized capacitors within a device plane as the vertical separation of active devices can be less than 40 μm [83, 316]. A test circuit is currently in fabrication by Tezzaron Semiconductor in a 130 nm CMOS technology with 1.2 μm diameter TSVs. Specific test structures have been developed to evaluate decoupling capacitor placement within a 3-D system.

In the following section, a description of the Tezzaron 3-D IC fabrication process is provided. An overview of the test circuit is described in Section 8.2. A description of the test circuit examining decoupling capacitor placement is provided in Section 8.3. Some conclusions are offered in Section 8.4.

8.1 3-D IC fabrication technology

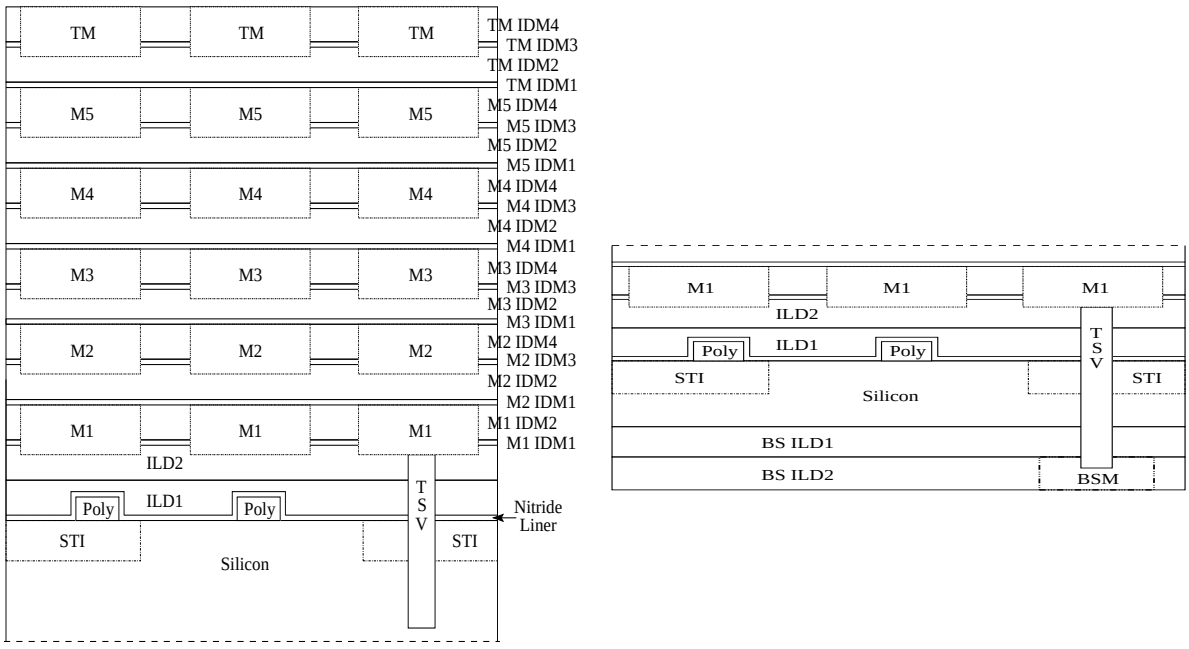
The manufacturing process developed by Tezzaron bonds five wafers/dies fabricated by Chartered Semiconductor (currently a subsidiary of GlobalFoundries) to form a five layer 3-D integrated stack. Two bulk wafers are wafer-to-wafer bonded to form the two layer logic stack, with the remaining three wafers used to produce the DRAM memory controller and two DRAM device planes [317–320]. Both logic wafers, WTOP and WBOTTOM, are initially at full thickness, about 760 μm, and not thinned before bonding. After bonding, the WTOP wafer, in preparation for connection with the memory stack, is thinned to reach

Table 8.1: Conductor layer properties for the two logic layers.

Conductor	Thickness (μm)	Variation (+/- %)	Min Width (μm)	Variation (+/- %)	Min Spacing (μm)
TM	0.86	15	0.44	10	0.46
M5	0.42	15	0.20	15	0.21
M4	0.42	15	0.20	15	0.21
M3	0.42	15	0.20	15	0.21
M2	0.42	15	0.20	15	0.21
M1	0.31	15	0.16	15	0.18
Poly	0.16	10	0.13	10	0.18
BSM	0.7	–	–	–	–

the through silicon via (TSV). After thinning, WTOP is about 12 μm thick, with about 6 μm of metal and wiring and 6 μm of bulk silicon with TSVs. The WBOTTOM wafer is thinned to about 750 μm, producing a stacked thickness of approximately 765 μm. A cross section of WTOP and WBOTTOM pre-thinned is shown in Fig. 8.1a, and a cross section after thinning and metallization is shown in Fig. 8.1b. The thickness, minimum width, and spacing of the five metal layers, top metal, polysilicon, and backside metal are listed in Table 8.1. The bond between WTOP and WBOTTOM is a face-to-face wafer bond using a thermodiffusive copper-to-copper bonding process [321, 322]. No TSVs are required for this bond. A preestablished bond point pattern in top metal (TM shown in Fig. 8.1) is provided. Each bond point is 3 μm by 3 μm with a 5 μm pitch between bond points. The bond points are present across the entire 5 mm by 5 mm area of the logic dies. The bond points used for interplane signaling between WTOP and WBOTTOM are selected by routing standard in-plane vias to the TM from any of the in-plane metal layers.

The memory controller layer and the two DRAM layers are similarly processed. Two DRAM wafers are first face-to-face bonded with the same wafer-to-wafer thermocompressive process used for the two logic device planes. No TSVs are required for the bond between the two DRAM layers. The top DRAM wafer is thinned to 12 μm from the



(a) (b)

Figure 8.1: Crosssection of (a) top of a non-thinned wafer, and (b) bottom of wafer after thinning and metallization. Depicted in the figure are metals 1 through 5 (M1 to M5), polysilicon (poly), through silicon vias (TSV), backside metal for TSVs (BSM), backstop interlayer dielectric (BS ILD), and shallow trench isolation (STI).

original nominal thickness of 750 μm, exposing the TSVs on the backside. The memory controller layer is bonded to the two DRAM layers in a face-to-back wafer-to-wafer process [317, 323, 324]. Thinning the memory controller layer to 12 μm is performed to bond the logic layers to the memory stack (two DRAM layers and one memory controller layer). The WTOP and memory controller layers are both 12 μm thick with TSVs exposed for inter-plane signaling. The exposed TSVs are covered with a backside copper layer. The backside metal is also a predefined pattern, with 2 μm by 2 μm landing pads for the 1.2 μm diameter TSVs. These backside metal pads provide both the interplane signaling paths and the bonding points between the logic layers and memory layers. The face-to-face bonded interface between the two logic dies and the back-to-back bonded interface between WTOP and the memory controller are shown in Fig. 8.2. The complete stack including the two logic layers, the memory controller layer, and the two DRAM layers is depicted in Fig. 8.3.

Through silicon vias are necessary for interplane signaling between the two logic layers and the three memory layers. In addition, TSVs are required for the wafer thinning process. The TSV diameter is 1.2 μm and surrounded by a 100 nm SiO_2 side wall passivation layer [317]. Each TSV is capped with backmetal on one end and metal 1 on the other end. As previously discussed, each TSV must align with the location of the regular pattern created by the back metal bond points, as the backside metal (BSM shown in Fig. 8.1b) is used as a landing pad for the TSVs.

Since both WTOP and WBOTTOM are fabricated on a single wafer, both logic layers incorporate a minimum density of TSVs. Design rules specify a maximum spacing of

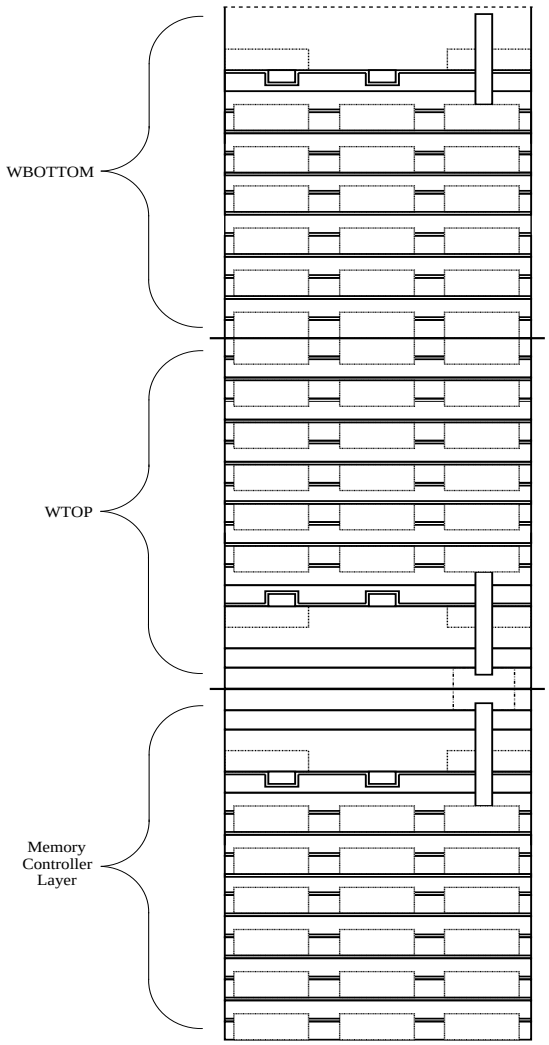


Figure 8.2: Stacked logic die and memory control layers.

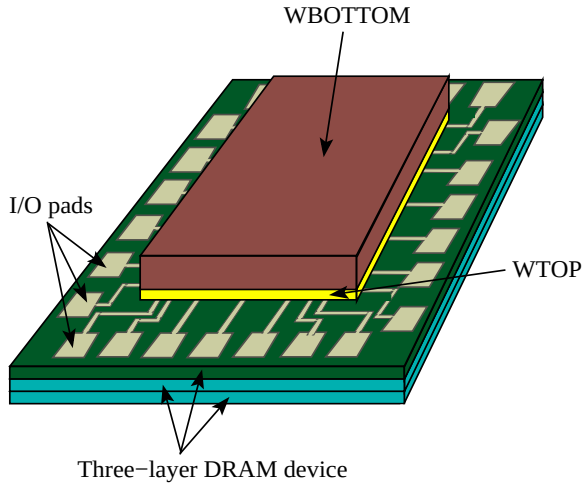


Figure 8.3: Complete stack including the two logic layers, memory controller layer, and two DRAM layers.

250 μm between TSVs [317]. WBOTTOM does not connect to another device plane, but satisfies the minimum TSV density requirements, implying that all TSVs in this device plane are ”dummy” TSVs. In addition, many TSVs in the WTOP logic die are ”dummy” TSVs to satisfy the minimum TSV density.

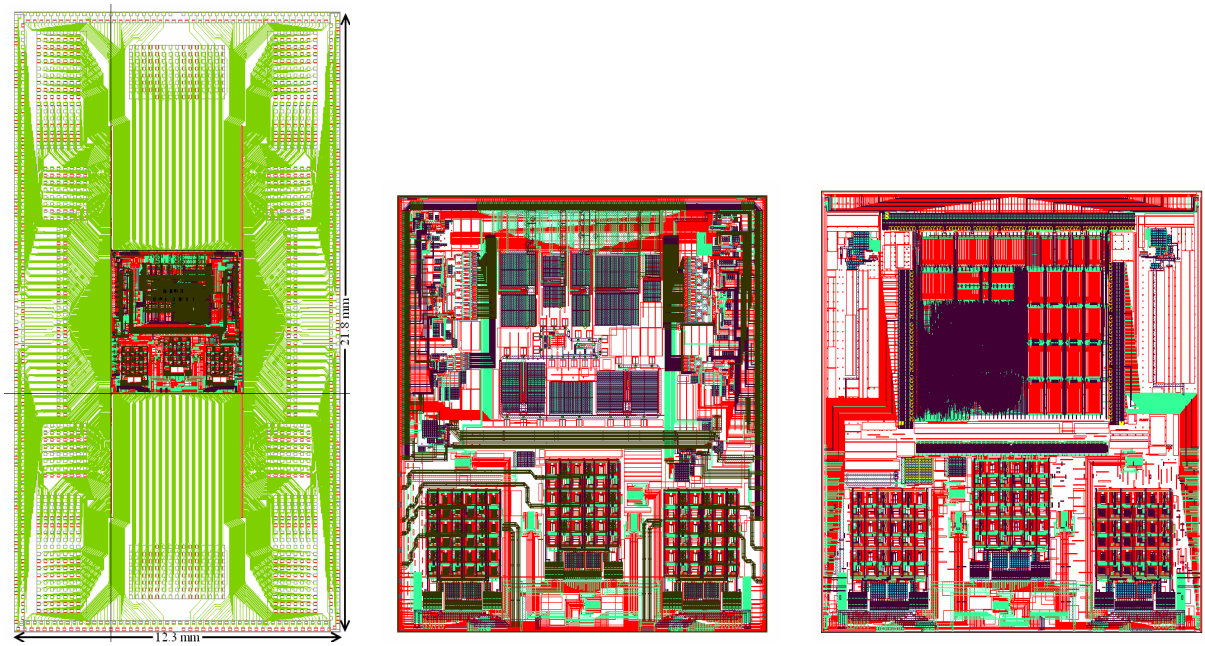
8.2 Overview of test circuit

The fabricated 3-D test circuit contains structures that examine a wide array of design issues in 3-D integrated systems. Some test circuits are designed to examine cross-plane thermal coupling to provide insight into thermal aware floorplanning. Other test circuits are designed as part of a 3-D integrated free-space optical interconnect system (see Chapter 9) as a demonstration that heterogeneous systems integration is feasible with 3-D IC technology. The focus of this chapter is on decoupling capacitor placement test circuits for minimizing noise within 3-D IC power distribution networks.

The fabricated 3-D integrated circuit contains five device planes stacked vertically as previously discussed. Three planes form the memory stack while two planes form the logic planes. The memory stack is designed by Tezzaron, and is accessible through eight pre-defined ports in the upper and lower central portions of WTOP in an area of about 2 mm by 0.5 mm. The DRAM memory stack is not utilized in the current system. Rather, the Artisan SRAM compiler is used to produce a set of caches for the microprocessor. The memory stack footprint is 21.8 mm by 12.3 mm. Bonded on top of this stack are the two 5 mm by 5 mm logic dies. A layout of the stacked memory and logic dies as well as WTOP and WBOTTOM is shown in Fig. 8.4.

The memory stack also behaves as the interface between the logic planes and external inputs and outputs (I/O). There are 800 I/O pads located on the backside metal of the memory controller layer. These 800 I/O pads are routed to the periphery of the logic dies, 200 I/O pads on each side. The length of the electrical paths from the edge of the logic layers to the various I/O pads located on the backside metal of the memory controller layer range from 2 mm to as long as 11 mm.

Tezzaron Semiconductor provided an additional option that does not include the memory stack, *i.e.*, the two logic dies are provided with a re-designed backside metal pattern. The copper backside metal is replaced with aluminum when the memory stack is removed from the five device layer 3-D IC. A microphotograph of the fabricated logic layers with the re-designed backside metal pattern is shown in Fig. 8.5. A description of the test circuit designed to examine the impact of decoupling capacitor placement on noise propagation in 3-D power distribution topologies is provided in the next section.



(a) (b) (c)

Figure 8.4: Layout of the (a) complete logic and memory stack, (b) WTOP logic layer, and (c) WBOTTOM logic layer.

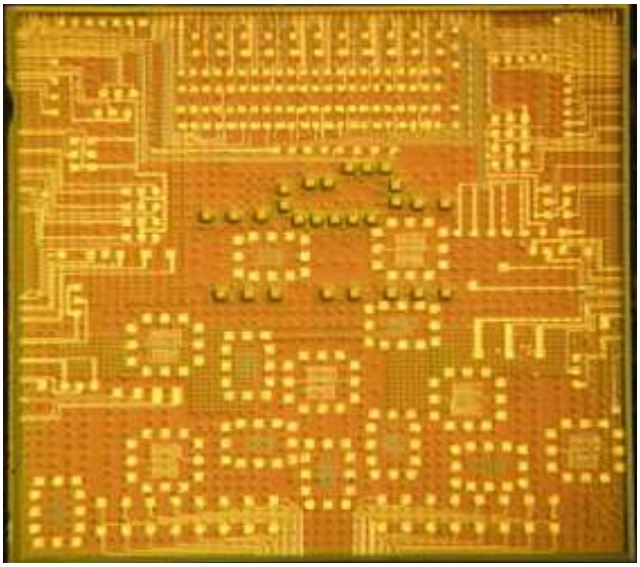


Figure 8.5: Microphotograph of the fabricated logic layers with the aluminum backside metal pattern (no memory layers).

8.3 3-D power delivery with decoupling capacitors

The test circuit is designed to evaluate the effect of both on-chip and board level placement of decoupling capacitors on *IR* and *L·di/dt* noise in 3-D circuits. A schematic diagram of the four topologies used to analyze decoupling capacitor placement is depicted in Fig. 8.6. The first two topologies do not include an on-chip decoupling capacitor. The structure pictured in Fig. 8.6b includes a board level capacitor. Test structures with on-chip capacitors are shown in Figs. 8.6c and 8.6d. The larger capacitors shared between the two planes are depicted in Fig. 8.6c, and a more standard 2-D approach to decoupling capacitor placement where localized point-of-load capacitors are distributed within each plane is shown in Fig. 8.6d. A breakdown of the various components comprising the test circuit evaluating decoupling capacitor placement in 3-D circuits is provided below.

The layout of the three different decoupling capacitor placement topologies is shown in Fig. 8.7. The topology of the power distribution network is identical, but differ with respect to the placement of the decoupling capacitors. The layout depicted in Fig. 8.7a includes distributed decoupling capacitors across both logic layers. The decoupling capacitors are depicted as black rectangles and are distributed throughout the figure. The test block depicted in the second figure, Fig. 8.7b, includes no on-chip decoupling capacitance, as shown by the white space in each sub-block between interdigitated power lines. A decoupling capacitor is shown in Fig. 8.7c, which is placed only on WBOTTOM and is shared between the two logic planes. The distributed capacitance shown in Fig. 8.7a and the shared capacitance shown in Fig. 8.7c are the same. The total on-chip capacitance for both of these test blocks is 0.35 nF, and is produced with a metal-insulator-metal (MIM)

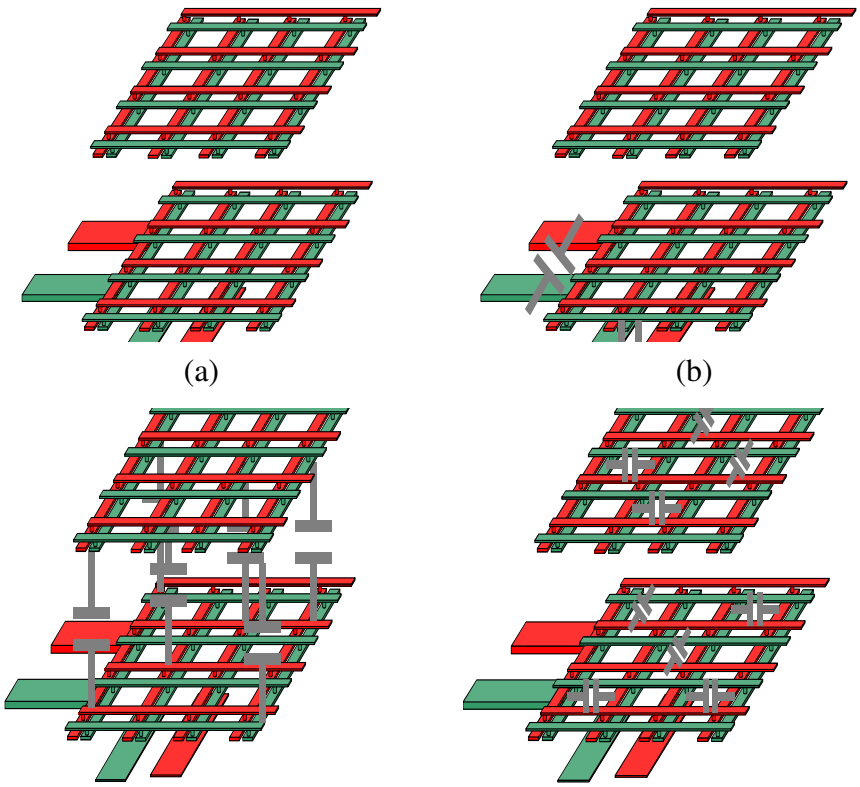
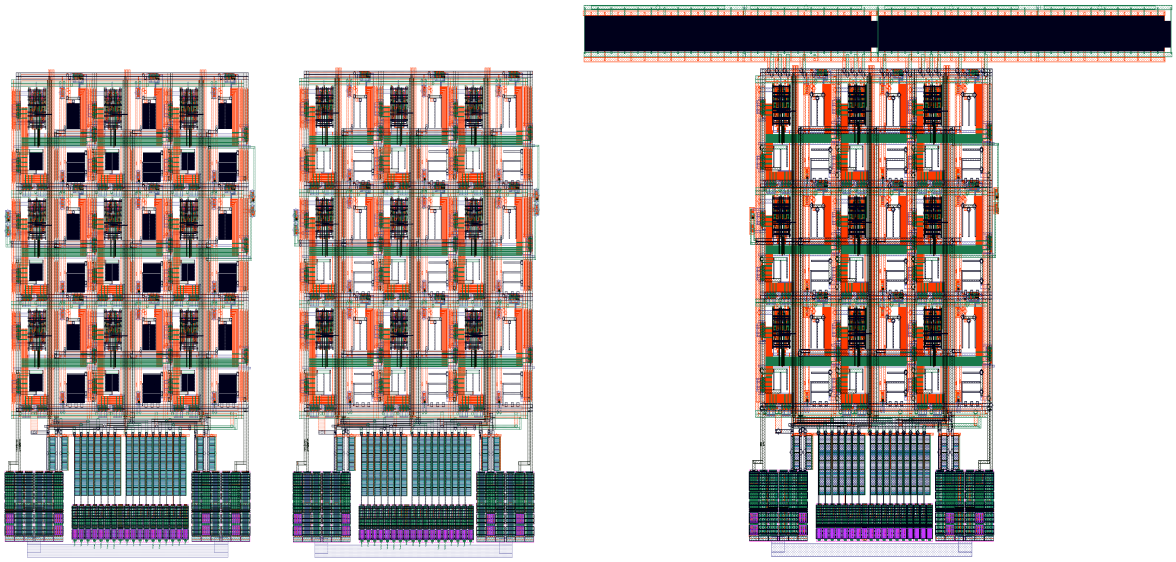


Figure 8.6: Schematic diagram of test circuit examining decoupling capacitor placement for a) no decoupling capacitors, b) board level decoupling capacitors, c) inter-plane shared capacitors, and d) point-of-load decoupling capacitors.

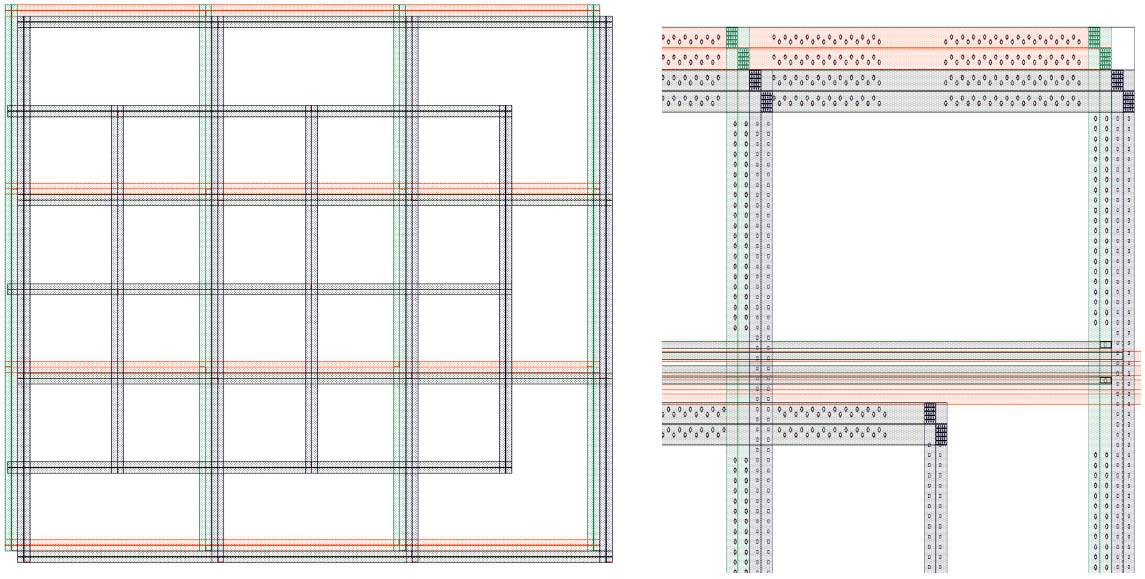


(a) (b) (c)

Figure 8.7: Layout of the three test circuit topologies to examine decoupling capacitor placement; (a) distributed decoupling capacitors, (b) no decoupling capacitors, and (c) shared decoupling capacitors. The black features are the capacitors. The topology shown in (a) only depicts a single logic layer, with the sum of the capacitance amongst the two logic layers being equivalent to the shared capacitance shown in (c).

topology using metal 1 through metal 5. The capacitance for a MIM with dimensions of 25 μm by 27 μm is 0.932 pF. Also note that all three blocks shown in Fig. 8.7 include buffering and ESD protection for each block input (the bottom portion of each figure). Buffering and ESD protection is required for all inputs other than the power and ground distribution network under test. Signal inputs include voltage controls for the ring oscillators, analog power supply for the sense circuits, isolated digital power supply to reduce noise for all other transistors (separate from power/ground network under test), reset signals, and bias voltages for the current mirrors.

Each of the three previously described test circuits includes three isolated power distribution networks. The three power networks are depicted in Fig. 8.8. The two larger

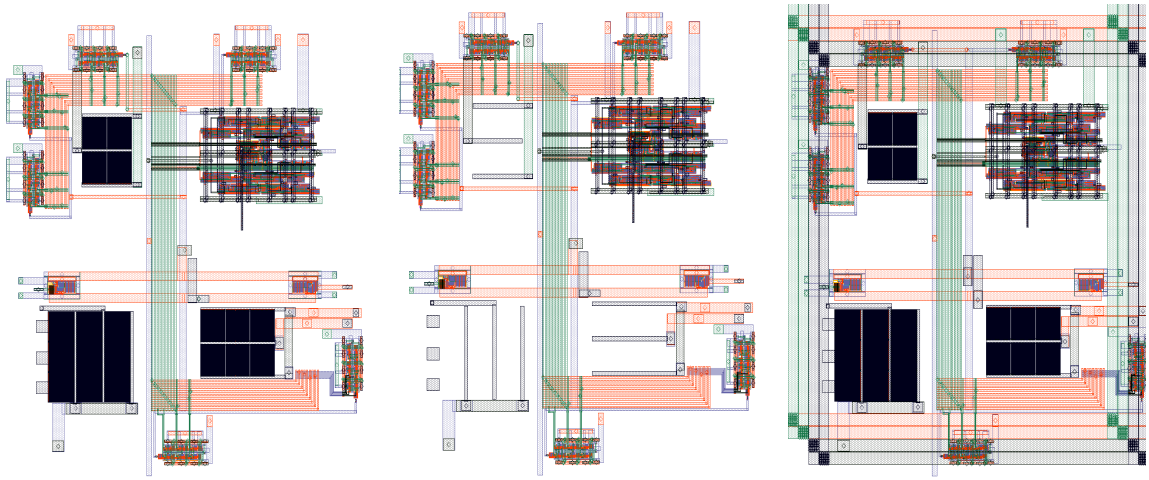


(a) (b)

Figure 8.8: Three independent power distribution networks implemented in each of the three test circuits examine the effects of on-chip and off-chip decoupling capacitances on noise propagation in 3-D integrated circuits. (a) The three power networks for each circuit block, and (b) an enlarged image of the upper right corner of the power networks revealing the TSVs connecting the power networks of the two logic planes.

interdigitated power networks with dimensions of 1025 μm by 1025 μm bound each test circuit. One of the two power networks provides a clean, non-noisy power source to all of the transistors, while the second network is the power network under test. The smaller internal interdigitated power network ensures an isolated power source for the sensitive analog source-follower based sense circuits. All three power distribution networks are composed of 10 μm wide interdigitated lines. Both power networks not under test include a 4 μF board level decoupling capacitance.

The design of the three test blocks with different decoupling capacitor topologies is further simplified by the use of a ”base” block that is repeated for all three topologies and both logic device planes. The larger interdigitated power networks include four sets of



(a) (b) (c)

Figure 8.9: Layout of the (a) noise generation circuit with decoupling capacitors, (b) noise generation circuit without decoupling capacitors, and (c) noise generation circuits within a base block of the power distribution network.

vertical and horizontal power and ground lines. These sets of lines sub-divide the active circuit area of each test block into nine 330 μm by 330 μm areas. Each of these nine smaller areas contain the same identical ”base” block of circuits. A layout representation of the base block with and without localized capacitors is shown, respectively, in Figs. 8.9a and 8.9b. A single base block with decoupling capacitors placed within one of the nine sub-divided areas of the larger interdigitated power network is depicted in Fig. 8.9c.

A schematic of the on-chip circuitry of each base block in each sub-divided area is shown in Fig. 8.10. The circuit is designed to emulate a power load drawing current, generating noise within the power and ground networks. Two sets of circuits are present, noise generating circuits and noise detection circuits. The noise generating circuits include six current mirrors, a voltage controlled ring oscillator with a maximum frequency of 2 GHz, divide by two, four, and eight frequency dividers, and five, ten, and 15- bit pseudorandom number generators (PRNG). Each row of the test block (three sub-blocks of the power

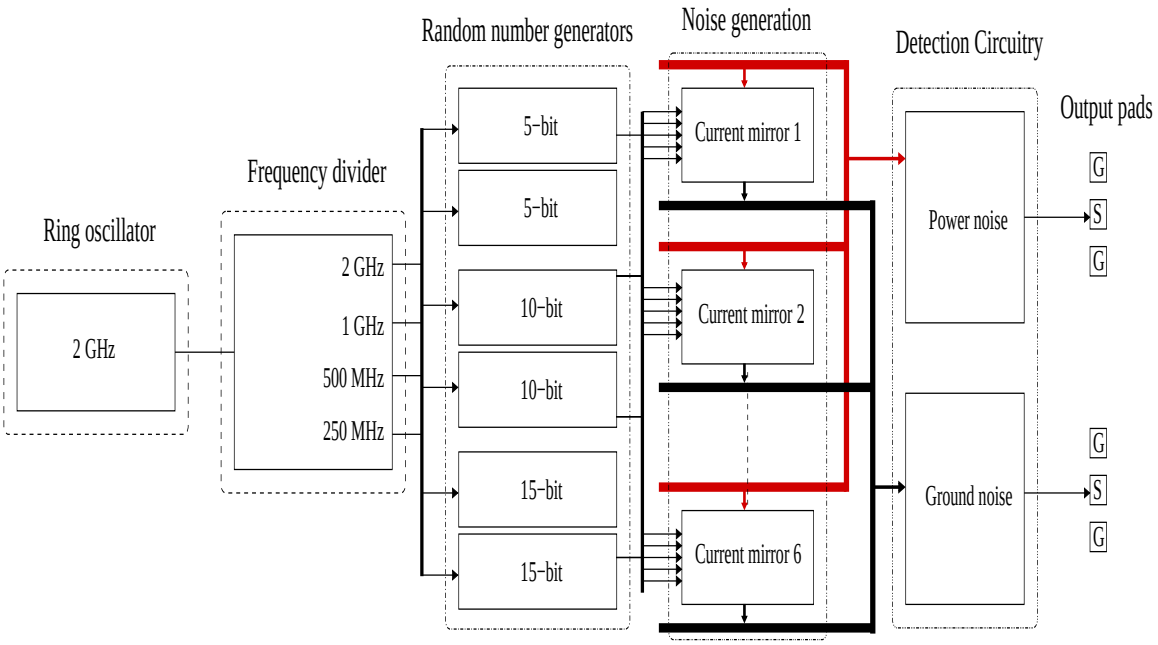


Figure 8.10: Block level schematic of noise generation and noise detection circuits.

network) includes independent voltage controls for the ring oscillator, permitting three different ring oscillator frequencies per device plane per topology. The clock pulses generated by the ring oscillators are buffered and drive the frequency divider, which are again buffered before distributed to the PRNGs. The output of the six PRNGs and the 2 GHz, 1 GHz, and 500 MHz clocks are buffered and randomly drive the current mirrors as enable bits. Each current mirror requires six enable bits, with each enable bit turning on one branch of the current mirror which draws a maximum 3 mA current at increments of 0.5 mA (0.5 mA per branch).

The layout of the noise generation circuits is shown in Fig. 8.11. The layout of the ring oscillator is shown in Fig. 8.12, and the layout of the current mirror and current mirror switch is shown in Fig. 8.13. The current mirror includes a 200 Ω polysilicon resistor (the segment of polysilicon is 1.25 μm long by 1 μm wide) shown in the far left portion of

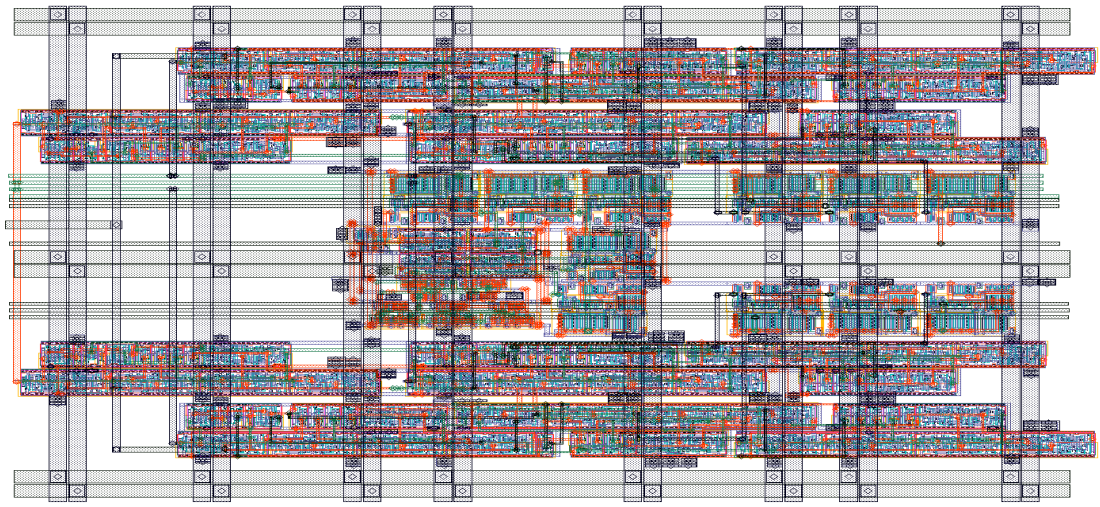
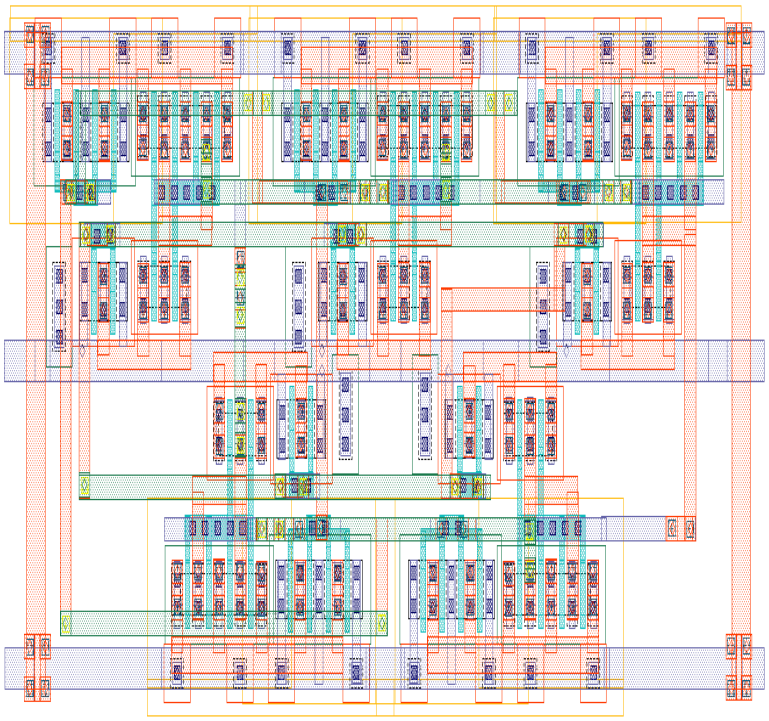


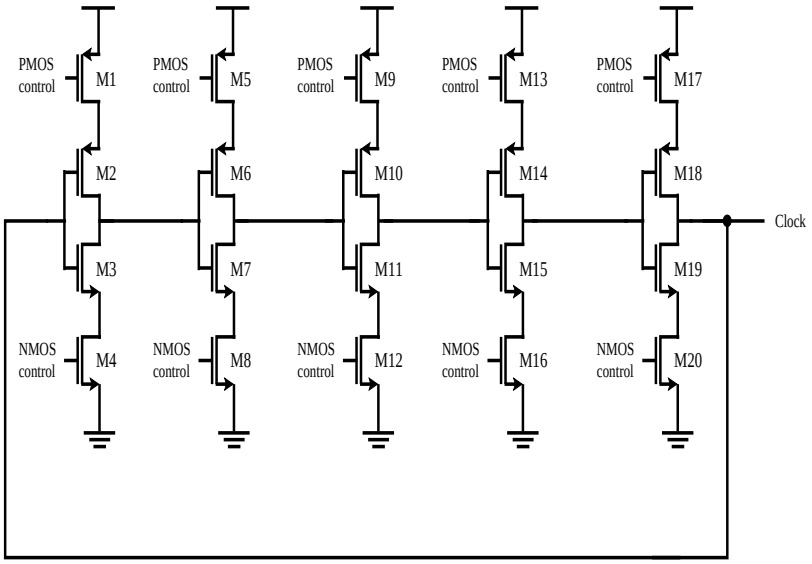
Figure 8.11: Layout of the ring oscillator, frequency divider, six PRNGs (two 5-bit, two 10-bit, and two 15-bit), and buffers for the noise generation circuits.

Fig. 8.13a. This resistor limits the current flow through the current mirror while providing a voltage bias for the circuit.

The noise detection circuits include separate circuits for evaluating power and ground noise. A schematic of the power and ground detection circuits is shown in Fig. 8.14. These single stage amplifier circuits utilize an array-based noise detection circuit, as suggested in [311]. Each 1/9 segment of the larger interdigitated power network contains both power and ground noise detection circuits. An off-chip resistor, labeled R_{bias} in Fig. 8.14, biases the output node of the PMOS-based current mirror, and ensures that the current mirror remains in saturation. The resistor is chosen to ensure that the maximum voltage at the output node of the current mirror does not exceed 400 mV, the voltage at which the current mirror enters the triode region. The layout of both the power and ground detection circuits is shown in Fig. 8.15. Both detection circuits include guard rings surrounding the active devices to ensure isolation from substrate noise.

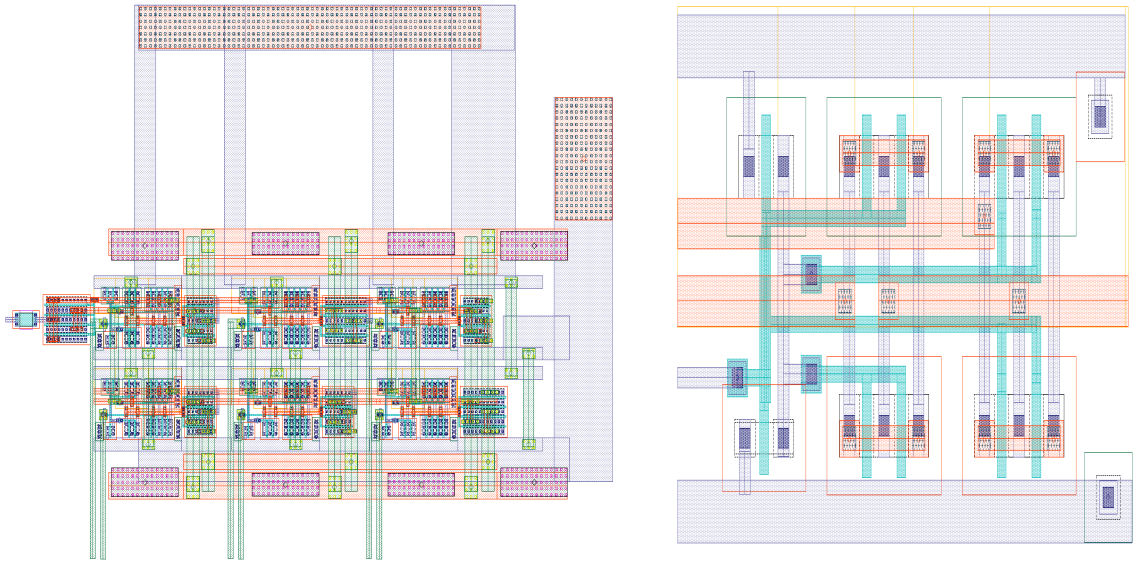


(a)



(b)

Figure 8.12: The voltage controlled, five-stage, two gigahertz ring oscillator used to produce the internal clock within each subsection of the power distribution network; (a) layout, and (b) schematic.

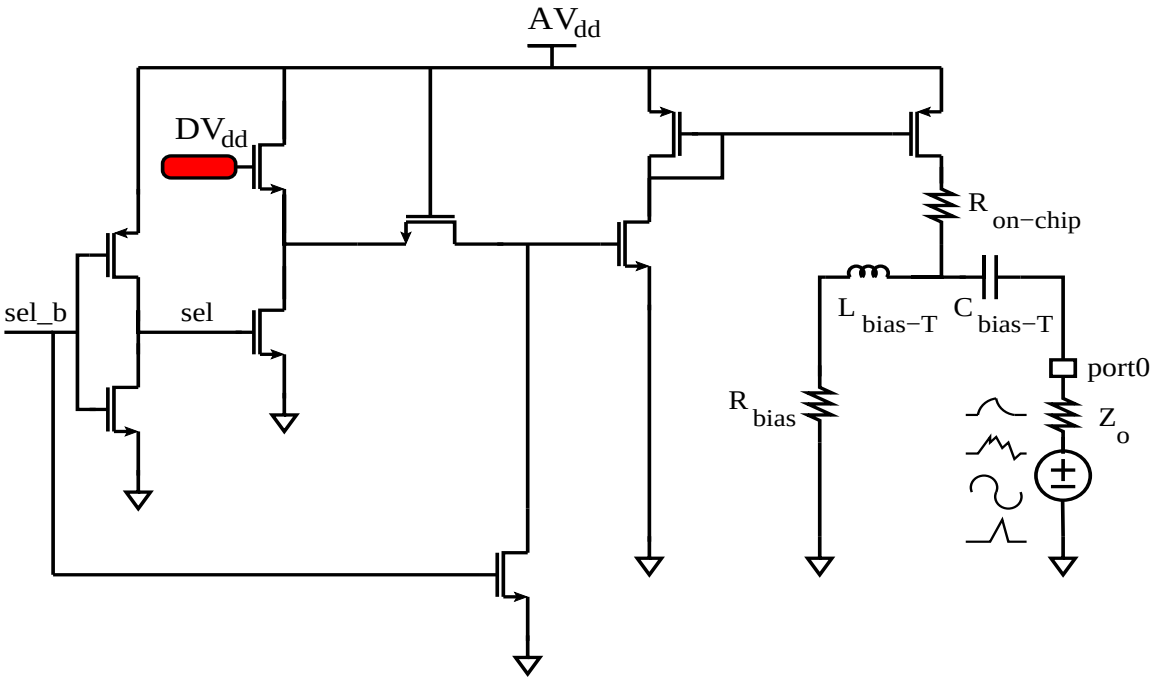


(a) (b)

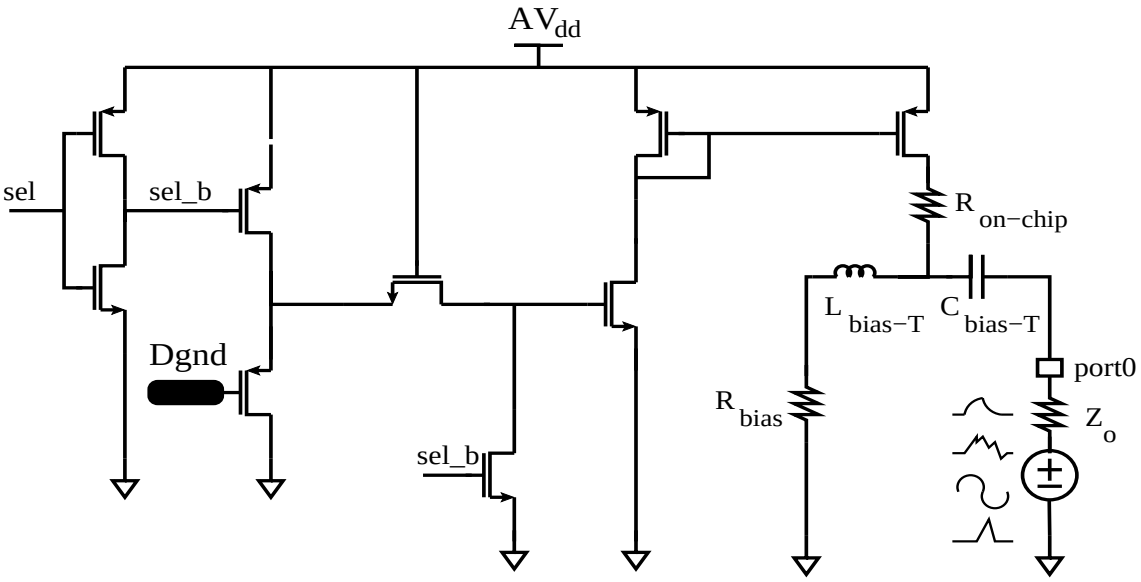
Figure 8.13: Layout of (a) the current mirror with six branches each drawing 0.5 mA of current from the power network, and (b) the electrical switch turning different branches of the current mirror on and off (the inputs to the switches are from the PRNGs).

8.4 Summary

The primary focus of this chapter is on the design of a 3-D test circuit that explores the placement of decoupling capacitors to minimize noise propagation in 3-D power distribution networks. The decoupling capacitor placement circuit requires the I/O pads located on the memory controller layer which is currently being fabricated by Tezzaron.

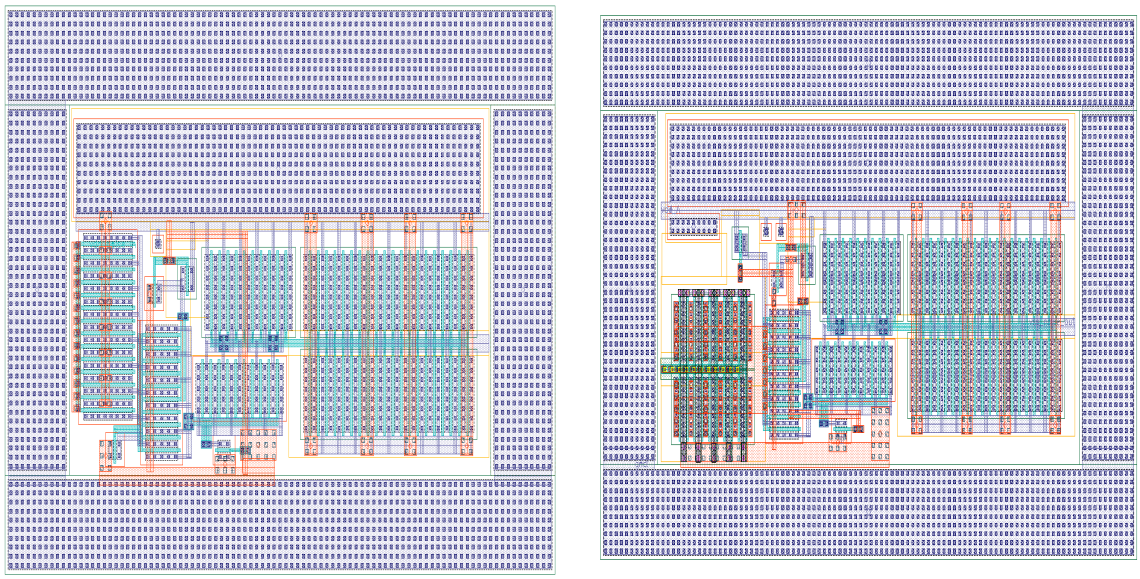


(a)



(b)

Figure 8.14: Source-follower noise detection circuits to detect noise on both the (a) digital power lines, and (b) ground lines.



(a) (b)

Figure 8.15: Layout of the (a) detection circuit to analyze noise on the power network, and (b) detection circuit to analyze noise on the ground network. Both layouts include a guard ring to isolate the detection circuits from substrate noise.

Chapter 9

Heterogeneous 3-D Integration of CMOS and Optical Devices

The research material included in this chapter is an overview of a large multi-group project exploring the integration of CMOS devices with free space optics. Five Professors (E. G. Friedman, H. Wu, M. Huang, G. Wicks, and D. Moore) and a team of twelve Ph.D. students (J. Xue, A. Garg, R. Parihar, B. Ciftcioglu, J. Hu, S. Wang, I. Savidis, M. Jain, R. Berman, P. Liu, J. Zhang, Z. Darling) worked on this project.

The author contributed high level knowledge and design expertise on the 3-D integration of the CMOS and optics device planes. He also contributed results on the thermal effect of VCSELs and the 3-D stacking of dies. No detailed design work or system specifications were completed by the author. Research on the optical transmitter and receiver circuits was performed by J. Hu. The receiver and transmitter circuits for the microprocessor were designed by R. Parihar. The synthesized microprocessor was designed by J. Xue. Design and fabrication of the VCSELs were performed by M. Jain, microlenses by R. Berman, and photodetectors by B. Ciftcioglu. General specifications were set by the group, while detailed circuit parameters were determined by each respective group member.

9.1 Introduction to heterogeneous 3-D integration

One substantial benefit of 3-D integration is the potential to integrate a variety of technologies without compromising yield. Disparate technologies such as analog, III-IV semiconductors, RF, memory, and optics, as well as emerging technologies such as nanowire and graphene can be integrated to form a 3-D integrated system. Separate processing of each device plane prior to bonding and TSV formation allows for wafer and die level testing to assure proper functionality of the devices. Merging disparate technologies with CMOS, however, often requires different currents, synchronization methodologies, and voltages among the device planes. Systems level integration with CMOS is therefore an on-going research objective [48, 76, 79, 325].

One technology that can potentially lead to improved circuit performance, reduced power consumption, and increased functionality is optics based interconnects. Integrating optical interconnects with traditional CMOS technology can reduce delay and increase bandwidth as compared to metal interconnects [326–328]. Integrating photonic devices with silicon has progressed significantly in recent years [329, 330]. However, much of this work has focused on optical interconnect using planar optical waveguides, which are integrated onto the same die with the CMOS devices. A limitation of optical waveguide interconnects is the lack of optical switches and storage devices. Without these switches and storage devices, routing and flow control in a packet-switched network requires repeated optoelectronic (O/E) and electro-optic (E/O) conversions, which significantly increases signal delay, energy consumption, and circuit complexity. In addition, efficient silicon electro-optic modulators are a challenge to manufacture due to the inherently poor

optical properties of silicon and other weaker physical properties such as the plasma dispersion effect [331]. The modulators, therefore, require a long optical length, resulting in large device sizes [332]. Techniques to reduce the size of the modulator devices, such as high-Q resonators (ring, micro-cavity, or micro-disk), effectively slow the light [333] at a cost of increased delay and reduced bandwidth.

An additional drawback of in-plane waveguides is the mode diameter, which determines the minimum distance between optical waveguides to avoid crosstalk. Existing distances are significantly larger than the electrical interconnect pitch at sub-micrometer technology nodes and is expected to become smaller as CMOS technologies scale [334]. Although wavelength division multiplexing (WDM) has proven effective in long distance fiber-optic communication systems, it is not practical for intra-chip optical interconnects due to the significant power and area overhead required for wavelength multiplexers and demultiplexers, and the requirement for multiple external laser sources operating at different wavelengths.

One solution is a heterogeneous 3-D system that includes an optical device plane integrated with a CMOS plane. The optical device plane operates as a free space interconnect system, thereby eliminating the switches and storage devices required with optical waveguides. Free space optics has been applied to both board-to-board [335, 336] and inter-chip communications [337, 338]. The added benefit of a 3-D IC hybrid system is that the optical devices can be fabricated on materials traditionally used for optics such as GaAs and SiGe, removing the need for the optics to be integrated in silicon. The lasers, modulators, and photodetectors are available in III-V semiconductors and can be assembled on a CMOS SoC using 3-D integration technology.

Technologies including emerging devices, CMOS circuits, and optical devices can be leveraged to build an interconnect architecture without network router nodes. The primary technologies used to develop a FSOI are:

- Signaling: Vertical cavity surface emitting lasers (VCSEL) provide light emission without the need for external laser sources and to route the optical power supply throughout the IC. VCSELs, photodetectors (PDs), and the supporting micro-optic components are implemented in GaAs, Ge, or silica technologies and integrated into a 3-D silicon system, which includes CMOS digital electronics as well as the transmitters and receivers for the optical communication network.

- Propagation medium: Free space optics using integrated micro-optic components provide high speed signal propagation with low loss and low dispersion.

- Networking: Direct communication links through dedicated VCSELs, PDs, and micromirrors (in small-scale systems) or via phase array beam-steering (in large-scale systems) allows a quasi-crossbar structure that avoids packet switching, offers ultra-low communication latency, and provides scalable bandwidth due to the fully distributed nature of the optical interconnect.

The 3-D integrated free space optical interconnect system is further described in this chapter. An overview of the system is provided in Section 9.2. A description of the optical device plane is provided in Section 9.3. The transceiver circuits necessary to interface between the electrical and optical communication network and the microprocessor core designed to verify core-to-core transmission are described in Section 9.4. Thermal effects

due to the VCSELs and the stacking of device planes are discussed in Section 9.5. Some concluding remarks are provided in Section 9.6.

9.2 Overview of 3-D integrated FSOI

An overview of the intra-chip optical interconnect system for multi-core processors incorporating free space optics with 3-D integrated photonic and CMOS devices is provided in this section. The objective is to provide point-to-point links constructing an all-to-all communication network with high bandwidth density, low latency, and high energy efficiency without the need for routing or switching optical data packets. The 3-D integrated free space optical interconnect (FSOI) system is illustrated in Figure 9.1. Three distinct layers of devices are integrated to produce a 3-D integrated FSOI. The three components comprise 1) a CMOS device plane consisting of the transceiver and microprocessor circuits, 2) a photonics layer consisting of vertical-cavity surface emitting lasers (VCSEL) and photodetectors (PD), and 3) a free space optical guiding medium constructed from micromirrors and microlenses.

The free space optics enables an N-to-N direct communication structure, where each core within a multi-chip module (MCM) has a dedicated VCSEL and PD. The optical signal generated by an electrically modulated VCSEL in one core is focused by a microlens on the backside of the GaAs substrate. Electrical modulation is performed by transmitter circuits in the CMOS device plane that are delivered to the VCSELs through the substrate with through silicon vias. The optical signal is guided across the IC by mirrors located on the IC and package (the top device layer shown in Figure 9.1). Once the signal is propagated to

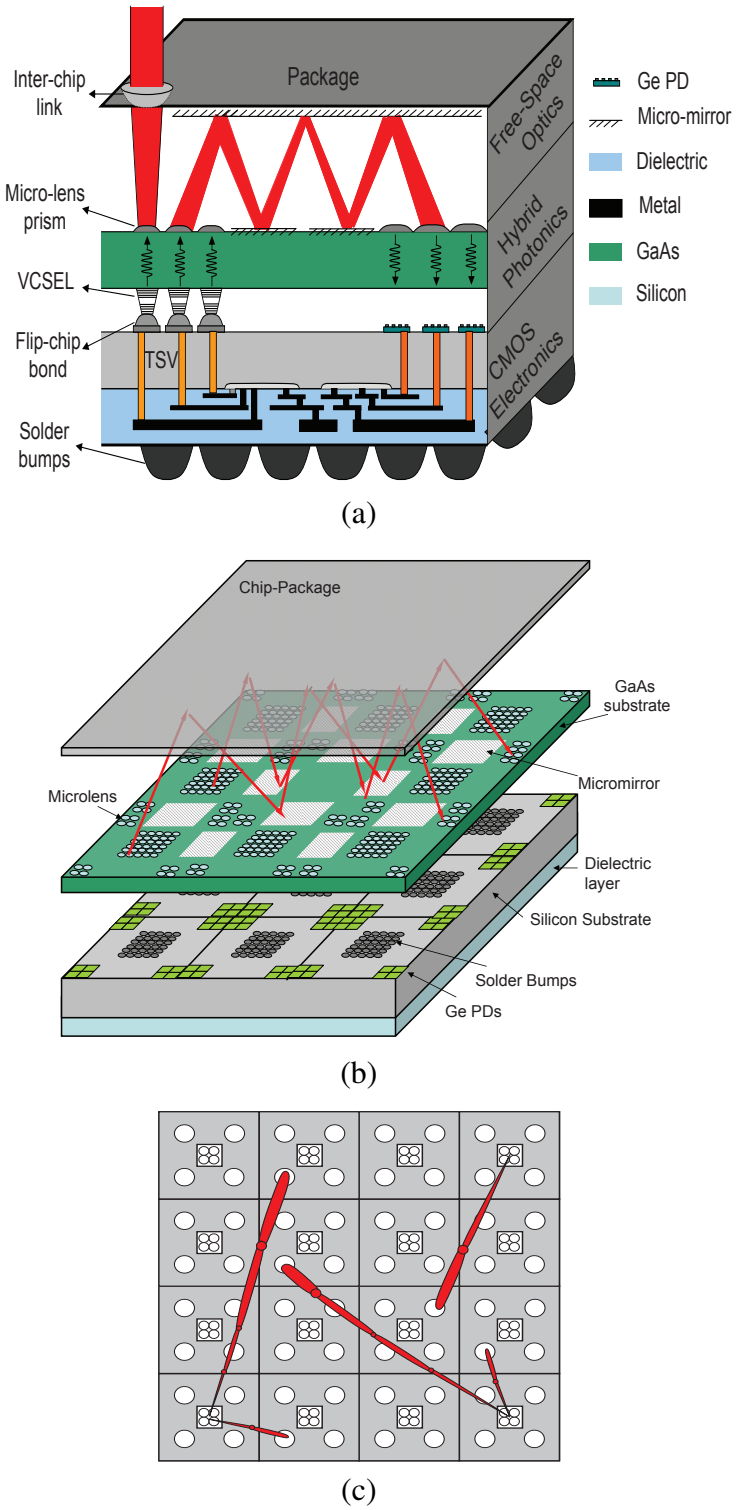


Figure 9.1: Schematic representation of the 3-D integrated FSOI including (a) a cross-sectional view of the FSOI with the intra- and inter-core communication links, (b) the FSOI system with VCSEL arrays (center) and photodetectors (periphery) to signal a multi-core microprocessor, and (c) a top view of core-to-core optical transmission [339–341].

the target core, a microlens focuses the signal onto a photodetector, where the optical signal is converted back into an electrical signal. The electrical signal is filtered and amplified by the receiver circuitry at the destination core. Note that, unlike waveguides, the electrical-to-optical and optical-to-electrical conversion only occurs once from the originating core to the destination core, reducing both delay and power consumption.

The free space optical interconnect exhibits several distinct advantages over other optical interconnect techniques. These advantages include:

- Low latency, as no packets are switched (therefore the associated intermediate routing, buffering, and arbitration delays are avoided).
- Low propagation loss, minimal dispersion, and low bandwidth degradation with transmission distance.
- Good signal integrity simplifies the CMOS transceiver electronics (single laser driver in the transmitter and single amplifier in the receiver).
- Reduced power consumption by eliminating packet-switching, powering down the VCSELs during low duty cycle operations, and avoiding thermal tuning of sensitive electro-optical modulators in WDM systems.
- Easily adjusted to included inter-chip communication pathways (as shown in Figure 9.1(a)).

A description of each of the optical components comprising the optical device planes is provided in the following section.

9.3 Components of the optical system

The primary optical components of the 3-D integrated free space optical interconnect system are the vertical cavity surface emitting lasers (VCSEL), the photodetectors (PD), and the micromirrors/microlenses. The optical components are designed for an operating wavelength of 850 nm. Each component of the 3-D integrated FSOI system was fabricated at the University of Rochester. The three components are described below.

9.3.1 Vertical cavity surface emitting laser (VCSEL)

The vertical cavity surface emitting laser is a fundamental technology that integrates the optical communication network with the CMOS device planes. The VCSEL converts electrical signals to light which is transmitted through the free space optical system to the destination core. The VCSELs described in this subsection operate at a 980 nm wavelength, whereas the rest of the components are designed to operate at 850 nm to match the commercial VCSELs used in the prototype circuit. Once in-house VCSELs are produced, the commercial VCSELs will be replaced and the PDs and other optical components will be adjusted to 980 nm.

The VCSEL includes an intra-cavity contact, an oxide aperture, and AlAs/GaAs distributed Bragg reflectors (DBR) mirrors. A schematic view of the VCSEL structure is shown in Figure 9.2, which also includes the substrate. The intra-cavity resonator is an arrangement of mirrors (AlAs/GaAs DBR) forming a standing wave cavity resonator by surrounding the gain medium that amplifies the light wave. Bragg reflectors consist of an alternating sequence of high and low refractive index layers with a quarter wavelength

thickness [342]. In addition, AlAs/GaAs DBRs have been shown to form uniform and smooth heterointerfaces, yielding mirrors with high reflectivity [342, 343]. The thick top DBR and bottom DBR outline the active region of the device. More than 20 Bragg pairs are typically required for each mirror [342]. The intra-cavity contact improves the speed characteristics of the VCSEL. The two metal square loops connect above and below the active region of the quantum well through the heavily doped (p++ and n++) GaAs contact layers. The doping concentration of the p++ and n++ contact layers are, respectively, 5×10^{18} cm^{-3} and 3×10^{18} cm^{-3}. Each contact layer is 208.6 nm thick and is located between the DBR mirrors and AlAs layers. This structure has both electrical and optical advantages. The vertical current injection path is reduced by as much as 90% without passing through the thick layer of the AlAs/GaAs mirrors. The two 83.8 nm AlAs lateral oxide layers shape the optical aperture and encapsulate the active region to reduce the threshold current while improving the quantum efficiency of the VCSEL. The active region is composed of three 8 nm $In_{0.2}Ga_{0.8}As$ quantum wells (QWs) and four 10 nm GaAs barrier layers, and is covered by two 116.8 nm thick $Al_{0.5}Ga_{0.5}As$ spacer layers. The total thickness of the core, approximately 300 nm, is about one third of a wavelength. In addition, since no current flows through the DBR mirrors during laser operation, the mirrors are undoped, which minimizes optical absorption and carrier scattering inside the mirror region. Note that the total thickness of the active region, aperture layers, and contact layers is one wavelength, about 980 nm [340, 341, 344].

A top view of the VCSEL is shown in Figure 9.2(b). The VCSEL is typically designed with a circular or square aperture. In this case, anisotropic oxidization of AlAs produces a

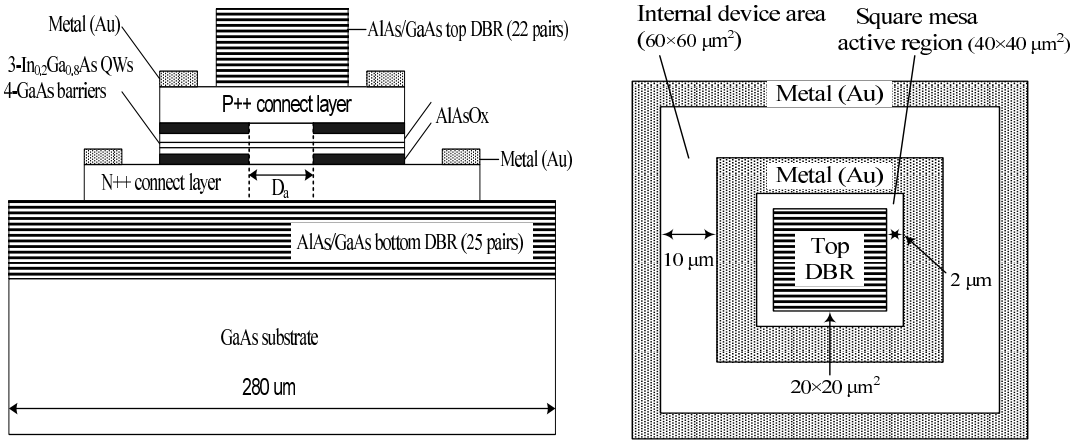


Figure 9.2: Schematic diagram of VCSEL structure (a) side view, and (b) top view.

non-circular aperture despite a circular active region mesa [345]. A square mesa is therefore fabricated rather than aligned with the sides along the <100> crystal lattice axes, which exhibits a higher oxidation rate than the <110> direction. The top mirror, active region (square mesa), and internal device area (without the N++ metal contact) are, respectively, 20×20 μm^2, 40×40 μm^2, and 60×60 μm^2. The minimum pitch size for the VCSEL array is equivalent to three times the width of the active area, and is 120 μm. Also shown in Figure 10.2(b) is the 2 μm space between the metal to the P++ contact and the top mirror sidewall to avoid metal contamination of the top mirror, and the 10 μm space between the metal contacts to facilitate the lift-off process. The lift-off process releases the substrate from the VCSEL structure that includes the epitaxial DBRs [346].

9.3.2 Photodetectors

The photodetectors (PD) receive optical signals at the receiver of the destination core. Light is absorbed and converted to an electrical signal by the PD. The photodetectors were

designed to support a 7.5 Gbps data rate per link at 850 nm. The semiconductor material used to produce the PDs is germanium. Germanium based PDs are easier to integrate with CMOS, and exhibit strong optical and electrical properties that increase sensitivity and bandwidth. A metal-semiconductor-metal (MSM) structure is chosen over a p-i-n structure to reduce the parasitic capacitance per unit area. With a reduced parasitic capacitance, the strict alignment requirements of the microlens with the PDs to efficiently couple the light can be loosened without degrading the bandwidth or responsivity [341, 347].

The physical and electrical properties of the germanium MSM PD were characterized after fabrication. Each PD is 62 by 62 μm^2 and includes a parasitic capacitance of 83 fF. The PD also exhibit a 0.315 A/W responsivity and a 7 μA dark current. The experimentally measured dark current is larger than the simulated. The higher experimental value is primarily due to accidental plasma damage at the hydrogenated amorphous silicon (a-Si:H) layer during etching of the Si_3N_4 layer, increasing surface trap states at the a-Si/Si_3N_4 interface [341]. The photodetector bandwidth is 9.3 GHz at a 7 V bias, and is primarily limited by the transit time of the carriers and the impedance of the wirebonds from the IC to the circuit board.

9.3.3 Micromirrors and microlenses

Microlenses and micromirrors are used twice within a single link of the FSOI system, as shown in Figure 9.3. Microlenses are located directly above the VCSELs to focus the output light signal onto the micromirror directly above which reflects the data signal through the free space optical interconnect system. A second micromirror reflects the signal back towards a photodetector on a second core. The reflected signal passes through a second

microlens that again focuses the lightwave onto the photodetector to minimize optical loss.

The microlenses are made from fused silica. Fused silica offers a number of advantages including very low optical transmission loss at 850 nm wavelength, is easily integrated with silicon substrates, and is compatible with CMOS processes. The lenses are built by melting and reflowing photoresist into a spherical shape and then dry etching the pattern into a 525 μm thick silica wafer. The thickness of the lens is 15.3 μm. The lens to lens pitch is 250 μm, matching the pitch of the VCSEL array. The fabricated microlenses for the VCSELs have a 200 μm aperture size with a corresponding focal point in air of 580 μm. Microlenses for the photodetectors have a 220 μm aperture and a 730 μm focal point (in air). All lenses exhibit a 1 dB optical scattering transmission loss. The total fused silica IC area is 0.84 by 0.84 cm^2 [340, 341, 348].

The micromirrors are fabricated directly on silicon or polymer substrates by micro-molding techniques [349]. A greater than 99% reflectivity is achieved as the light reflects off the mirrors.

9.3.4 FSOI prototype test circuit

A prototype circuit has been tested to evaluate the optical device planes that includes the VCSELs, microlenses, and photodetectors. The micromirrors are replaced by a glass prism to simplify the testing process. In addition, commercial VCSELs are used for the prototype as the VCSELs produced by members of the University of Rochester team are not yet available. The commercial VCSELs used are Finisar V850-2092-001S, sold as a 1 by 4 array with a pitch of 250 μm. Each VCSEL provides a 2 mW optical power output at 850 nm with a 5 GHz modulation bandwidth. The commercial VCSELs also had an

aperture size of 8 μm and a full width half-maximum far-field divergence angle of 20° in free space at the operating bias point.

The prototype includes the 3-D integration of the VCSELs, microlenses, and photodetectors, as shown in Figure 9.3. The photodetectors were fabricated on a germanium (Ge) substrate. The VCSELs are integrated with the PDs on the germanium carrier with high horizontal accuracy and minimal tilt and rotational error [341]. A non-conductive epoxy was used to bond the commercial VCSELs to the Ge carrier and wirebonding was used to electrically connect the devices. The alignment tolerance is limited to a few micrometers due to the 0.5 μm optical stage resolution and a maximum ±5 μm axial placement uncertainty of the VCSELs. After bonding, the VCSELs and PDs are 0.75 cm apart. In addition, the VCSELs are wirebonded to 1 mm long 50 Ω transmission lines on the Ge substrate. Each PD is also bonded to a 1 mm long feed line for testing.

The VCSELs and PDs are integrated onto the Ge carrier with the microlenses and micromirrors. A 380 μm spacer layer of silica is bonded to the Ge carrier to assure the vertical space between the bondwires for the VCSELs and PDs that are located between the carrier and microlens. The fused silica microlens die is aligned to the Ge carrier and glued with a non-conductive epoxy. The measured tilt of the microlens die after integration is less than 4 μm from one edge to the other edge [341]. As there is an approximately 200 μm difference in height between the VCSELs and PDs (as shown in Figure 9.3), the microlens aperture and focal length have been adjusted to match the required conditions of each device. As previously mentioned, the microlenses had a 200 μm aperture size and 580 μm focal point for the VCSELs, and a 220 μm aperture size and 730 μm focal point for the PDs.

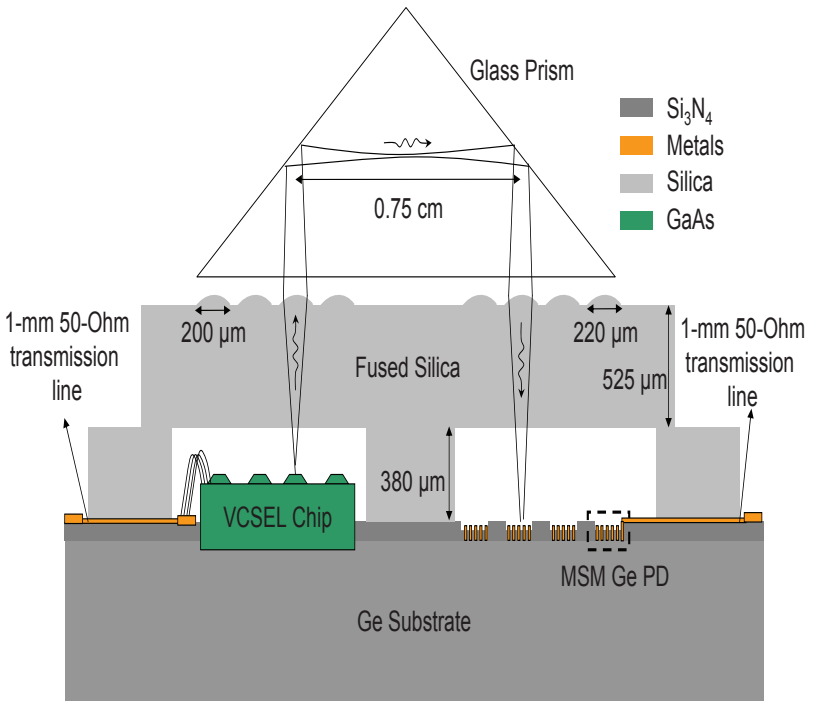


Figure 9.3: Schematic of the test circuit that experimentally verifies the optical components of the 3-D integrated free space optical interconnect system [341].

Experimental results from the prototype indicate a 3 dB bandwidth for the FSOI link of 3.3 GHz with a total power consumption of 5.1 mW. The electrical-to-electrical current gain of the optical link was measured as -24.4 dB at a 1 cm distance and -26.6 dB at 1.4 cm (total path distance from VCSEL to the PD). The crosstalk between adjacent links is -23 dB for a path length of 1 cm and -16 dB for a 2 cm long path [341].

9.4 CMOS based 3-D integrated circuitry

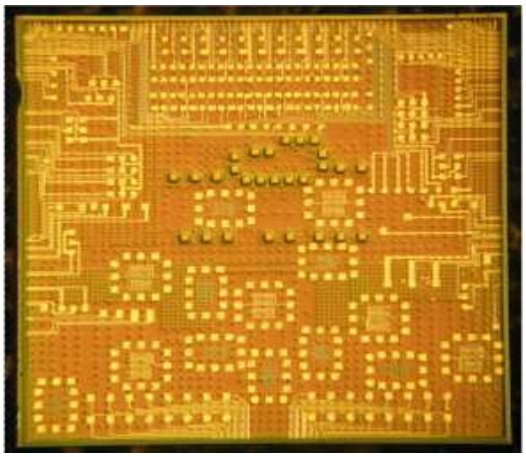
In addition to the optical components of the 3-D integrated free space optical interconnect system, transceiver circuitry and a microprocessor are also included to provide, respectively, signal control (electric-to-optic and optic-to-electric conversion, amplification,

and filtering) and generation of the electrical data for transmission over the FSOI. Both the transceiver and microprocessor are currently in fabrication from Tezzaron Semiconductor. Tezzaron provides two logic device planes that are integrated within a 3-D system with 1.2 μm diameter TSVs to form a stacked die. The 3-D integrated IC also includes three device layers of memory (two 1 GB memory layers, and one memory controller layer). Each logic device plane is separately processed by Chartered Semiconductor (currently a subsidiary of GlobalFoundries) in a 130 nm CMOS process technology prior to 3-D bonding, TSV fabrication, and wafer thinning, which is completed by Tezzaron.

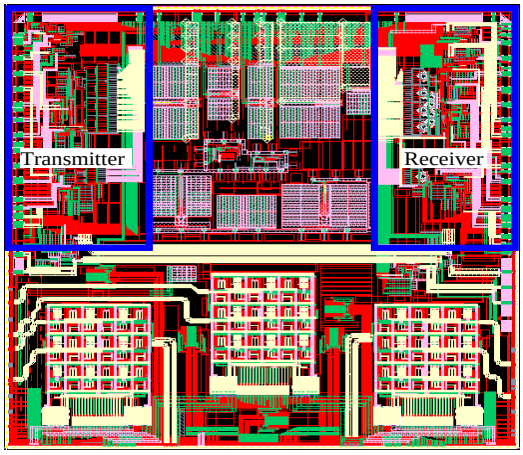
The Chartered fabrication process includes low power 1.5 volt and 2.5 volt transistors, six metal layers per device plane, a single polysilicon layer, dual gates for the 2.5 volt transistors, and low and nominal threshold devices [317, 350]. The sixth metal level on each die is allocated for face-to-face bonding to vertically stack the two logic device planes. A microphotograph of the completed 3-D integrated logic planes is shown in Figure 9.4(a). The location of the transmitter and receiver circuits on the top logic die, and the microprocessor on the bottom logic die are shown, respectively, in Figures 9.4(b) and 9.4(c). A description of the FSOI tranceiver circuitry is provided in the following subsection, while an overview of the microprocessor is provided in Subsection 9.4.2.

9.4.1 FSOI transceiver circuitry

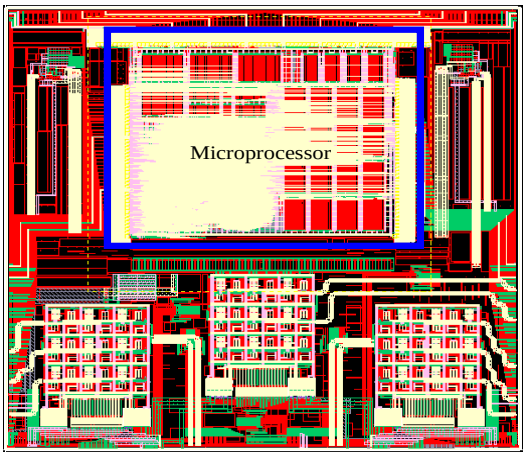
The transceiver circuitry is composed of transmitter circuits that convert an electrical signal from an originating logic core to an optical signal for transmission by the free space optical network, and receiver circuits that convert an optical signal to an electrical signal at a destination core. A schematic representation of the transmitter, FSOI, and receiver



(a)



(b)



(c)

Figure 9.4: A microphotograph of (a) the fabricated 3-D integrated logic dies, (b) the top logic die, and (c) the bottom logic die. The two individual logic dies comprise the 3-D IC.

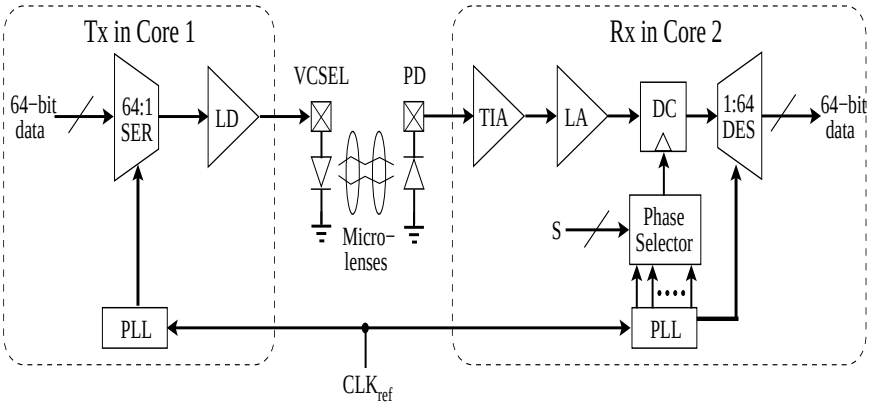


Figure 9.5: Schematic representation of the transceiver circuitry included in the CMOS device plane [353].

circuits is shown in Figure 9.5. A shared-clock transceiver has been designed, where the reference clock is simultaneously shared between the transmitter and receiver [351, 352]. A shared-clock architecture is well suited for systems with multiple transceivers, such as multi-channel or multi-core systems. The clock signal is distributed among the different cores through a global distribution network. A brief description of both the transmitter and receiver circuits follows.

The transmitter circuitry consists of three components: serializer (SER), laser driver (LD), and phase-locked loop (PLL). The SER serializes the n-bit low speed parallel outputs from the microprocessor core and feeds the high speed serialized output to the laser driver. In the FSOI system, a 64 bit to 1 bit SER produces one high speed 10 Gb/s data signal. The PLL provides the reference clock frequency (10 GHz) for the SER. The SER is based on a binary tree architecture consisting of a set of 2-to-1 multiplexer circuits, where each stage shares one clock for control and synchronization. Based on the serialized data from the SER, the laser driver delivers the corresponding modulation current to the VCSEL, which converts the electrical signal to an optical signal transmitted by the FSOI [353].

The receiver circuitry consists of six components: 1) transimpedance amplifier (TIA), 2) limiting amplifier (LA), 3) decision circuit (DC), 4) phase selector, 5) phase-locked loop (PLL), and 6) deserializer (DES). An optical signal is converted into an electrical signal by the photodetector. The TIA converts the output current from the photodetector to a voltage with sufficient gain and bandwidth. The output signal from the TIA is further amplified by the LA to satify the input sensitivity requirement of the data recovery circuit (the decision circuit). As there is substantial noise in the received data, a decision circuit is placed between the limiting amplifier and the deserializer to sample the amplified signal from the limiting amplifier and generate a clean output to the deserializer. The phase selector generates the optimal phase to trigger the decision circuit sampling the data [353]. The PLL is identical to the circuit used in the transmitter circuit, and can be shared among the two. A single PLL is therefore needed per core. In the receiver, the PLL is used for both the deserializer and the multi-phase clocks in the decision circuitry. The DES converts the high speed 10 Gb/s serial data to 64 bit low speed parallel data that is sent to the microprocessor core.

9.4.2 Integrated microprocessor cores

The microprocessor core generates and receives data from the FSOI through the transceiver circuits located on the top CMOS logic plane. The core is a 32 bit RISC (RISC32E) processor modified for the 3-D integrated FSOI system. The core is a high performance, low power, 32 bit RISC core designed for custom system-on-silicon applications [354]. The RISC32E is intended for semiconductor manufacturing companies, ASIC developers,

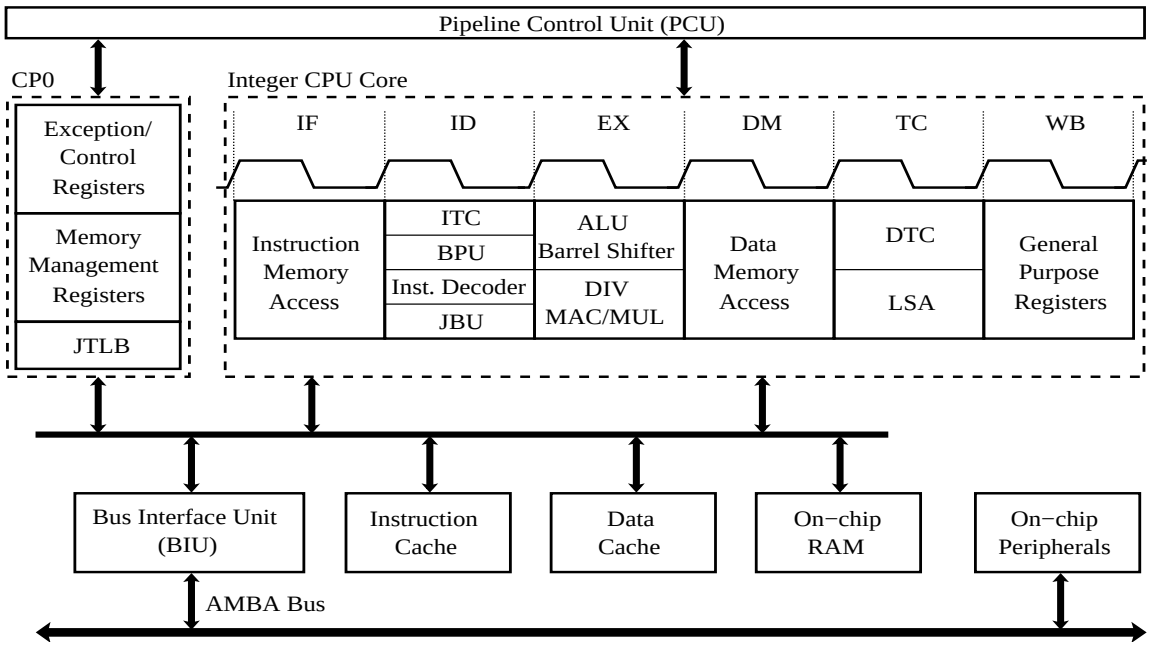


Figure 9.6: Schematic block representation of the RISC microprocessor included in the CMOS device plane.

and system OEMs to rapidly integrate custom logic and peripherals with a high performance RISC processor. It is highly portable across processes, and can be integrated into full system-on-silicon circuits.

The RISC32E core architecture is shown in Figure 9.6. The core utilizes the MIPS32 architecture and contains all MIPS II instructions, including modified multiply-accumulate (MAC), conditional move, prefetch, wait, and leading zero/one detect instructions. The RISC32E also includes the 32 bit privileged resource architecture and the R4000-style memory management unit which contains three-entry instruction and data translation lookaside buffers (TLBs - ITLB/DTLB) and a 16 dual-entry joint TLB (JTLB) with variable page sizes [354].

The RISC32E core utilizes a six stage pipeline, as shown in Figure 9.6. The pipeline allows the processor to achieve a high frequency while minimizing device complexity, which

reduces both cost and power consumption. The six stage pipeline is: 1) Instruction Fetch (IF stage), 2) Instruction Decode (ID stage), 3) Execution (EX stage), 4) Data Memory (DM stage), 5) Tag Compare (TC stage), and 6) Writeback (WB stage). During the IF stage an instruction is fetched from the instruction cache. An instruction is decode based on operands fetched from a register file during the ID stage. The execution stage includes the arithmetic or logical operation for register-to-register instructions by the ALU, as well as multiply and divide operations. During the DM stage, data cache fetch and data virtual-to-physical address translation is performed for the load and store instructions. In addition, data cache look-up is performed and a hit or miss determination is made. If a data miss occurs, a miss request is sent to the data cache during the TC stage. If no data miss occurs, a partial word load and store are aligned. Finally, the instruction result is written back to the register file during the WB stage for register-to-register or load instructions.

The system control coprocessor (CP0 in Figure 9.6) is responsible for the virtual-to-physical address translation and cache protocols, the exception control system, the processor diagnostics, the operating modes (kernel, user, and debug), and interrupt control. Coprocessor 0 also contains the logic for identifying and managing exceptions. Exceptions are caused by a variety of sources, including data, external events, or program errors.

The synthesizable RISC32E core implements four cycle MAC instructions. The multiply/divide unit (MDU) allows 32 bit by 32 bit MAC instructions to be issued every cycle. Instruction and data caches are 8 Kbytes each in size. In addition, each cache is organized as 4-way set associative. Both caches are virtually indexed and physically tagged to allow

access during the same clock cycle that the address is translated. There are 256 sets of instruction and data caches. The line size is 16 bytes for both the instruction and data cache. Both the instruction and data cache include a 32 bit read unit, and the data cache also permits write through without write allocation (a write policy is not applicable to instruction cache). An enhanced JTAG (EJTAG) block allows for single stepping of the processor as well as instruction and data virtual address breakpoints.

The microprocessor core includes transmit and receive circuitry that interface, respectively, with the transmitter and receiver circuitry of the free space optical interconnect. The complete circuitry of both the microprocessor transmitter and receiver are shown in Figure 9.7. The transmitter circuit includes two sets of data that can be transmitted via the FSOI. The first data set is from the microprocessor core stored on a dual port SRAM (Bank 0 in Figure 9.7). The SRAM is a 32 bit wide and 1 KWord long memory. The second data set, produced by a 32 bit counter, bypasses the microprocessor completely and is used to generate data for diagnostics or is used if the core is not functioning properly. Since the serializer on the optical transmitter is 64 bits and both the microprocessor and counter produce 32 bit data, the 32 bit data set is copied by doubling the interconnect of the original 32 bit data set. This 64 bit data set is sent to the optical transmitter for serialization, as shown in the top half of Figure 9.7.

The receiver circuit performs similarly to the transmit circuit, but in the reverse direction. The optical receiver sends 64 bit deserialized electrical data to the receiver circuit of the microprocessor core, as shown in the bottom right corner of Figure 9.7. The 64 bit

data includes a copy of the original 32 bit data set, and either one of the two sets is discarded. The 32 bit data set is clocked into D flip flops, which are written into another dual port SRAM (Bank 1 in Figure 9.7). The microprocessor reads the SRAM into local cache memory.

Circuits are also included to provide control signals to the transmitter and receiver circuits. Both the transmitter and receiver control units include write, clear, and ten bit address lines. In addition, the transmitter control unit also includes a clear line for the counter, and a transmitter select line that determines which data set is chosen for transmission (either the microprocessor or the counter data). Both the transmitter and receiver control logic also include a loopback line which verifies the functionality of the microprocessor transmitter and receiver circuits by completely bypassing the free space optical interconnect components. The loopback is a 64 bit bus that loops from the output of the 1:2 multiplexer in the transmitter to the input of the 2:1 multiplexer in the receiver, as shown on the right side of Figure 9.7.

The microprocessor transmitter and receiver circuits operate at one half the core frequency. The core frequency is set to 333 MHz, mainly limited by the two dual port SRAM banks, and therefore, the transmitter and receiver operate at 166 MHz. Since 64 bits are serialized and transmitted per optical link, an overall link bandwidth of approximately 10.65 Gbps is achieved.

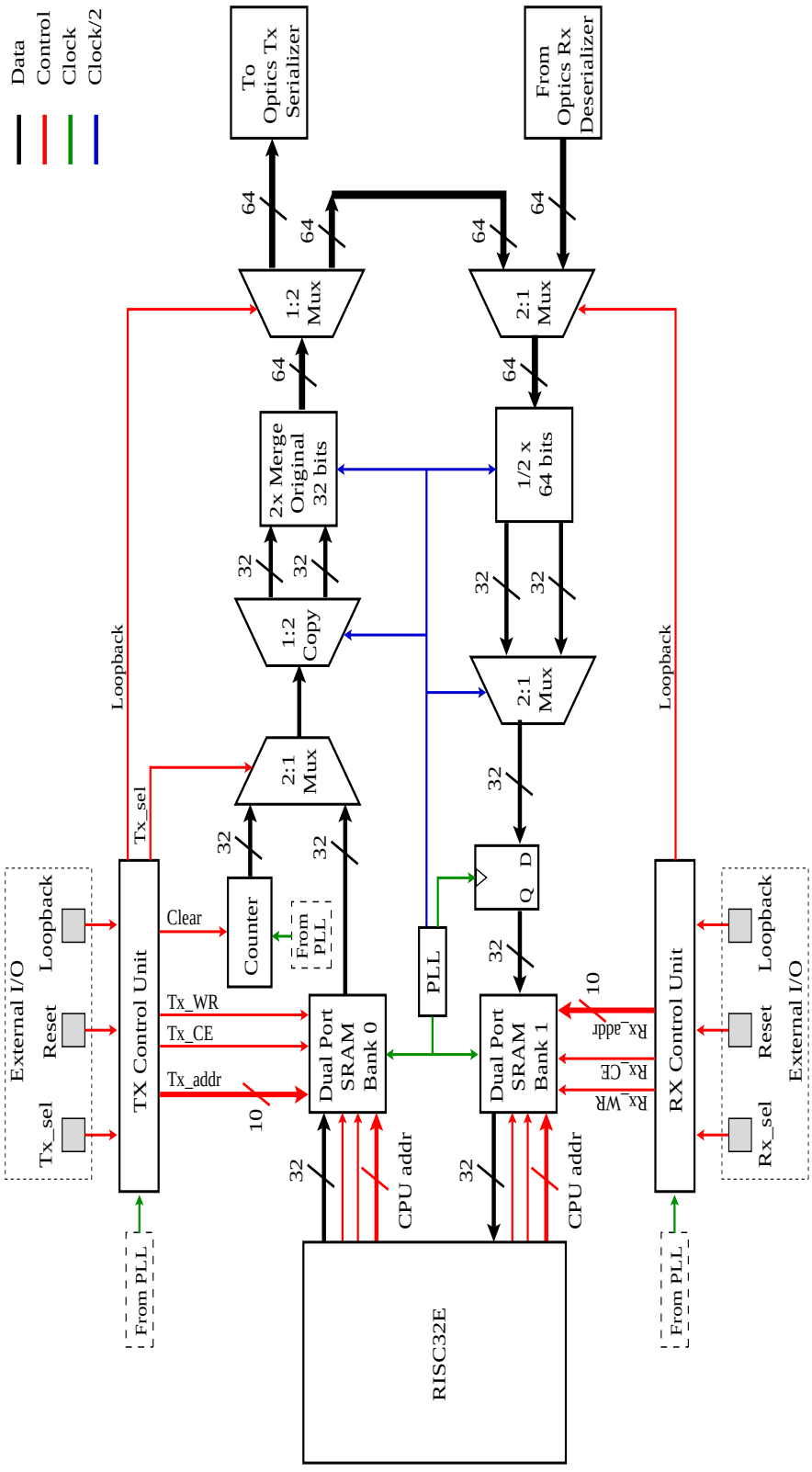


Figure 9.7: Schematic block representation of the transmit and receive circuitry at the interface between the RISC32E and, respectively, the 64 bit serializer and deserializer. The transmit circuitry is on the top half of the figure, while the receive circuits are on the bottom half.

9.5 Heat dissipation in stacked hybrid 3-D systems

3-D integrated circuits suffer from increased thermal gradients due to increased power densities, as discussed in Chapter 4. In particular, device planes farthest from the heat sink experience much higher temperatures as compared to devices planes adjoining the heat sink. The thermal issue is further magnified when the 3-D integrated free space optical interconnect system is considered. With the 3-D FSOI system, the top surface used as a bonding surface to the heat sink is replaced with a free space optical interconnect. A heat sink is present after the packaged free space optical network, but the thermal resistance for the heat generated by the CMOS layers and the VCSELs is significantly higher than if a traditional CMOS device plane is present.

Two important considerations are required to better understand the effect of the FSOI on hot spot formation and dissipation. The first requirement is an analysis of the heat generated due to the operation of the high current density VCSELs. The second requirement is better understanding of the thermal coupling between device planes as a means for heat removal. Two chapters are therefore dedicated in this disseration to these important issues. A thermal model and analysis of single and bundled VCSELs are provided in Chapter 10, while experimental results on interplane thermal coupling are discussed in Chapter 11.

9.6 Summary

One fundamental application of 3-D ICs is the integration of disparate technologies into a single stacked system. A 3-D integrated free space optical interconnect system is described that merges CMOS technology with a free space optical interconnect. The system

utilizes an all-to-all direct link network, where two cores are optically linked for core-to-core signaling. The FSOI also eliminates the constraints of the silicon optical modulators and switches, thereby increasing bandwidth while reducing the total power consumed by a link.

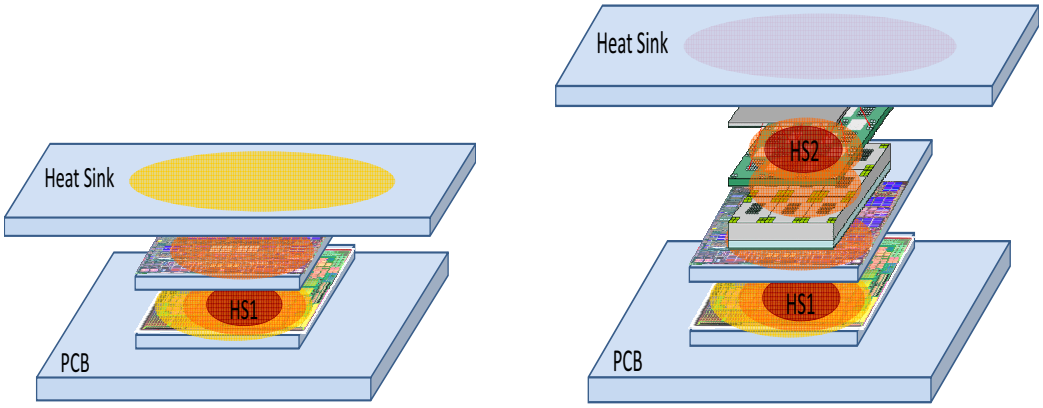
The optical components are described. The vertical cavity surface emitting lasers fabricated by the research group as well as the commercial VCSELs included for the prototype that experimentally verified the functionality of the 3-D integrated optical components are reviewed. The photodetectors fabricated by the research group are also described, which have an operational bandwidth of 9.3 GHz. The microlenses and micromirrors are the remaining optical components that are discussed.

The CMOS logic planes containing the transmitter and receiver circuits for the FSOI and the microprocessor core used to generate and receive data are also discussed. The maximum bandwidth of the FSOI transmitter and receiver circuits is 10 GHz. A microprocessor operating at 333 MHz and the interface circuits required to transmit signals through the FSOI are described. A 64 bit data word is sent from the microprocessor transmitter circuit to the serializer of the FSOI transmitter circuit, which is sent via optical link through modulated light waves at an 850 nm wavelength produced by a VCSEL. On the receiving core, a photodetector converts the optical signal back into an electrical signal which is amplified and filtered before deserialization and storage in SRAM memory for the core to read.

Chapter 10

Thermal Analysis of VCSEL Arrays

A 3-D integrated circuit is under increased thermal stress as compared to a 2-D IC. The thermal effects of including a high current density optical device plane within a 3-D stack only exasperates the issue. Two primary effects of integrating a free space optical interconnect (FSOI) system with 3-D integrated CMOS device planes include the insertion of an additional heat generating device plane, and an increase in the thermal resistive path of the CMOS device planes to the heat sink above. The two thermal issues are illustrated in Figure 10.1. The two logic planes comprising the CMOS device planes are shown in Figure 10.1(a), where a single heat generating plane exists and a heat sink is in close proximity to both planes. A significant challenge remains in controlling thermal effects if the heat source is on the bottom logic die. The challenge of managing the thermal effects in a 3-D integrated FSOI is significantly more demanding, as shown in Figure 10.1(b). In this case, a second heat source (HS2 in Figure 10.1(b)) and a longer resistive path to the heat sink from the logic planes increases the number and magnitude of hot spots and produces larger thermal gradients.



(a) (b)

Figure 10.1: Thermal profile of (a) two 3-D integrated logic dies, and (b) the two logic dies, optical device plane, and free space optics network. Both diagrams include a PCB and heat sink.

The thermal effects of including the FSOI require enhanced understanding of two thermal principles: 1) the additional heat generated by the vertical cavity surface emitting lasers (VCSELs), and 2) the propagation of heat through the device planes within a 3-D integrated system. This chapter is dedicated to the first thermal issue, where the effect of the different sized VCSEL arrays on the thermal characteristics is investigated. The second thermal principle is explored in Chapter 11.

A commercial software tool, specifically COMSOL, which is based on the finite element method (FEM), provides an appealing approach to accurately analyze the thermal behavior of VCSEL arrays [355–359]. COMSOL considers the material properties, mesh density at different locations, and irregularly shaped structures such as polyhedral domains and multilateral boundaries, while providing flexible control over the accuracy, memory, and computational time [356]. Different size VCSEL arrays with several oxide aperture diameters (D_a), substrate thicknesses, and current densities are considered in this chapter.

These parameters are used to obtain a precise distribution of the thermal behavior.

The chapter is organized as follows: An introduction to prior research on VCSEL modeling is provided in Section 10.1. The material and thickness of each layer within the VCSEL structure are described in Section 10.2. The physical characteristics of the materials and a model of the VCSEL arrays are discussed in Section 10.3. Based on an analysis of the simulations, the effects of different parameters on the internal temperature are described in Section 10.4. Several concluding remarks are provided in the final section of this chapter.

10.1 Research on VCSEL modeling

Due to small size, low divergence, non-astigmatic circular output beams, and inherent dynamic single longitudinal mode operation, vertical cavity surface emitting laser (VCSEL) arrays delivering optical energy are commonly applied in areas such as printing, engraving, sensors, and free space data transmission [360–362]. However, given the small oxide apertures, relatively thick and low thermal conductive substrates, and distributed Bragg reflectors (DBR), VCSEL structures are usually afflicted with severe thermal effects. Although a single VCSEL emitting a few milliwatts at room temperature was demonstrated in the late 1970's [363], large area VCSEL arrays emitting high output power beams producing significant heat have recently been developed [364, 365]. An understanding of the thermal behavior of VCSEL arrays is therefore crucial since the heat limits device performance, optical output power, threshold current, modulation speed, as well as the efficiency of the heat sink.

A number of research teams have modeled VCSEL thermal effects. The accuracy of these models is different and depends upon the tradeoff among flexibility, precision, and computational runtime [355–358, 366–375]. Although work exists on modeling VCSELs, there is significant inaccuracy. Some researchers have investigated simple 2-D structures [355, 366–373]; the VCSEL however, is a complex 3-D structure that includes electrical, thermal, and optical properties. Other researchers only consider a single VCSEL, not an oxide-confined array of VCSELs [356–358, 374, 375], or use simple physical parameters, removing details describing some of the layers comprising a VCSEL to reduce the complexity of the VCSEL model [355, 375]. In addition, research on VCSEL arrays has focused on "etched well" VCSELs, which are significantly different from the modern oxide confined VCSELs considered in this chapter [376]. However, since the size of the arrays directly determines the distribution of heat, and VCSEL arrays are comprised of different layers of materials in both the lateral and vertical directions, a physically based 3-D simulation model is preferable. To the authors' knowledge, this work is the first investigation of the thermal properties of oxide-confined VCSEL arrays.

10.2 VCSEL array structure

A 980 nm intra-cavity contacted oxide aperture VCSEL array with heterogeneous layers of aluminum arsenide and galium arsenide (AlAs/GaAs) distributed Bragg reflector (DBR) mirrors is considered. A cross view of the VCSEL structure is shown in Figure 10.2(a), and a top view is shown in Figure 10.2(b). The minimum pitch size for the VCSEL arrays is equivalent to two times the width of the active area, and is 80 μm. The

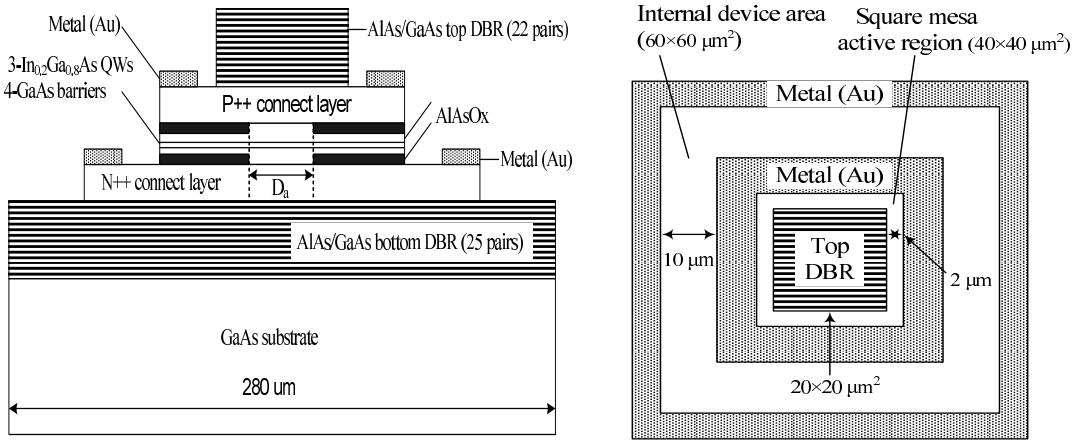


Figure 10.2: Schematic diagram of VCSEL structure (a) side view, and (b) top view.

general properties of the VCSEL are provided in Subsection 9.3.1. A discussion of the DBRs including the mirror reflectivity, threshold gain, and losses is provided below.

The top and bottom DBR mirrors consist of an alternating sequence of high refractive index layers, GaAs (84 nm), and low refractive index layers, AlAs (70 nm). The total number of AlAs/GaAs pairs on the top and bottom layers are respectively, 22 and 25. The reflectivity of the top and bottom mirrors is, respectively, 98.8% and 99.7%, based on (10.1), (10.2), and (10.3) [342, 377]

$$R_{t,b} = \frac{1 - b_{t,b}}{1 + b_{t,b}}, \tag{10.1}$$

$$b_t = \frac{n_{air}}{n_{GaAs}} \left(\frac{n_{AlAs}}{n_{GaAs}} \right)^{2M_t}, \tag{10.2}$$

$$b_b = \frac{n_{air}^2}{n_{air} n_{GaAs}} \left(\frac{n_{AlAs}}{n_{GaAs}} \right)^{2M_b}, \tag{10.3}$$

where $R_{t,b}$ is the top and bottom mirror reflectivity, n_{air} is the material reflectivity index of

air, n_{GaAs} is the material reflectivity index of GaAs, n_{AlAs} is the material reflectivity index of AlAs, M_t is the peak reflectivity of the top mirror layer, and M_b is the peak reflectivity of the bottom mirror layer. The thickness of the GaAs (d_{GaAs}) and AlAs (d_{AlAs}) layers is [342]

$$d_{GaAs,AlAs} = \frac{\lambda_B}{4n_{GaAs,AlAs}}, \tag{10.4}$$

where λ_B is the Bragg wavelength.

The threshold gain G_{th} is accurately determined by considering the losses within the mirror and active region and the mirror reflectivity. The gain is compensated for these losses within the cavity. In the threshold condition, the gain enhancement and penetration of the waves into the Bragg reflectors are described by G_{th} [342],

$$G_{th} = \alpha_a + \frac{1}{\Gamma_r d_a}\left[\alpha_i\left(L_{eff} - d_a\right) + ln\left(\frac{1}{\sqrt{R_t R_b}}\right)\right], \tag{10.5}$$

where λ_r is the gain enhancement factor, d_a is the total thickness of the active layers, L_{eff} is the effective mirror length, R_t is the intensity reflection coefficient of the top mirror, R_b is the intensity reflection coefficient of the bottom mirror, α_i describes the intrinsic losses in the passive section, and α_a describes the intrinsic losses in the active section. In simple structures, R_t and R_b can be determined from (10.1). The intrinsic losses in the active region do not contain band-to-band absorption which is included in the optical gain of the active region. The value of these parameters is listed in Table 10.1.

Alternatively, when considering losses, the threshold condition can be formulated using the maximum reflectivity, expressed as [342]

Table 10.1: Value of parameters used to calculate the threshold gain G_{th}.

Parameter	Value
Γ_r	1.8
d_a	24 nm
α_i	20 cm^{-1}
α_a	20 cm^{-1}
L_{eff}	1.3 μm
R_t	0.984
R_b	0.995

$$R_a = R_{t,b}\ exp(-2\alpha_i l_{eff}) \approx R_{t,b}(1 - 2\alpha_i l_{eff}), \tag{10.6}$$

where l_{eff} is the penetration depth. For lightly doped AlAs and GaAs layers in the top and bottom mirrors, the losses are much smaller, $2\alpha_i l_{eff} \ll 1$. The mirror reflectivity therefore does not change substantially.

From (10.1) - (10.6), G_{th} of the triple QW active region is 2500 cm^{-1}. The losses of the mirrors at 980 nm are low, 10 cm^{-1}. The threshold current density J_{th} is exponentially dependent on the gain threshold and inverse exponentially dependent on the material gain G_o [378],

$$J_{th} = \frac{n_w J_o}{\eta_i}\ exp\left(\frac{G_{th}}{G_o}\right), \tag{10.7}$$

where J_o, η_i, and n_w are, respectively, the transparency current density, internal quantum efficiency, and number of wells. J_o and G_o are, respectively, 80 to 90 A/cm^2 and 2400 cm^{-1} for an 8 nm thick $In_{0.2}Ga_{0.8}As$/GaAs layer emitting at 980 nm [379]. Based on these parameters, J_{th} is 666 A/cm^2.

10.3 Model of VCSEL array

VCSEL arrays are characterized in this chapter for aperature diameters D_a (shown in Figure 10.2) of 5, 10, 15 and 20 μm, substrate thicknesses of 100 μm, 200 μm and 300 μm, and array sizes ranging from 1×1 to 4×4. The thermal analysis utilizes a steady-state three-dimensional heat dissipation model, from which a profile of the thermal distribution is generated. This profile depends upon the size of the simulation domain since the heat can originate from a single or multiple hot spots and can spread across the much larger substrate. If thermal spreading is minimal, the ambient temperature will rise rapidly. Each side of the square substrate is as large as L_s = 280 μm [357]. The basic steady-state 3-D heat dissipation models are [380, 381]

$$-\nabla (k\nabla T) = Q, \tag{10.8}$$

$$k\nabla T = q_o + h\left(T_{inf} - T\right) + \epsilon\sigma\left(T_{amb}^4 - T^4\right), \tag{10.9}$$

where Q, k, T, q_o, h, ϵ, σ, T_{inf}, and T_{amb} are, respectively, the heat source density, thermal conductivity, initial temperature (assumed to be 298 K), inward heat flux, heat transfer coefficient, emissivity of the surface, Stefan-Boltzmann constant, ambient bulk temperature, and temperature of the surrounding radiation environment. The heat produced by the active region within the VCSEL arrays is primarily transferred from the device to the environment by two means: through the GaAs substrate to a copper heat sink by conduction, as modeled by (10.8), and convection and radiation through the top electrodes, as modeled by (10.9).

Since the device contacts directly to the copper heat sink, and one end of the heat sink is assumed to be at room temperature (about 298 K), the heat is mostly released by conduction, and convection and radiation are assumed negligible [357].

The VCSEL arrays are composed of multilayer thin film structures, which results in many internal and external boundary interfaces. The thermal boundary resistance (TBR) generated by these interfaces reduces the thermal conductivity of the material. This issue is known as interface phonons [382]. These interface phonons increase scattering and reduce the overall thermal resistance of the VCSEL arrays [383, 384]. Currently, however, there is minimal understanding of the effect of the phonons at the interface. In this chapter, to determine the effective thermal conductivity for an anisotropic thermal propagation process, the lateral and vertical thermal conductivities are determined from the following two expressions [385],

$$k_L = \frac{d_1 k_1 + d_2 k_2}{d_1 + d_2}, \tag{10.10}$$

$$k_V = \frac{d_1 + d_2}{\frac{d_1}{k_2} + \frac{d_2}{k_1}}, \tag{10.11}$$

where k_L is the lateral thermal conductivity, k_V is the vertical thermal conductivity, k_1 (k_2) is the thermal conductivity for layer 1 (layer 2), and d_1 (d_2) is the thickness of layer 1 (layer 2).

Both the AlAs/GaAs DBR regions and the InGaAs/GaAs QW regions comprise many multilayer structures in VCSEL arrays [386, 387]. Greater accuracy in the thermal analysis is achieved in these regions by determining the lateral thermal conductivity k_L and vertical

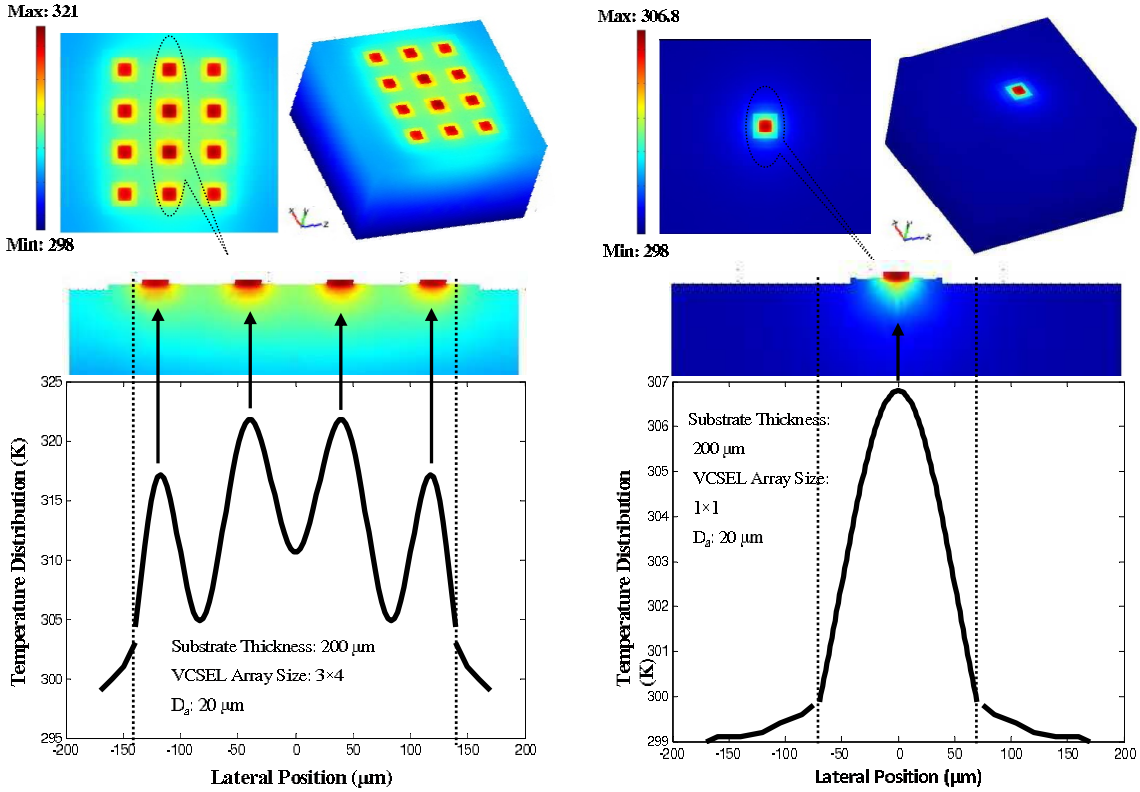
Table 10.2: Thermal conductivity of composite materials in VCSEL arrays mounted on a copper heat sink. [357, 358, 388–392]

Layer	Material	TC $\left(\frac{\mathrm{W}}{\mathrm{m\,K}}\right)$
p-contact metal	Au	315
Top DBR	AlAs/GaAs	$k_L = 32$, $k_V = 29$
p++ connect layer	GaAs	44
Oxidation	AlAs	0.7
Spacer layer	$Al_0.5Ga_0.5As$	11
3 MQW	$In_0.2Ga_0.8As$ (well)/GaAs (barrier)	$k_L = 29$, $k_V = 14$
Spacer layer	$Al_0.5Ga_0.5As$	11
Oxidation	AlAs	0.7
n++ connect layer	GaAs	44
Bottom DBR	AlAs/GaAs	$k_L = 32$, $k_V = 29$
Substrate	GaAs	44
n-contact metal	Au	315

thermal conductivity. Other than within the QW and DBR regions, isotropic thermal conductivity ($k = k_L = k_V$) is assumed. The thermal conductivity of the different materials in the VCSEL arrays attached to the copper heat sink is listed in Table 10.2 [357].

10.4 Simulation results and analysis

The heat flux and temperature distribution in the lateral direction for a 3×4 and 1×1 VCSEL array are shown in Figure 10.3. Note in this figure, since the VCSEL array is bonded with a copper heat sink below the bottom DBR and GaAs substrate, heat conducts radially from the heat sources through the substrate to the heat sink. The maximum internal temperature induced by the heat generated inside the device is 321 K for a 3×4 VCSEL array, and 306.8 K for a 1×1 VCSEL array, assuming the heat sink is initially at room temperature (298 K). As the size of the VCSEL array increases, each individual VCSEL contributes heat that accumulates within the device body, increasing the maximum internal temperature. The four VCSELs in the cross-sectional slice with the two center VCSELs



(a) (b)

Figure 10.3: Heat flux and temperature distribution in lateral direction for (a) a 3×4 VCSEL array, and (b) a 1×1 VCSEL array.

are also shown in Figure 10.3(a) and exhibit a higher maximum internal temperature. The higher temperature is due to the smaller temperature difference between the center VCSELs and the surrounding material. These center VCSELs are surrounded by other VCSELs that also generate heat and produce a smaller temperature gradient.

The maximum internal temperature as a function of the oxide aperture diameter (D_a), current density, and VCSEL array size is depicted in Figure 10.4. Note that the generated heat density increases with decreasing D_a. The reason is that the oxide layer adjacent to the heat generation region extends over a larger area, preventing the generated heat from

spreading [357]. Thus, for the example shown in Figure 10.4(d) with a 300 μm substrate, as D_a decreases from 20 μm to 5 μm, the maximum internal temperature difference is about 9 K for a 1×1 VCSEL array, and reaches 40.5 K for a 4×4 VCSEL array.

As illustrated in Figure 10.4, the maximum internal temperature also depends significantly on the substrate thickness. The VCSEL array with a substrate thickness of 300 μm exhibits a higher maximum internal temperature as compared to substrates with a 100 μm and 200 μm thickness. This characteristic can be attributed to the long thermal path from the QW and DBR regions through the thick substrate to the heat sink. For the example shown in Figure 10.4(a), where D_a = 20 μm and Jth = 666 A/cm^2, the substrate thickness varies from 100 μm to 300 μm, ∇T increases by 4 K for a 1×1 VCSEL array, and increases by 25 K for a 4×4 VCSEL array. This behavior indicates that VCSEL arrays formed by the bottom emitting VCSEL and epi-layer bonding scheme with a thin substrate are expected to exhibit a lower thermal resistance due to the shorter transfer path between the heat source and heat sink.

The current density also affects the internal temperature of the VCSEL arrays. The dependence of the internal temperature on the injection current is [357, 392]

$$T_{int} = T_{hs} + R_{th}\left(VI + R_s I^2 - P_{opt}\right), \tag{10.12}$$

where T_{int} is the internal temperature, T_{hs} is the temperature of the heat sink, R_{th} is the thermal resistance (K/W), R_s is the series resistance (Ω), V is the forward voltage, I is the injection current, and P_{opt} is the optical output power. The resistive term R_s is an electrical parameter given in ohms, while R_{th} is a thermal resistance in Kelvins per watt. According

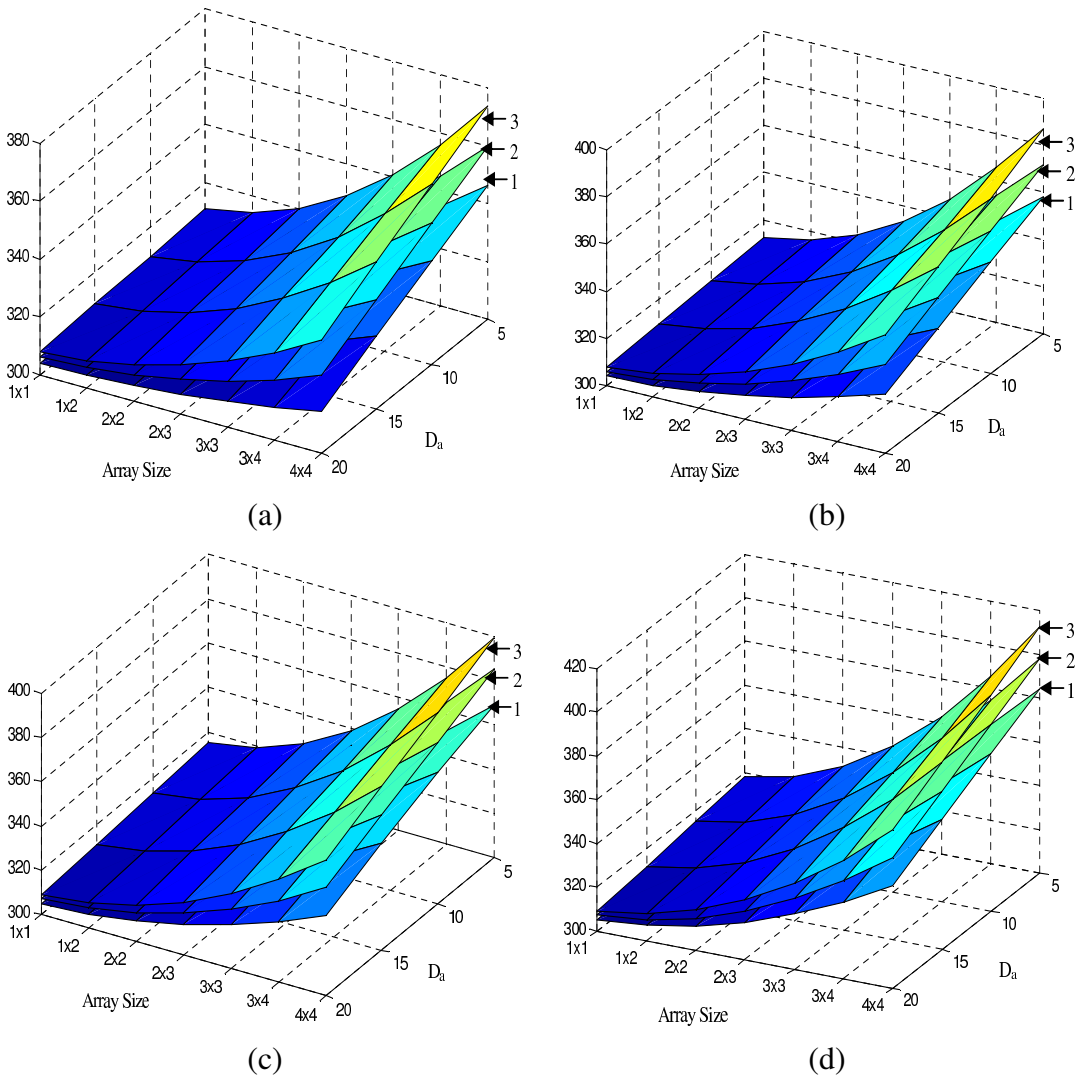


Figure 10.4: Maximum internal temperature for several D_a, VCSEL array sizes, and current densities, (a) J = 666 A/cm^2, (b) J = 800 A/cm^2, (c) J = 1000 A/cm^2, and (d) J = 1200 A/cm^2. Substrate thicknesses of 1 = 100 μm, 2 = 200 μm, and 3 = 300 μm.

to [372], the maximum internal temperature increases with higher injection current. A VCSEL array with a narrow D_a experiences a significant rise in temperature caused by an increase in the injection current, leading to poor heat removal and internal heating. Accordingly, at an injection current density of 1200 A/cm^2, ΔT is 115 K for D_a = 5 μm, a 300 μm substrate thickness, and a 4×4 VCSEL array. ΔT is only 86 K when the injection current is decreased to 666 A/cm^2 for the same parameters, as illustrated in Figs. 10.4(a) and 10.4(d).

In summary, the maximum internal temperature of intra-cavity contacted oxide aperture VCSEL arrays is affected by many factors including the aperture and array size. The maximum internal temperature is lower with decreasing current density. It is also worth noting that a thicker substrate can further degrade the thermal characteristics of the structure due to an increase in the thermal resistance.

10.5 Summary

Complex VCSEL arrays are modeled and the thermo-electrical behavior is analyzed, describing the heat flux and temperature distribution in a array structure. These results indicate that the maximum internal temperature of intra-cavity contacted oxide aperture VCSEL arrays depends strongly on the oxide aperture diameter, array size, current density, and substrate thickness. The diameter of the oxide aperture, array size, and injection current determine the magnitude of the generated heat; a thinner substrate reduces the physical distance to the heat sink. The maximum internal temperature of the VCSEL arrays can range from 306.5 K for a 20 μm D_a, 100 μm substrate thickness, 666 A/cm^2 current density,

and 1×1 array size to 412 K for a 5 μm D_a, 300 μm substrate thickness, 1200 A/cm^2 current density, and 4×4 array size.

An analysis of the heat generated by VCSEL arrays indicates that large microprocessors composed of multiple cores, requiring more VCSELs, produce significant increases in the peak temperature and thermal profile of a heterogeneous 3-D integrated circuit. In addition, traditional heat dissipation techniques such as a heat sink are not sufficient to manage the increased heat. Novel techniques such as microfluidic channels, as discussed in Chapter 4, are necessary to mitigate hotspot formation and reduce peak temperatures generated by the CMOS logic planes as well as the VCSEL arrays. An experimental study of the thermal coupling between device planes including a heat source located on the backside of the silicon is provided in the following chapter. Placing a heat source on the backside of the silicon die emulates the placement of VCSELs, as the VCSELs are bonded to the silicon backside in the 3-D integrated free space optical interconnect system.

Chapter 11

Inter-Die Thermal Coupling in 3-D Integrated Circuits

Two of the most omnipresent and challenging issues in high performance 3-D systems are power delivery and thermal management. The interdependence of these issues is of critical importance to 3-D systems, as high current loads on the power network can produce severe hot spots within a 3-D stack. The effect of hot spots on circuit operation is well documented in 2-D ICs [218] and is greatly exacerbated in 3-D ICs, requiring novel thermal mitigation and management techniques. Enhanced understanding of these interrelated design challenges in 3-D integration is therefore necessary to develop design techniques and methodologies to effectively deliver power while managing thermal effects.

TSV-based 3-D integrated circuits have rapidly progressed over the past decade to continue the trend of higher transistor density. Although possible to reduce the system-level power budget with 3-D integrated circuits by reducing the length of the global interconnect, vertical stacking of computational blocks within a smaller footprint increases the power density within the 3-D IC [393–395]. In addition, stacking device planes limits the thermal pathways available to efficiently remove heat from those dies farthest from the heat sink. A

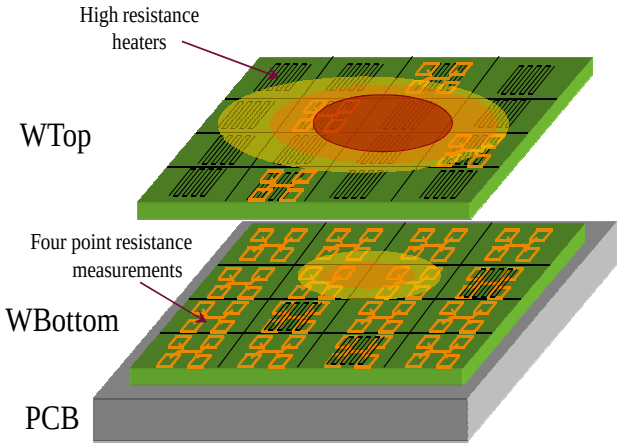


Figure 11.1: Heat propagation from one device plane spreading into a second stacked device plane.

schematic depiction of a two die stack with a hot spot on one die affecting the second die is shown in Figure 11.1. The increased thermal profile due to the higher power density and the lack of thermal conduits remain a limiting issue for 3-D ICs.

Prior work has focused on the simulation [396] and modeling [227, 397, 398] of hot spot formation and propagation within 3-D ICs. Tools, such as 3D-ICE [262] and a 3D extension to HotSpot [399, 400], to model thermal profiles of 3-D circuits have been developed and provide a visual interpretation of the hot spots based on the power requirements of each device plane. Models and simulations have been extended to block level floorplanning including global wire congestion, permitting the location of the highly active blocks within a 3-D IC floorplan to be adjusted based on the location of the thermal hot spots [247, 248, 401, 402]. Additional mitigation techniques, including the use of passive techniques such as thermal through silicon vias (TTSVs) [224] and active techniques such as microchannel or microfluidic cooling [263, 403], have been proposed to address heat removal in 3-D ICs.

Although extensive theoretical work has provided an understanding of heat flow in 3-D

ICs, there has been limited experimental results quantifying the flow of heat between device planes. Meindl *et al.* experimentally characterized the impact of microfluidic cooling techniques on both 2-D and 3-D circuits [256, 259, 404]. Additional experimental results have characterized microfluidic cooling methods [405]. Experiments characterizing an interplane cooling system for a vertically stacked DRAM/multiprocessor system-on-chip have been described [406]. Numerical and experimental characterization of thermal hot spots of a packaged DRAM on logic 3-D IC has also been described [407]. Similar to this work, interplane thermal propagation was investigated. However, the primary purpose of the results described in this paper is to characterize intra- and inter-plane thermal coupling to improve design methodologies and techniques for stacked ICs. The experimental results discussed herein provide insight into the effects of the location of the heat source and active cooling on heat flow within 3-D ICs.

The test circuit has been fabricated by Tezzaron Semiconductor in a 130 nm CMOS technology with 1.2 μm diameter TSVs. A face-to-face bonding technique to vertically stack the two logic device planes is used. This test circuit is designed to also evaluate the effects of inter- and intra-plane thermal resistance on hot spot formation.

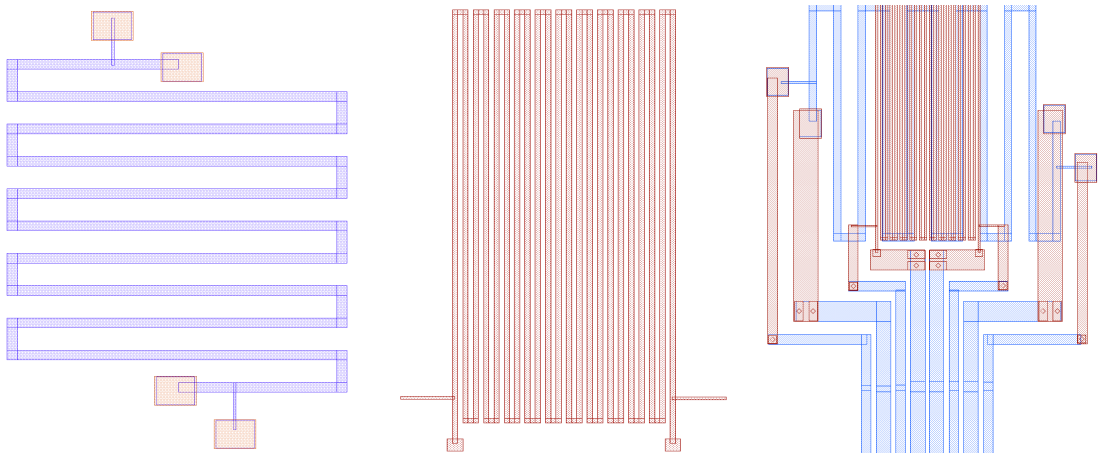
The 3-D test circuit is described in the following section. A brief summary of the Tezzaron 3-D process is provided in Section 11.1. Experimental characterization of thermal coupling for a set of test configurations are presented in Section 11.3. A discussion of the experimental results including the effects of hot spot formation and mitigation techniques is provided in Section 11.4. Some conclusions are offered in Section 11.5.

11.1 3-D IC fabrication technology

Each device plane is individually processed in a 130 nm CMOS technology, provided by Chartered Semiconductor, before 3-D bonding, TSV fabrication, and wafer thinning by Tezzaron. The Chartered fabrication process includes low power 1.5 volt and 2.5 volt transistors, six metal layers per device plane, a single polysilicon layer, dual gates for the 2.5 volt transistors, and low and nominal threshold devices [317,350]. The sixth metal level on each device plane is used to face-to-face bond the two dies to form the 3-D IC.

11.2 Thermal propagation test circuit

The test structures have been used to investigate thermal coupling between adjacent planes and include both resistive thermal sources and thermal sensors. The thermal sensors use four-point voltage measurements. Each thermal source is paired with a resistive thermal sensor on an adjacent metal level, and these pairs are distributed throughout each plane within a two layer 3-D logic stack. The heaters are 200 μm by 210 μm, similar to the dimensions of the heaters in [407], and are in metal 2. Within this area, the total length of the heater is 2,120 μm. The thickness of metal 2 is 0.42 μm, the width is 6 μm, and metal 2 has a nominal sheet resistance of 0.053 Ω/□. The resistance of the heaters is therefore 18.7 Ω. Joule heating through the resistive heater is adjusted by controlling the current flow, and therefore the I^2R power consumed within the 200 by 210 μm^2 area. The thermal sensors, with dimensions of 200 μm by 86 μm, are placed directly above the heaters in metal 3. The thermal sources are heater resistors with a maximum applied voltage of 28 volts, producing a maximum applied current of 1.5 amperes. The total length of the sensors is 4,442 μm.



(a) (b) (c)

Figure 11.2: Physical layout of the (a) on-chip resistive heater, (b) on-chip four-point resistive thermal sensor, and (c) overlay of the resistive heater and resistive thermal sensor.

The resistance of the sensors is 117.7 Ω based on a width of 2 μm, a metal 3 thickness of 0.42 μm, and a nominal sheet resistance of 0.053 Ω/□. The temperature sensor provides a calibrated four point measurement tested at a low current to avoid joule heating. The resistive heater, thermal sensor, and combined heater and sensor are shown in Figure 11.2.

Similar resistive heaters and thermal sensors are included in the aluminum back side metal layer, as shown in Figures 11.3(a) and 11.3(b). The differences between the heaters and sensors located in the logic planes and the back side metal include: 1) The heaters and sensors are not vertically stacked on adjacent metal layers as in the logic planes, as there is only a single back side metal layer, 2) the back side metal layer is almost three times the thickness of either metal 2 or 3 (1.2 μm as compared to 0.42 μm), and 3) due to the greater width and thickness, larger currents can pass through the back side metal. The heaters and sensors on the back side metal support thermal coupling through the thinned silicon to the thermal sensors on the internal logic layer, providing enhanced understanding of the effects

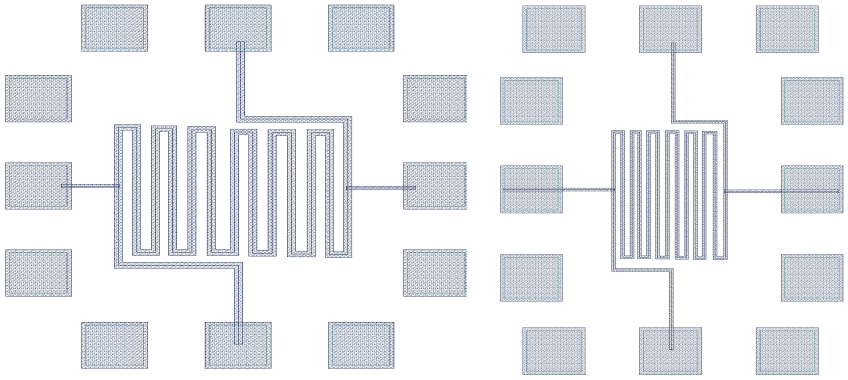


Figure 11.3: Physical layout of the (a) back metal resistive heater, and (b) back metal four-point resistive thermal sensor.

of thermal spreading through the silicon to the neighboring device planes

The location of the on-chip thermal sensors and resistive heaters with respect to the backside sensors is shown in Figure 11.4. A microphotograph of the 5 mm by 5 mm 3-D IC depicts two locations from which thermal data are collected. The center-to-center distance between the back metal sensors is 1.1 mm, while the on-chip sensors are 1.3 mm apart. A cross-cut view of the complete 3-D IC stack is shown in Figure 11.5. Each device plane, labeled as WTop and WBottom in the figure, includes a thermal sensor on metal 3 and a resistive heater on metal 2. The backside metal heaters and sensors are at the top of the stack, as shown in Figure 11.5. In addition, the thickness of both the WBottom silicon and the active portion of the test structure are included in the image shown in Figure 11.5, indicating a significantly smaller thermal resistive path to the top of the 3-D stack than to the board below.

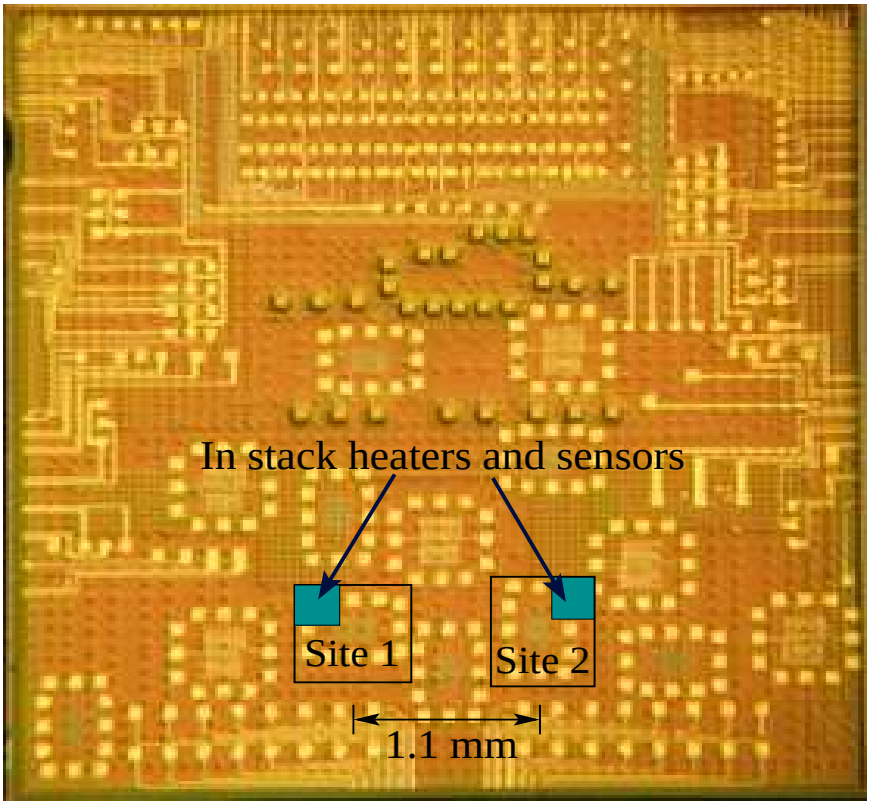


Figure 11.4: Microphotograph of the test circuit depicting the back metal pattern with an overlay indicating the location of the on-chip thermal test sites.

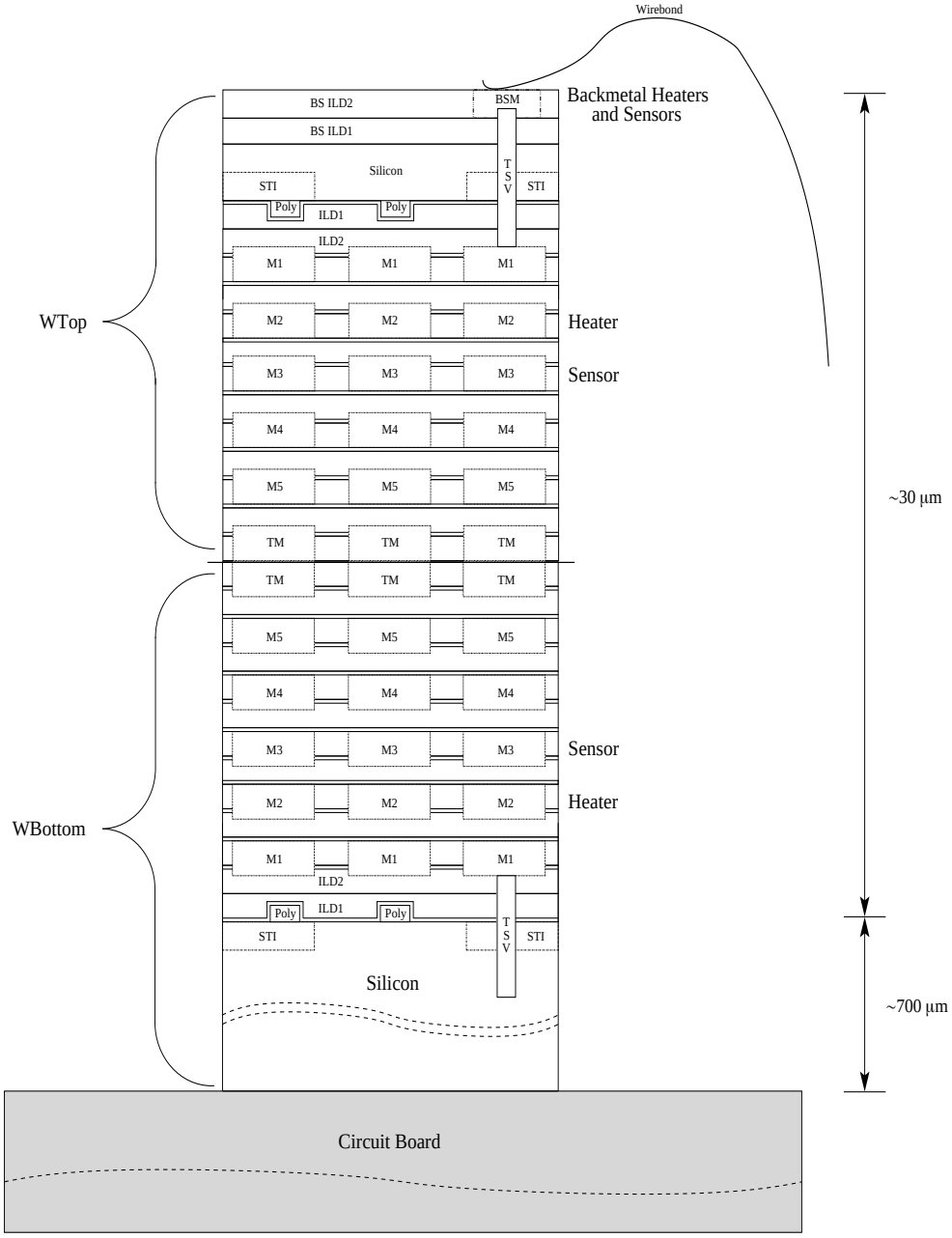
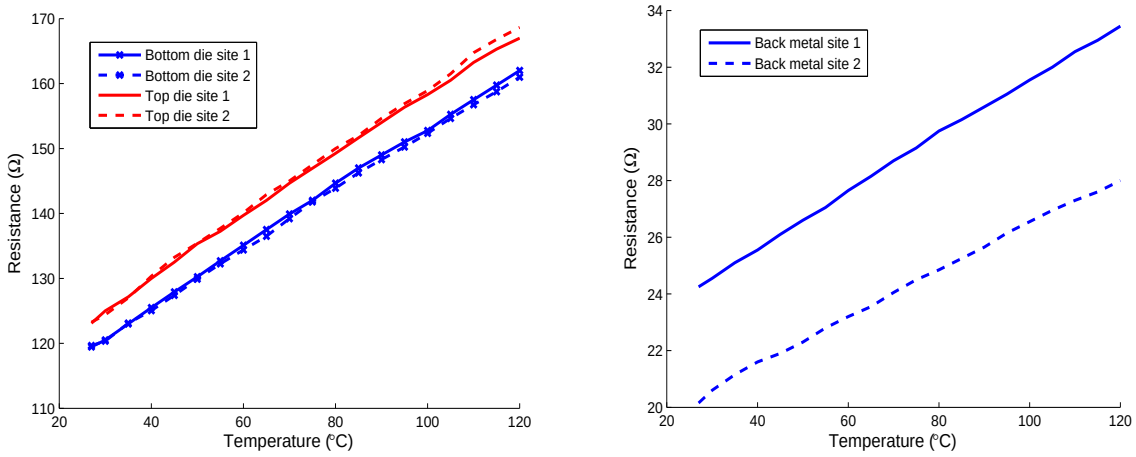


Figure 11.5: Placement of thermal heaters and sensors, respectively, in metals 2 and 3 in the two stacked device planes.

11.3 Setup and experiments

The on-chip and back metal thermal sensors require calibration prior to experimental analysis of thermal coupling between logic planes. Calibration is performed by setting the die temperature through a thermal chuck, and measuring the resistance of the sensors at each temperature from room temperature (27° C) to 120° C. The resistance as a function of temperature for each calibrated sensor is shown in Figure 11.6. All of the sensors exhibit a linear response to temperature. The on-chip sensors, shown in Figure 11.6(a), produce consistent results on the same logic layer. A difference in resistance exists between the top and bottom logic layers, as shown in Figure 11.6(a). Within a given die, the sensors produce consistent results, demonstrating that the thermal sensors can be calibrated from a single sensor. This behavior, however, is not the case with the back metal sensors, as shown in Figure 11.6(b). The difference in resistance between the two sites on the back metal reveals greater process variations on the back metal layer than the on-chip metal layers, requiring each thermal sensor to be individually calibrated and normalized at room temperature.

The experimental setup includes the use of an HP 4145B Semiconductor Parameter Analyzer and an HP 16058-60003 Personality Board. In addition, Interactive Characterization Software from Metric Technology Inc is used to define the settings for the parametric analyzer. A Keithley 2420 SourceMeter is used as a current source to supply the on-chip and back metal heaters with 0 to 110 mA of current (130 mA for the back metal). The parametric analyzer sweeps the voltage on the sensor from 0.1 volts to 0.6 volts in 0.01 volt increments, which permits the average resistance to be determined across this voltage



(a) (b)

Figure 11.6: Calibration of (a) on-chip thermal sensors, and (b) back metal thermal sensors.

range. Measurements have been made at each site and each sensor (a total of six data locations) for each current value. In addition, there are six different heater locations, as shown in Figure 11.5. The resistance measurements have been converted to temperature based on the results depicted in Figure 11.6.

The resistive heaters are controlled to provide different test conditions to investigate thermal propagation within the 3-D stack. A current is individually supplied to the heaters on WBottom and WTop. These results are shown in Figures 11.7(a) and 11.7(b). Another experiment examines the effects of limited in-plane thermal spreading by removing the metal heat spreaders surrounding the thermal sensors and resistive heaters, as described by comparing Figure 11.7(b) to Figure 11.7(c). The effect of placing two highly active device blocks on the thermal gradient is explored. In this case, two separate conditions are examined: 1) Heaters on WBottom and WTop are simultaneously active and stacked directly above each other, and 2) heaters on WBottom and WTop are simultaneously active

and physically nonaligned. These results are shown in Figures 11.7(d) and 11.7(e). A fourth test condition investigates the flow of heat from a back metal heat source to the on-chip thermal sensors. One of the two back metal sensors operates as a heater while the other sensor detects the temperature. The effects of placing the CMOS 3-D IC in a location where heat may couple from the backside of a thinned silicon substrate are shown in Figure 11.7(f). The final test condition examines the effects of active cooling on heat dissipation in a stacked IC. A 12 volt, 0.13 ampere, 6,500 RPM, and 8 CFM (cubic feet per minute) fan is placed one inch above the 3-D stack for convective heat removal. The fan is used for three different resistive heater conditions: 1) Active heater in WBottom site 1, 2) active heater on WBottom and WTop site 1, and 3) active heater on back metal site 1. The results of this cooling experiment are illustrated in Figures 11.7(g), 11.7(h), and 11.7(i).

11.3.1 Calibration of the four-point thermal sensors

Resistance data for calibrating the on-chip and back metal four-point thermal sensors are provided in Table 11.1. A temperature controlled hot plate is used to measure the resistance as a function of temperature for the thermal sensors on WTop, WBottom, and the back metal at both sites 1 and 2. The data are fitted to a second-order polynomial expression which is used to determine the temperature from the resistance measurements.

Table 11.1: Resistance as a function of temperature for calibrating the on-chip and back metal thermal sensors.

Heater location	Temperature (°C)																			
	27	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120
WBottom, Site 1	119.5	120.5	123.0	125.5	127.9	130.3	132.7	135.1	137.5	139.9	142.0	144.7	147.0	149.0	151.0	152.8	155.3	157.5	159.8	162.0
WBottom, Site 2	119.7	120.4	123.1	125.1	127.4	129.9	132.3	134.4	136.5	139.2	141.8	143.9	146.3	148.3	150.3	152.4	154.7	156.8	158.8	161.0
WTop, Site 1	123.1	125.0	127.2	130.0	132.5	135.4	137.3	139.7	142.0	144.7	147.0	149.3	151.7	154.0	156.4	158.3	160.5	163.3	165.3	167.0
WTop, Site 2	123.3	124.4	127.0	130.4	133.3	135.4	137.7	140.1	142.9	145.0	147.5	150.0	152.0	154.7	156.9	158.9	161.5	164.8	166.8	168.7
Back metal, Site 1	24.25	24.55	25.10	25.55	26.10	26.60	27.05	27.65	28.15	28.70	29.15	29.75	30.15	30.60	31.05	31.55	32.00	32.55	32.95	33.45
Back metal, Site 2	20.15	20.60	21.15	21.60	21.90	22.30	22.80	23.20	23.55	24.05	24.50	24.85	25.25	25.65	26.15	26.55	26.95	27.30	27.60	28.00

11.3.2 Temperature data from thermal sensors

Temperature data for all test conditions from the four on-chip and two back metal resistive thermal sensors are listed in Table 11.2. The data is from converted measured resistances. The location of the resistive heaters determines the particular testing condition. The temperature at the six thermal sensors is listed as a result of activity from the single heaters, double heaters, and active cooling.

11.4 Design considerations based on experimental results

The experimental test conditions described in Section 11.3 provide insight into thermal propagation and hot spot formation in 3-D ICs. Design considerations to minimize hot spot formations are discussed below.

11.4.1 Effects and mitigation of placement on hot spot formation

The effects of stacking two dies on the temperature profile of a 3-D IC are significant, as shown in Figures 11.7(a) and 11.7(b). The maximum observed temperature on WBottom site 1 increases by 65.7% when the resistive heater is active on WBottom as compared to an active heater on WTop (100.9° C from 60.9° C), as indicated in Table 11.3. The maximum observed temperatures occur for the maximum applied current of 110 mA. For heater currents less than 110 mA, the per cent increase in temperature for WBottom as compared to WTop decreases due to the exponential drop in temperature as the current is reduced. Data from the remaining five thermal sensors reveal a temperature increase of 6.1% to 13.0%

(see Table 11.3), with the 13.0% increase occurring on the sensor located directly above the active heater on WTop. Although the measured temperature is lower at the thermal sensors 1.3 mm away from the hot spot, the drop in temperature is not as significant as reported in [407], where over a 50% drop in temperature is measured at a distance of 500 μm. The observed temperature on WBottom when a resistive heater is active on WBottom site 1 is within 20% of the measured temperature at the same WBottom sensor when an active heater is on WTop site 1 for all heater currents less than 50 mA. Placing a highly active device block in WBottom requires special consideration as the thermal resistance along the path to the top of the 3-D stack is higher than to a block placed in WTop.

Three potential mitigation techniques can be considered. During placement and routing, blocks with high current loads can be placed in locations that minimize the thermal resistance between the block and the heat sink. Since the current load greatly affects the thermal profile (as revealed by less than a 20% change in temperature for currents less than 50 mA), another approach is to control the activity of the block to ensure that the average current load remains below an acceptable value, thereby not forming hot spots. Passive and active cooling techniques such as TTSVs and microfluidic channels can also be used to remove heat from the internal device planes by changing the thermal resistive paths of the highly active circuit blocks.

11.4.2 Horizontal and vertical thermal conduits

The experiment examining in-plane thermal spreading (Figures 11.7(b) and 11.7(c)) requires additional design consideration. By removing the in-plane metal heat spreaders

surrounding the thermal sensors and resistive heaters, the ability of heat to spread horizontally is diminished. The maximum temperature from a sensor directly above an active heater in WTop site 2 is 16.4% higher than an active heater in WTop site 1 for a heater current of 110 mA. All other sensors exhibit temperatures 14.9% to 27.7% higher than when metal heat spreaders are present and a 110 mA heater current is applied, as listed in Table 11.4. The increase in temperature at site 2 when the heater on WTop site 2 is active is expected; however, the lower temperature indicated by the negative change for the on-chip sensors at site 1 is not as great as expected. The thermal resistive path from site 2 to the heat sink is larger and therefore produces higher on-chip temperatures. Placing thermal spreaders to horizontally distribute the heat and thereby lower the effective thermal resistance reduces the maximum temperature experienced by those blocks 1.3 mm and a device plane away. It is therefore critical to reduce both the inter- and intra-plane thermal resistances by providing a path (or multiple paths) for the heat to flow from the hot spots to the heat sink.

11.4.3 Multiple aligned active blocks

The placement of two highly active circuit blocks aligned directly above one another has a significant effect on the thermal profile of an entire 3-D IC. Placing an active block on WBottom produces a higher thermal resistive path than a block placed on WTop. An analysis of two vertically aligned active circuit blocks has been performed by comparing the temperature for a single heater placed on WBottom with the increase in temperatures

caused by placing two active heaters located at WTop and WBottom site 1. The largest temperature increase occurs at WTop site 1, where a 79.4% increase in temperature is observed (from 69.6° C to 124.8° C) for an applied current of 110 mA. The maximum on-chip temperature occurs on WBottom site 1, where the maximum temperature increases from 100.9° C to 161.0° C when a current of 110 mA flows through, respectively, one active heater and two active heaters, producing a 59.6% increase in temperature. The remaining four sensors reveal an increase in temperature of 56.5% to 66.9%, corresponding to absolute temperatures of 81.4° C to 91.9° C from 52.1° C to 55.6° C. The per cent increase in temperature when two active circuit blocks are vertically stacked as compared to a single active block is listed in Table 11.5, and the corresponding temperatures are listed in Table 11.2. The magnitude of the current has a significant effect on the thermal profile within a 3-D IC. When applying 40 mA through both heaters, the increase in temperature remains below 20% for all thermal sensors as when applying 40 mA to just the WBottom heater. Two mitigation techniques include limiting the current flow, or deactivating the circuit block to allow heat to flow from the hot spot. Deactivation is particularly useful in high activity circuit blocks. Active and passive heat removal techniques such as microfluidic channels and thermal TSVs also apply, similar to the case of a single heater.

11.4.4 Multiple non-aligned active blocks

The effect of increased spacing on the maximum temperature between vertically non-aligned active circuits has been investigated. A comparison is made between two vertically aligned resistive heaters (both heaters in site 1) and two heaters separated by 1.3 mm

(WBottom heater site 1 and WTop heater site 2). A 0.0% to 5.4% reduction in temperature for all six thermal sensors up to a current of 30 mA is observed when shifting the WTop heater from site 1 to site 2. The on-chip thermal sensors at site 1 reveal temperature changes of -2.9% to 0.5% for currents up to 90 mA, and a maximum increase of 11.8% at WBottom and 9.9% at WTop for a peak current of 110 mA, as listed in Table 11.6. The on-chip thermal sensors at site 2 detect an exponentially increasing temperature from an applied current of 40 mA (0.6% for WTop and 1.8% for WBottom) to a peak current of 110 mA (54.3% and 51.0%). The exponential increase in temperature at the back metal sensors for currents above 30 mA indicates strong thermal coupling through the back metal layer. A maximum on-chip temperature of 179.9° C occurs at WBottom site 1. The maximum temperature at WTop site 1 is 137.2° C, while the remaining sensors exhibit a maximum temperature ranging from 121.0° C to 130.6° C.

Two primary design issues are noted: 1) Moving highly active circuit blocks during floorplanning reduces the maximum temperature as compared to the vertically aligned circuits. 2) Although nonaligned circuit blocks reduce the maximum temperature, the effective thermal resistance to the heat sink (air in this case) greatly affects the maximum on-chip temperature, as indicated by the increase in temperature for currents above 30 mA.

11.4.5 Additional design considerations

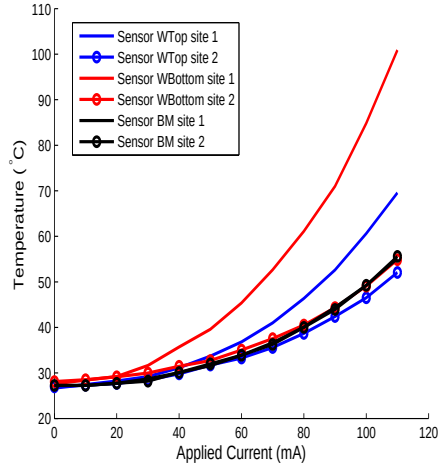
The fourth and fifth test conditions evaluate, respectively, heat flow from a source on the back side of the silicon to the on-chip thermal sensors and the reduction in the maximum temperature when a fan is used to convectively cool the 3-D IC. All thermal sensors produce

results within 6% of each other for all current load conditions through the back metal heater. The heat flows from a source on the back metal, evenly dispersing heat to the rest of the 3-D IC. Thermal spreading from a hot spot on the back metal is more pronounced. It is, however, more difficult to isolate a circuit block from thermal effects caused by a back metal heat source. For the fifth test condition, a fan one inch above the 3-D IC reduces the temperature from 0.7% for a 10 mA current through a back metal heater to 21.9% for a current load of 130 mA. In the case where both on-chip heaters in site 1 are active, the maximum temperature is reduced by 0.7% and 11.9% for, respectively, currents of 10 mA and 110 mA. The per cent reduction in temperature for three different heater configurations and heater currents of 0 mA to 110 mA is listed in Table 11.7. Despite an approximately 12% reduction in peak temperature with the use of a fan, additional heat removal techniques like TTSVs and microfluidic channels are necessary for those hot spots located deep within the 3-D IC.

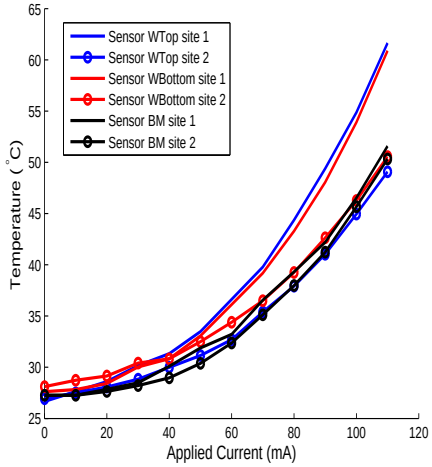
11.5 Summary

A primary challenge in TSV-based 3-D integrated circuits is the management of large thermal gradients across different device planes. The performance and reliability of a 3-D integrated circuit is significantly affected by large heat gradients. Thermal effects can potentially alter the performance of the clock and power networks due to hot spots. Proper block placement and active and passive heat removal techniques are critical to ensure a 3-D IC operates within the specified thermal design power (TDP) envelope.

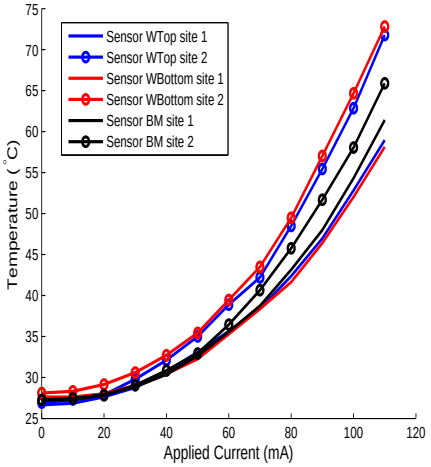
A three-dimensional test circuit examining thermal propagation within a 3-D integrated stack has been designed, fabricated, and tested. Design insight into thermal coupling in 3-D ICs through experimental means is provided, and suggestions to mitigate thermal effects in 3-D ICs are offered. Two wafers are vertically bonded to form a 3-D stack. Intra- and inter-plane thermal coupling is investigated through single point heat generation using resistive thermal heaters and temperature monitoring through four-point resistive measurements. Five test conditions are examined to characterize thermal propagation in 3-D integrated circuits. The five conditions are: 1) the location of the active circuit, 2) the impact of intra-plane thermal spreading, 3) the relative vertical alignment of the two active circuits, 4) heat flow from a thermal source at the back side of the silicon, and 5) convective cooling of the 3-D IC. Design suggestions are provided to better manage hot spot formation while reducing any effects on neighboring circuit blocks. The position of a block relative to a heat sink significantly affects the thermal resistance and therefore the flow of heat from the hot spots. The proper design of both inter- and intra-plane thermal conduits provides an additional means for removing heat. This test circuit provides enhanced understanding of thermal hot spot formation and propagation within 3-D integrated circuits.



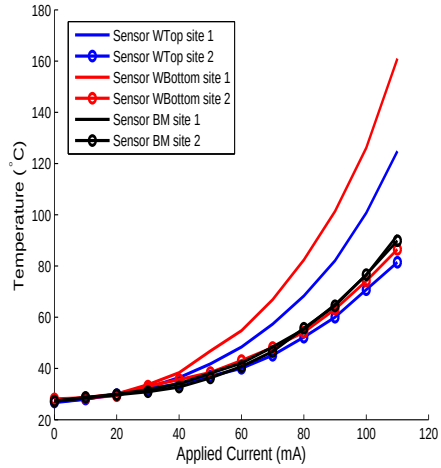
(a) WBottom site 1; no fan



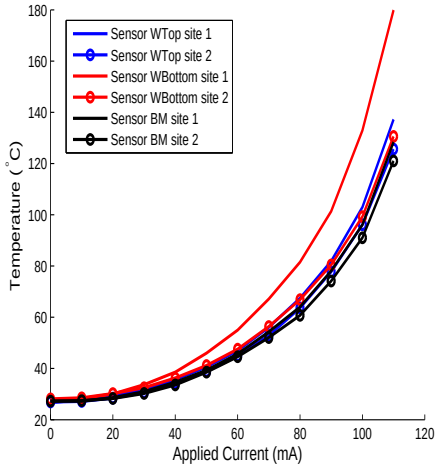
(b) WTop site 1; no fan



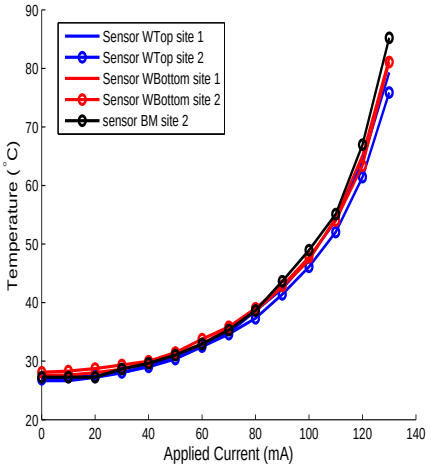
(c) WTop site 2; no fan



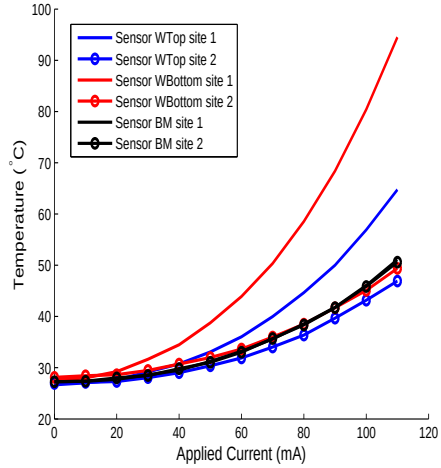
(d) WBottom site 1, WTop site 1; no fan



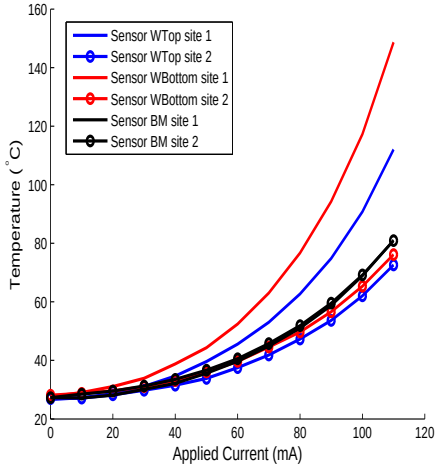
(e) WBottom site 1, WTop site 2; no fan



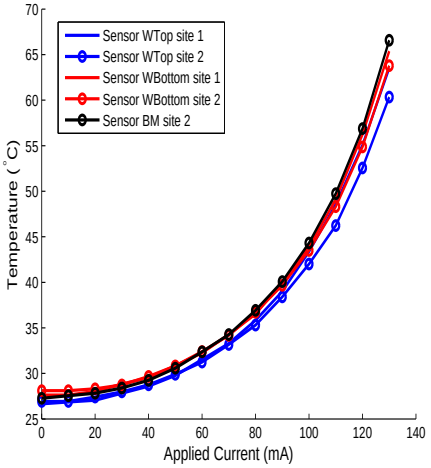
(f) Back metal site 1; no fan



(g) WBottom site 1; fan



(h) WBottom site 1, WTop site 1; fan



(i) Back metal site 1, fan

Figure 11.7: Experimental results for the different test conditions. Each label describes the device plane, site location of the heater, and whether active cooling is applied.

Table 11.2: Temperature measurements from the four on-chip and two back metal thermal sensors for different heater activities.

Heater location	Convective cooling	Sensor location	Current (mA)													
			0	10	20	30	40	50	60	70	80	90	100	110	120	130
WBottom, Site 1	No	WTop, Site 1	26.7	27.4	28.2	29.2	31.1	33.7	36.8	41.0	46.4	52.6	60.6	69.6	—	—
		WTop, Site 2	26.9	27.3	27.7	28.5	29.8	31.7	33.3	35.6	38.7	42.4	46.5	52.1	—	—
		WBottom, Site 1	27.6	28.4	29.2	31.6	35.7	39.6	45.4	52.7	61.1	71.0	84.8	100.9	—	—
		WBottom, Site 2	28.1	28.5	29.1	30.0	31.5	32.7	35.0	37.6	40.5	44.3	49.1	54.9	—	—
		Back Metal, Site 1	27.3	27.3	27.7	28.7	30.1	31.6	33.9	36.6	40.1	44.4	49.2	55.1	—	—
		Back Metal, Site 2	27.3	27.3	27.7	28.2	30.1	32.0	33.9	36.3	40	43.9	49.2	55.6	—	—
WBottom, Site 1	Yes	WTop, Site 1	26.7	27.0	27.6	29.2	30.7	33.1	36.0	40.0	44.7	50.0	56.9	64.8	—	—
		WTop, Site 2	26.8	27.1	27.3	28.1	29.0	30.4	31.9	34.0	36.4	39.7	43.2	46.9	—	—
		WBottom, Site 1	27.8	28.0	29.2	31.6	34.5	38.8	43.9	50.3	58.6	68.4	80.4	94.5	—	—
		WBottom, Site 2	27.9	28.4	28.6	29.5	30.7	32.0	33.7	36.0	38.5	41.7	45.1	49.4	—	—
		Back Metal, Site 1	27.2	27.4	27.7	28.6	29.5	31.2	33.2	35.6	38.5	41.8	46.0	51.1	—	—
		Back Metal, Site 2	27.0	27.4	28.0	28.5	29.8	31.0	33.0	35.7	38.4	41.7	45.9	50.6	—	—
WTop, Site 1	No	WTop, Site 1	27.0	27.6	28.6	30.2	31.3	33.5	36.6	39.8	44.4	49.4	54.9	61.6	—	—
		WTop, Site 2	26.8	27.5	28.1	28.8	30.0	31.1	32.7	35.4	37.9	41.0	44.9	49.1	—	—
		WBottom, Site 1	27.5	27.8	28.4	30.0	30.8	33.1	36.1	39.2	43.3	48.1	53.9	60.9	—	—
		WBottom, Site 2	27.8	28.7	29.1	30.4	30.8	32.5	34.4	36.5	39.2	42.6	46.3	50.6	—	—
		Back Metal, Site 1	27.1	27.3	27.6	28.2	29.0	30.4	32.4	35.1	38.0	41.2	45.7	50.3	—	—
		Back Metal, Site 2	27.2	27.3	27.8	28.5	30.1	31.9	33.2	36.5	39.3	42.2	46.5	51.6	—	—
WTop, Site 2	No	WTop, Site 1	26.5	26.9	27.6	28.8	30.7	33.1	35.6	38.4	42.4	47.0	52.8	59.0	—	—
		WTop, Site 2	26.7	27.3	27.9	29.8	32.1	35.0	38.9	42.2	48.5	55.4	62.8	71.8	—	—
		WBottom, Site 1	27.5	27.6	28.0	29.0	30.4	32.3	35.3	38.4	41.6	46.4	52.0	58.1	—	—
		WBottom, Site 2	27.4	28.3	29.1	30.6	32.7	35.4	39.5	43.5	49.5	57.0	64.7	72.8	—	—
		Back Metal, Site 1	27.1	27.3	27.9	28.9	30.4	32.6	35.6	38.7	43.2	48.0	54.3	61.4	—	—
		Back Metal, Site 2	26.9	27.5	27.8	29.1	30.8	32.9	36.4	40.7	45.8	51.7	58.1	65.9	—	—
WBottom, Site 1 WTop, Site 1	No	WTop, Site 1	27.1	28.0	30.2	32.5	36.4	41.8	48.4	57.3	68.3	82.1	100.7	124.8	—	—
		WTop, Site 2	27.3	28.1	29.8	31.9	34.0	36.9	40.1	45.1	52.2	60.0	70.8	81.4	—	—
		WBottom, Site 1	27.9	28.6	30.0	33.7	38.4	46.8	54.8	66.9	82.4	101.4	126.1	161.0	—	—
		WBottom, Site 2	28.0	28.5	29.4	33.1	35.7	38.4	43.1	48.2	54.2	63.1	74.2	86.5	—	—
		Back Metal, Site 1	27.2	28.5	29.9	31.7	34.1	38.2	42.1	48.4	55.3	64.9	76.5	91.9	—	—
		Back Metal, Site 2	27.6	28.7	29.7	30.9	32.7	36.4	40.8	46.6	55.7	64.5	76.6	89.9	—	—
WBottom, Site 1 WTop, Site 2	No	WTop, Site 1	27.2	27.8	29.6	31.7	35.6	40.6	47.0	56.1	67.5	81.9	103	137.2	—	—
		WTop, Site 2	27.0	27.1	28.5	30.6	34.2	38.9	44.9	52.8	63.7	77.7	96.1	125.6	—	—
		WBottom, Site 1	27.8	28.2	30.0	33.7	38.6	46.0	55.0	67.1	81.5	101.4	132.9	179.9	—	—
		WBottom, Site 2	28.0	28.5	30.2	32.5	36.3	41.2	47.6	56.4	66.9	80.4	99.3	130.6	—	—
		Back Metal, Site 1	27.3	27.5	28.3	30.3	33.6	38.5	44.6	52.1	60.7	74.1	91.0	121.0	—	—
		Back Metal, Site 2	27.2	27.3	28.6	30.9	34.6	39.3	45.7	54.3	64.0	78.1	96.1	128.4	—	—
WBottom, Site 1 WTop, Site 1	Yes	WTop, Site 1	26.9	27.2	28.6	31.1	34.7	39.6	45.6	53.0	62.7	74.8	90.8	112.0	—	—
		WTop, Site 2	26.8	27.1	28.3	29.8	31.5	33.8	37.5	41.8	47.3	53.6	62	72.6	—	—
		WBottom, Site 1	27.9	29.0	31.0	33.9	38.8	44.3	52.4	63.0	76.7	94.3	117.3	148.6	—	—
		WBottom, Site 2	28.1	28.7	29.6	31.0	33.1	36.1	39.7	44.6	49.7	56.6	65.3	76.2	—	—
		Back Metal, Site 1	27.1	27.2	28.1	30.1	32.2	35.6	39.7	44.9	51.0	58.7	68.9	81.0	—	—
		Back Metal, Site 2	27.4	28.5	29.5	31.2	33.5	36.6	40.6	45.8	51.9	59.5	69.2	80.9	—	—
Back metal, Site 1	No	WTop, Site 1	26.6	26.7	27.2	28.0	29.2	30.5	32.5	34.9	38.6	42.4	47.6	54.1	64.1	79.3
		WTop, Site 2	27.1	27.3	27.5	28.1	29.0	30.4	32.5	34.6	37.3	41.4	46.1	52.1	61.4	75.9
		WBottom, Site 1	27.5	27.6	28.0	28.6	29.8	31.2	32.9	35.7	38.6	42.7	47.2	54.5	65.2	81.7
		WBottom, Site 2	28.0	28.3	28.7	29.4	30.0	31.5	33.8	35.9	39.0	42.9	47.8	54.0	63.4	81.1
		Back Metal, Site 1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Back Metal, Site 2	26.9	27.3	27.3	28.7	29.6	31.0	32.9	35.3	38.6	43.6	49.0	55.1	67.0	85.2
Back metal, Site 1	Yes	WTop, Site 1	27.0	26.9	27.0	27.8	28.6	29.8	31.5	33.3	35.8	39.0	43.4	48.6	55.1	63.5
		WTop, Site 2	27.1	26.9	27.4	28.0	28.7	29.9	31.2	33.2	35.3	38.4	42.0	46.2	52.5	60.3
		WBottom, Site 1	27.4	27.6	28.0	28.8	29.5	30.8	32.3	34.3	36.5	40.0	43.9	49.1	56.2	65.4
		WBottom, Site 2	27.9	28.1	28.3	28.7	29.7	30.8	32.4	34.2	36.7	39.7	43.5	48.3	54.9	63.8
		Back Metal, Site 1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
		Back Metal, Site 2	27.1	27.5	27.8	28.4	29.2	30.6	32.4	34.3	36.9	40.1	44.3	49.7	56.9	66.7

Table 11.3: Per cent increase in temperature when active circuit is located in internal plane (WBottom).

Heater location	Sensor location	Current (mA)											
		0	10	20	30	40	50	60	70	80	90	100	110
WBottom, Site 1	WTop, Site 1	0.00	-0.72	-1.39	-3.31	-0.64	0.60	0.55	3.02	4.50	6.48	10.38	12.99
	WTop, Site 2	0.00	-0.73	-1.42	-1.04	-0.67	1.93	1.83	0.56	2.11	3.41	3.56	6.11
% increase as compared to	WBottom, Site 1	0.00	2.16	2.82	5.33	15.91	19.64	25.76	34.44	41.11	47.61	57.33	65.68
	WBottom, Site 2	0.00	-0.70	0.00	-1.32	2.27	0.62	1.74	3.01	3.32	3.99	6.05	8.50
WTop, Site 1	Back Metal, Site 1	0.00	0.00	0.36	1.77	3.79	3.95	4.63	4.27	5.53	7.77	7.66	9.54
	Back Metal, Site 2	0.00	0.00	-0.36	-1.05	0	0.31	2.11	-0.55	1.78	4.03	5.81	7.75

Table 11.4: Per cent increase in temperature from intra-plane thermal spreading.

Heater location	Sensor location	Current (mA)											
		0	10	20	30	40	50	60	70	80	90	100	110
WTop, Site 2	WTop, Site 1	0.00	-2.81	-3.39	-4.52	-1.87	-1.17	-2.69	-3.48	-4.50	-4.89	-3.71	-4.35
	WTop, Site 2	0.00	-0.70	-0.68	3.33	7.06	12.39	18.97	19.24	27.89	35.09	39.83	46.23
% increase as compared to	WBottom, Site 1	0.00	-0.72	-1.41	-3.35	-1.31	-2.45	-2.25	-2.09	-3.81	-3.46	-3.50	-4.53
	WBottom, Site 2	0.00	-1.45	0	0.69	6.12	9.06	14.73	19.17	26.08	33.75	39.77	44.05
WTop, Site 1	Back Metal, Site 1	0.00	0.00	0.34	1.32	0.94	2.38	7.15	6.01	9.77	13.73	16.74	19.05
	Back Metal, Site 2	0.00	0.69	0.68	3.01	6.52	8.41	12.62	15.78	20.50	25.35	27.14	30.94

Table 11.5: Per cent increase in temperature when two active circuit blocks are vertically aligned.

Heater location	Sensor location	Current (mA)											
		0	10	20	30	40	50	60	70	80	90	100	110
WTop, Site 1 and WBottom, Site 1	WTop, Site 1	0.00	2.12	6.89	11.35	17.01	24.05	31.40	39.88	47.29	55.98	66.19	79.43
	WTop, Site 2	0.00	2.81	7.62	12.15	14.22	16.45	20.41	26.85	35.03	41.58	52.10	56.47
% increase as compared to	WBottom, Site 1	0.00	0.71	2.76	6.40	7.43	18.24	20.72	27.04	34.85	42.77	48.64	59.55
	WBottom, Site 2	0.00	0.00	0.72	10.48	13.36	17.40	22.97	28.33	33.81	42.36	51.18	57.61
WBottom, Site 1	Back Metal, Site 1	0.00	4.49	7.82	10.53	13.22	20.76	24.25	32.03	37.98	46.12	55.42	66.86
	Back Metal, Site 2	0.00	5.18	7.13	9.70	8.80	13.66	20.28	28.60	39.29	46.81	55.63	61.65

Table 11.6: Per cent increase in temperature when two active circuit blocks are not vertically aligned.

Heater location	Sensor location	Current (mA)											
		0	10	20	30	40	50	60	70	80	90	100	110
WTop, Site 2 and WBottom, Site 1	WTop, Site 1	0.00	-0.69	-1.94	-2.41	-2.16	-2.86	-2.91	-2.14	-1.23	-0.26	2.22	9.94
	WTop, Site 2	0.00	-3.41	-4.51	-4.22	0.57	5.26	12.21	17.06	21.83	29.47	35.85	54.27
% increase as compared to	WBottom, Site 1	0.00	-1.40	0.00	0.00	0.53	-1.77	0.38	0.32	-1.08	0.00	5.41	11.75
	WBottom, Site 2	0.00	0.00	2.85	-1.90	1.77	7.17	10.41	16.99	23.32	27.35	33.85	51.01
WTop, Site 1 and WBottom, Site 1	Back Metal, Site 1	0.00	-3.63	-5.36	-4.47	-1.39	1.00	5.95	7.65	9.80	14.22	18.90	31.63
	Back Metal, Site 2	0.00	-4.92	-3.81	0.00	5.80	8.14	12.05	16.50	14.90	21.06	25.41	42.87
WTop, Site 2 and WBottom, Site 1	WTop, Site 1	0.00	1.41	4.82	8.68	14.48	20.51	27.57	36.89	45.49	55.57	69.89	97.28
	WTop, Site 2	0.00	-0.70	2.77	7.42	14.87	22.58	35.11	48.49	64.51	83.32	106.63	141.39
% increase as compared to	WBottom, Site 1	0.00	-0.71	2.76	6.40	8.00	16.14	21.18	27.45	33.39	42.77	56.68	78.29
	WBottom, Site 2	0.00	0.00	3.59	8.38	15.37	25.82	35.77	50.14	65.02	81.31	102.34	138.01
WBottom, Site 1	Back Metal, Site 1	0.00	0.69	2.04	5.59	11.64	21.97	31.64	42.12	51.49	66.90	84.81	119.64
	Back Metal, Site 2	0.00	0.00	3.05	9.70	15.11	22.92	34.77	49.82	60.04	77.73	95.18	130.95

Table 11.7: Per cent decrease in temperature with convective cooling.

Heater location	Sensor location	Current (mA)													
		0	10	20	30	40	50	60	70	80	90	100	110	120	130
WBottom, Site 1	WTop, Site 1	0.00	-1.41	-2.06	0.00	-1.25	-1.75	-2.14	-2.42	-3.67	-5.00	-6.10	-6.91	—	—
	WTop, Site 2	0.00	-0.70	-1.38	-1.35	-2.58	-4.25	-4.06	-4.35	-6.02	-6.45	-7.17	-9.87	—	—
% decrease as compared to	WBottom, Site 1	0.00	-1.41	0.00	0.00	-3.42	-2.07	-3.19	-4.37	-4.17	-3.66	-5.23	-6.34	—	—
	WBottom, Site 2	0.00	-0.37	-1.79	-1.74	-2.33	-2.25	-3.90	-4.22	-4.97	-6.00	-8.07	-10.01	—	—
WBottom, Site 1 (fan)	Back Metal, Site 1	0.00	0.34	0.00	-0.33	-1.88	-1.20	-1.96	-2.86	-3.82	-5.85	-6.50	-7.27	—	—
	Back Metal, Site 2	0.00	0.34	1.02	1.00	-0.94	-2.96	-2.52	-1.58	-4.07	-5.04	-6.89	-8.97	—	—
WTop, Site 1 and WBottom, Site 1	WTop, Site 1	0.00	-2.77	-5.16	-4.21	-4.86	-5.23	-5.81	-7.48	-8.23	-8.85	-9.88	-10.23	—	—
	WTop, Site 2	0.00	-3.41	-5.15	-6.63	-7.37	-8.39	-6.31	-7.38	-9.46	-10.65	-12.31	-10.87	—	—
% decrease as compared to	WBottom, Site 1	0.00	1.41	3.36	0.60	1.07	-5.31	-4.22	-5.77	-6.96	-6.98	-7.03	-7.67	—	—
WTop, Site 1	WBottom, Site 2	0.00	0.73	0.71	-6.33	-7.08	-6.05	-7.89	-7.54	-8.34	-10.35	-11.91	-11.94	—	—
WBottom, Site 1 (fan)	Back Metal, Site 1	0.00	-4.62	-5.99	-5.07	-5.57	-6.75	-5.70	-7.20	-7.78	-9.47	-10.01	-11.88	—	—
	Back Metal, Site 2	0.00	-0.66	-0.63	0.92	2.32	0.79	-0.47	-1.87	-6.85	-7.69	-9.74	-9.96	—	—
Back Metal, Site 1	WTop, Site 1	0.00	0.73	-0.71	-0.69	-2.00	-2.55	-3.01	-4.50	-7.16	-7.97	-8.84	-10.12	-14.12	-19.92
	WTop, Site 2	0.00	-1.40	-0.35	-0.34	-0.99	-1.58	-3.85	-4.19	-5.45	-7.29	-8.93	-11.19	-14.47	-20.48
% decrease as compared to	WBottom, Site 1	0.00	0.00	0.00	0.70	-1.01	-1.29	-1.85	-3.99	-5.30	-6.26	-7.02	-9.99	-13.72	-20.00
	WBottom, Site 2	0.00	-0.74	-1.45	-2.13	-1.05	-2.00	-4.04	-4.70	-5.96	-7.43	-8.94	-10.56	-13.39	-21.32
Back Metal, Site 1 (fan)	Back Metal, Site 1	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	Back Metal, Site 2	0.00	1.03	2.07	-0.99	-1.27	-1.52	-1.73	-2.96	-4.45	-8.15	-9.49	-9.75	-15.09	-21.88

Chapter 12

Concluding Remarks

Although scaling remains a significant driving force in the development of CMOS integrated circuits, a paradigm shift to low power mobile computing incorporating multiple functionality into systems-on-chip has driven research efforts into alternate technologies that both continue scaling and satisfy power and computational goals. 3-D integration provides a path to incorporate disparate technologies while reducing the total power by decreasing the length of the global interconnects through the vertical placement of interdependent circuit blocks.

The integration of disparate technologies using 3-D integration provides unique research opportunities. Industry and academic research initiatives in integrating memory on processors have progressed rapidly over the past few years. In addition, a push to integrate emerging technologies such as optics, MEMS, carbon nanotubes, and graphene, into a 3-D stack is on-going. Although much research effort has focused on manufacturing these complex vertical systems, research on circuit design methodologies and techniques for synchronization, power delivery, and signaling has only just begun. In addition, there is a significant research focus on a specific problem exacerbated in 3-D ICs: The effect on

hot spot formation in stacked device layers not immediately in contact with a heat sink. The thermal ”problem” in 3-D ICs requires novel physical mitigation techniques as well as design techniques such as thermal-aware floorplanning or the placement of thermal TSVs to alleviate, or at least lessen, the effects of hot spots on 3-D circuit performance and reliability.

The research efforts described in this dissertation discuss various aspects of 3-D integrated circuit design. One primary research focus is on the characterization and modeling of the through silicon vias. The TSV is a critical technology for the formation of 3-D ICs. Proper characterization and modeling of the TSV is necessary to accurately design and analyze 3-D circuits. The TSV electrical impedance affects the design of multiple critical components in 3-D circuits including the clock network, affecting both skew and slew, the power network with greater impedance, and signaling with increased path loads. Characterization of the resistance, capacitance, and inductance of the TSV provides a basic understanding of the TSV electrical impedance with respect to the TSV diameter, length, dielectric thickness, distance to ground, substrate material, and distance to neighboring TSVs. After characterizing the TSV resistance, capacitance, and inductance across different parameters, closed-form expressions of the TSV electrical characteristics have been developed, providing a computationally efficient means to quickly and accurately determine the TSV impedance. The fundamental goal of this work is the development of an equivalent π-model of the TSV for insertion into CAD tools for accurate system-level simulation.

A critical component of 3-D systems is synchronization of data across device planes.

In the simplest context, homogeneous device planes, dies fabricated on different wafers but of the same technology are bonded together, and can potentially suffer from cross-plane synchronization errors due to wafer-to-wafer process variations. The more challenging synchronization scenario occurs when disparate technologies such as RF, MEMS, optics, CMOS, and mixed signal/analog are bonded to form a 3-D system. In this case, synchronization is affected not only by cross-plane process variations but also from widely varying synchronization criteria that can occur when integrating completely disparate technologies; for example, the integration of CMOS with optical devices or emerging technologies such as graphene. Proper modeling of the clock distribution network is therefore necessary to account for the disparity among integrated technologies. The clock models must also incorporate the equivalent impedance model of the TSV, as the TSV impedance can greatly affect the timing requirements of each device plane.

In addition to synchronizing disparate technologies, power delivery in 3-D systems provides a unique set of challenges that are similar to 2-D ICs but are more pronounced. As in current 2-D SoC ICs that require ten to twenty different voltages, 3-D ICs have the potential to integrate many device planes, requiring tens to hundreds of voltage domains. In addition, the current demands of each device plane can vary significantly, producing unintended IR and $L\frac{di}{dt}$ noise coupled between vertically stacked dies. Disparate device planes such as digital, RF, analog, mixed-signal, MEMS, and optics all require reliable and stable power supplies. The difficulty in distributing power in 3-D circuits arises due to the integration of independent power distribution networks with different impedance and load characteristics, all of which can be physically integrated within a single multi-plane

system. Similar to the clock network, the through silicon vias are critical components of these vertically integrated power networks. The TSV electrical characteristics must therefore be considered when determining the overall impedance of the power network. In addition, since a significant current can flow through the TSVs within the power network, the current capability of the TSV is of critical importance; therefore, design considerations such as redundant TSV placement on the power network must be considered.

Experimental test circuits and models of different power distribution topologies, as described in Chapter 7, provide insight into the effect of the power distribution topology on noise generation and propagation within 3-D systems. A test circuit has been designed in a 150 nm SOI process and fabricated by MIT Lincoln Laboratory. Two significant results of this work are: 1) the total number of TSVs affects the noise generation and propagation process within the power network, and 2) novel power network topologies such as dedicated power and ground planes, not readily available in 2-D ICs, can further reduce the average and peak noise on the power network. SPICE models of the power delivery networks have also been developed to analyze noise generation and propagation within the power network. The models provided added insight into the analysis of the peak noise voltage, voltage range, average noise voltage, and resonant frequency characteristics for both the power and ground networks.

A primary benefit of 3-D ICs is the potential to integrate disparate technologies into a stacked system with greater functionality, while not compromising yield. One potential application, described in Chapter 9, is the integration of a high speed optical device plane with CMOS circuits. A free-space optical interconnect (FSOI) system is integrated

with traditional CMOS devices into a stacked 3-D IC to provide a dedicated high bandwidth core-to-core pathway. The FSOI, based on vertical-cavity surface-emitting lasers, germanium based photodetectors, and microlenses, provides a high speed communication backbone for high bandwidth core-to-core signaling, while 3-D integration supports the placement of front end circuits and drivers for both the transmitter and receiver circuitry directly below the FSOI devices. Although the maximum bandwidth of the IC prototype does not exceed 3.5 GHz, an approximately 5.4 Gbps data rate, the 3-D FSOI test circuit demonstrates that it is possible to combine disparate technologies into a stacked system while providing enhanced functionality within a small form factor.

Thermal effects in 3-D ICs are a primary critical issue that can potentially constrain the vertical integration of the individual dies. The deleterious effects of hot spots on circuit operation in 2-D ICs have been evaluated in a 3-D test circuit. Heat degrades the maximum operating frequency of transistors while also increasing leakage current. Enhanced understanding of inter- and intra-plane thermal propagation, as described in Chapter 11, provides the groundwork for removing heat in stacked ICs. The experimental results on inter- and intra-plane thermal coupling indicate: 1) a substantial thermal effect for device planes farthest from the heat sink, 2) horizontal heat spreaders within a device plane are effective in lowering the peak temperature, and 3) the relative alignment of two active circuit blocks significantly affects hot spot formation. Novel design techniques including the dealignment of highly active vertical circuit blocks, and novel techniques such as active cooling through microfluidic channels are being developed to thermally manage 3-D integrated circuits.

The topics presented in this dissertation focus on critical systems components and applications of 3-D integrated circuits: synchronization, power delivery, TSV modeling, thermal effects, and heterogeneous systems integration are all discussed. Through experimental test circuits on noise propagation through 3-D IC power networks and thermal propagation, and models of clock topologies, power distribution networks, and the electrical impedance of the TSV, the work described in this dissertation chips away at the complexity of 3-D ICs. Although significant research efforts remain in many aspects of the 3-D IC design process, much groundwork has been laid through the results described in this dissertation as well as the work of many others in making 3-D integration a practical reality.

Chapter 13

Future Research

The emergence of TSV-based 3-D systems integration during the past decade as a potential technology candidate for beyond Moore devices and heterogeneous systems integration has lead to a plethora of manufacturing and design challenges. The standardization of TSV fabrication processes, including wafer and die bonding, wafer alignment, thin wafer handling, and TSV metallurgy, amongst other TSV manufacturing challenges has become an active research area. Research on 3-D circuits, physical design techniques, test methodologies, and methodologies for 3-D heterogeneous system integration has also been heavily pursued in both academia and industry. Distributing the clock signal and ensuring that each transistor in a 2-D circuit is properly powered requires a significant amount of computational and human resources. The complexity of the clock and power delivery systems in a 3-D stacked circuit is further complicated by each device plane often requiring different synchronization and power delivery capabilities, as described in Chapters 3, 6, and 7.

Clock and power delivery as well as interplane signaling require circuits at the interface between planes. These interface circuits are unique to 3-D systems as inter-die communication, power delivery, and synchronization between two or more dies from disparate

technologies are stacked into a tightly packed vertical system. Consider that two dies produced from completely disparate process technologies are vertically aligned and bonded to form a 3-D system. The packaging company performing the bonding and TSV fabrication no longer requires specific technology dependent knowledge such as the frequency of operation, device voltage, I/O requirements, and other characteristics of each process. Packaging houses that produce these stacked systems now have the potential to bond integrated circuits from two separate foundries with ”off-the-shelf” dies based on the needs of the customer. For such ”off-the-shelf” systems integration to occur, it is necessary to develop standard requirements for interfacing between device planes when circuit information is unavailable. In addition, industry wide general interface guidelines and standardization requirements are necessary to establish basic interface circuit properties such as interplane port locations, bus widths, interplane pin allocations, and other port characteristics. Once these basic guidelines are established, interface circuits for multiplane synchronization, power delivery, and signaling can be developed that support a ”plug-and-play” style of 3-D integration, where disparate technologies from different vendors can be combined to form a stacked vertical system within a common package.

Another area of research is the development of place and route algorithms that minimize the area of a 3-D circuit and the interconnect length among the cells. With 3-D integration, an issue arises in deciding whether two cells sharing a large number of interconnects require less area if closely placed within the same 2-D plane or on an adjacent physical plane. Placing circuit blocks on adjacent planes can often produce a shorter wire between two blocks. Placement algorithms must also consider the higher, more inductive

interconnect impedance of the 3-D vias used for interplane communication. More severe, however, is the threat posed by the large thermal gradients among planes within a 3-D stack. Thermal problems are not unique to 3-D integration. Elevated temperatures and hot spots within 2-D circuits can greatly degrade the performance and reliability of an integrated circuit [218]. Thermal analysis of 3-D ICs indicates that escalated temperatures are more problematic than in 2-D counterparts [219]. Peak temperatures within a 3-D system can exceed thermal limits of existing packaging technologies. Placement algorithms must consider these thermal gradients while reducing circuit area and interconnect length among the cells.

An issue related to circuit placement and routing while accounting for thermal effects is the location of the decoupling capacitors within a 3-D system. The placement of the on-chip decoupling capacitors in a 2-D integrated circuit is highly dependent upon the impedance of the power and ground lines connecting the capacitors and power supplies to the current loads. Although prior practice has been to place decoupling capacitors in the available free space, placing the decoupling capacitors irrespective of where the loads are located does not lower IR and $L \cdot di/dt$ noise while leaking current. This problem is further exacerbated in 3-D integrated circuits as placing both the decoupling capacitors and point-of-load power supplies not only affects the device plane the capacitors are located on, but can potentially affect the noise propagating to adjoining dies located both above and below that plane.

Several future research topics are described in this chapter for developing the necessary circuitry and design methodologies to interface between dies fabricated from disparate

technologies and integrated to form a 3-D IC. Research on interface circuits for interplane synchronization, power delivery, and signaling is discussed in Section 13.1.

13.1 Universal interface circuits for 3-D systems

Leveraging the benefits of traditional CMOS with emerging technologies such as nanowire or graphene based devices to form an integrated heterogeneous 3-D hybrid system requires novel methodologies and circuits at the interface between disparate device planes. One objective is to produce interface circuitry that is not unique to any particular technology, but is flexible and adaptive when merging different technologies. General purpose interface circuits are envisioned that will support a ”plug-and-play” approach to 3-D integration, where any combination of heterogeneous device planes are integrated through a general purpose 3-D interface. Circuit techniques can be developed that address multi-plane synchronization, power delivery, and signaling among currently available disparate technologies such as CMOS, RF, analog, and micro-/nano-electromechanical systems (MEMS/NEMS) as well as emerging technologies such as nano-FET and graphene-based device planes.

In addition, tools and methodologies for heterogeneous systems integration are important. These tools and methodologies will be needed to ensure functionality at the interface between vertically stacked device planes. Models to accurately describe signal propagation and power dissipation within 3-D heterogeneous systems, supporting system level optimization and performance analysis, are necessary. Optimal signaling techniques between planes are also important [408–411]. A breakdown of important research on interfacing

between disparate technologies in terms of synchronization, power delivery, and signaling is provided in the following subsections.

13.1.1 Interface circuits for synchronizing multiple device planes

Synchronization among emerging technologies such as nano-FETs and CMOS device planes is complicated by the difference in operating speeds between high speed devices which can operate at frequencies exceeding 200 GHz [412, 413], and much slower CMOS circuits. Integrating these devices with CMOS technologies operating between 5 to 10 GHz requires a two stage approach. The development of frequency multiplier and divider circuits to synchronize between high speed devices such as nano-FETs and the CMOS device layers is necessary. Once the interface circuits are available, system design issues that consider the large differences in clock frequency need to be addressed. In this system, frequency division is performed on the device plane operating at a frequency higher than D flip flop based frequency divider circuits. After frequency division, phase comparators combine with voltage controlled phase locked loops (PLL) to locally tune the frequencies [414]. Voltage controlled current regulators set the frequency of the PLL. The interface circuits on both the high speed device plane and the CMOS planes auto-tune to each local frequency at runtime.

In addition, several clock network topologies need to be explored, and both globally asynchronous, locally synchronous (GALS), and fully synchronous clock topologies need to be investigated. Extensive prior work exists on on-chip clocking [220,286,287,415,416]

and, more recently, in analyzing the clock skew and power consumption of 3-D clock distribution topologies [290, 291, 417]. Finally, design methodologies and techniques for synchronizing between two or more clock signal domains are required, providing guidelines for best design practices when merging technologies operating at different clock frequencies.

13.1.2 Interface circuits for power delivery in 3-D technology

In a heterogeneous system, multiple circuit domains require several reliable power supplies while providing sufficient and stable current. Building on the knowledge gained from a previous 3-D test circuit examining noise within power delivery networks [417, 418], multiple techniques to efficiently deliver power in 3-D integrated systems should be explored. 3-D power delivery must consider multi-voltage domains, different power noise requirements ($L \cdot di/dt$ and IR drops) between the heterogeneous device planes, and parasitic impedance mismatches within the power distribution networks, while ensuring that thermal and current density limits are not exceeded. In addition, built-in electrostatic discharge (ESD) protection at the interface between planes is essential to minimize damage to the interplane power network caused by static discharge during manufacturing and packaging. An interface is therefore required to ensure effective power delivery across the disparate technologies forming a stacked 3-D system.

Interface circuits for power delivery in 3-D systems should be developed that exploit techniques developed for 2-D circuits and packaging, such as voltage shifters and ESD protection. Two potential schemes to address multi-voltage device planes include; 1) novel

area efficient voltage converter circuits which support voltage tuning to match yet to be determined voltage requirements from black box IP circuits, and 2) dedicated power distribution networks with the corresponding interface connections for different multi-plane voltages. To ensure noise is not coupled from one power plane to another plane through the 3-D universal interface circuit, local decoupling capacitance should be included in the interface circuit to reduce transient $L \cdot di/dt$ noise. On-chip power supplies should internally generate the local voltages required by the different device planes within the heterogeneous 3-D system. A two level power delivery system is shown in Figure 13.1, where a dedicated device plane (middle plane in Figure 13.1) with large and efficient power supplies, such as buck converters [419, 420], convert the off-chip voltage to a lower voltage [198, 310, 421] determined by voltage detection circuits for each device plane and within each power domain. This lower voltage is distributed to the small point-of-load voltage regulators within the global power distribution network on each device plane, where the point-of-load regulators locally generate the required voltages. This method should support different types of on-chip voltage regulators (low-dropout regulators, switched capacitor regulators, buck converters, and hybrid voltage regulators) and decoupling capacitors (PIP capacitors, MOS capacitors, and MIM capacitors) as the noise regulation characteristics can vary significantly for the different technologies within a 3-D stack.

Standardization protocols to properly integrate disparate technologies from different IC fabrication houses should also be investigated. These standards address the physical constraints between the power and ground ports necessary for 3-D systems integration, and set guidelines to determine the optimal number of ports necessary to satisfy the current density

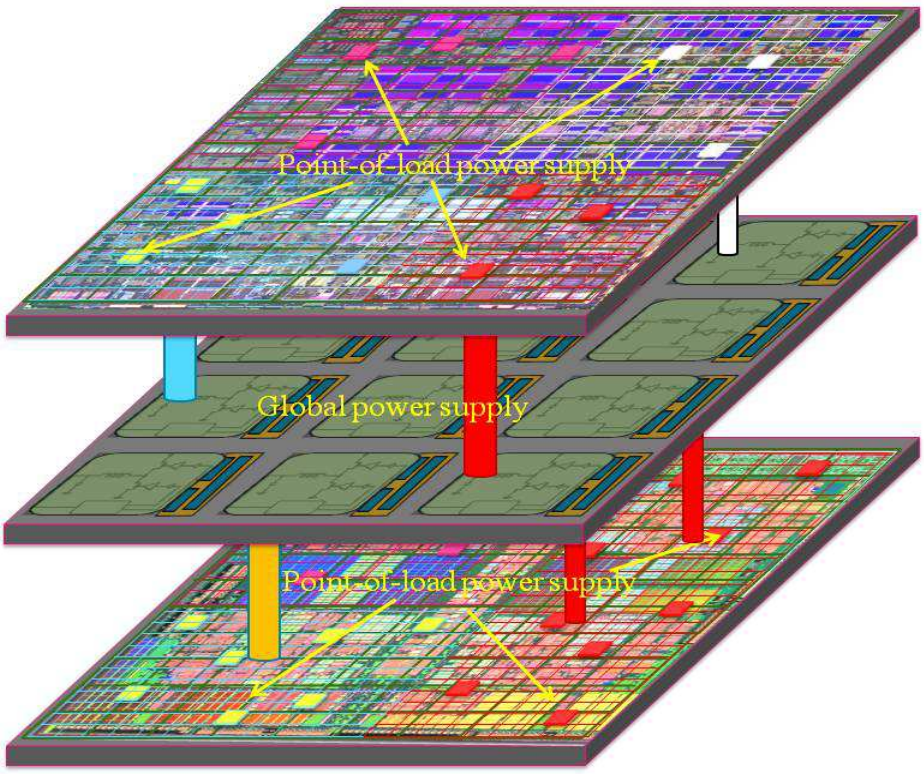


Figure 13.1: Distributed power delivery system with a dedicated global power supply plane and point-of-load power supplies for each device plane of a heterogeneous 3-D system.

requirements of each device plane. Standardized port circuitry guarantee interchangeability and interoperability among planes of different technologies. These circuits include voltage detection circuits that automatically determine the voltage levels between the power generation plane and the device planes, adjustable power supplies, and control circuitry to set the power supply voltages based on the detected voltages.

Existing circuits and methodologies for 2-D power delivery manage the interface between disparate device planes within a 3-D stack. The circuits should include tunable voltage converters for power delivery, voltage level shifters, and sense circuits to self-detect the voltage requirements of each device plane at runtime. Methodologies to determine the proper topology for interplane power delivery, ESD protection, and decoupling capacitance at each interface should also be developed.

13.1.3 Interface circuits for interplane signaling

A primary requirement of 3-D integrated systems is plane-to-plane signal communication between device planes. Circuits necessary to ensure proper signaling between planes require conversion across a wide range of voltages. The voltage levels should be determined at run time by voltage sensing circuits similar to those described in the previous section. Buffering and ESD protection circuits to ensure proper signal behavior should also be developed. A schematic representation of a group of circuits including level shifters, voltage detection circuits, and ESD protection is shown in Figure 13.2. A second research objective is to examine the use of nanowire-based, graphene-based, and other non-standard

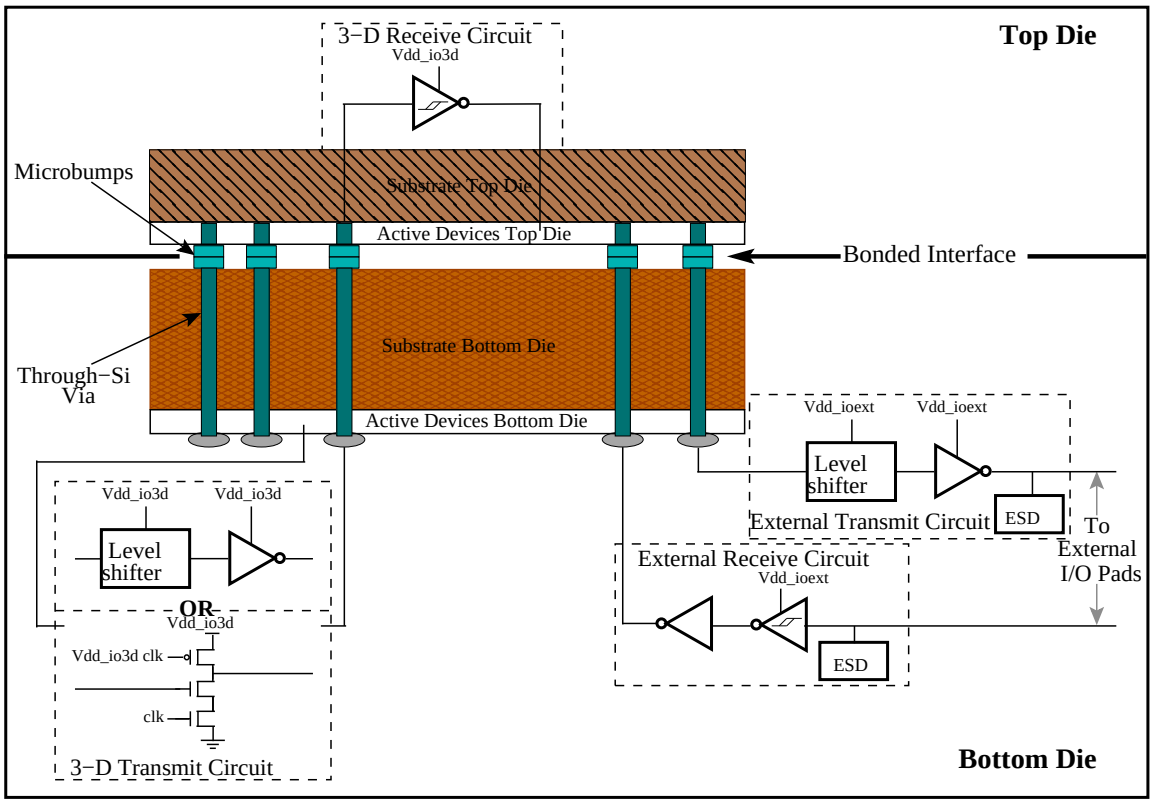


Figure 13.2: Interface circuits for interplane signaling that include level shifters, buffers, and ESD protection [422].

high speed materials behaving as an interposer plane to produce high speed, high bandwidth dedicated routing for point-to-point communication. The objective is to determine how best to utilize the interposer plane within a 3-D integrated circuit. The interposer layer will adhere to the basic design objective, the formation of a general interface between heterogeneous technologies for plug-and-play 3-D integration.

Algorithms and methodologies to configure the high speed interposer are necessary. Two potential routing methodologies should be considered, a point-to-point routing algorithm that minimizes crossbar arbitration, and a topology similar to a network-on-chip (NoC) [423], where the metal links between the router nodes are replaced with nanowires or

other high speed materials. Methodologies that address cross-die placement of the switches and links are required. Arbitration algorithms [424, 425] should be developed since data may attempt to occupy the same nanowire link. Based on the effects of the crossbar topology and the design of the router circuitry, modifications to arbitration and congestion algorithms are needed to ensure high speed interplane communication with fewer data collisions.

In the likely case that a device plane with high speed transistors (e.g., GaAs, nano-FETs) is stacked with much slower CMOS devices, architectural level techniques should be developed to ensure high throughput computation. The interface between these device planes will require a mechanism to both queue data and, in high speed computational systems, properly serialize multiple data paths for computation by the nano-FETs. Once the computation is performed, de-serialization of the data will interface with the CMOS logic. Methodologies should consider thermal effects between multiple stacked dies. The use of nanowires and other circuit techniques as thermal conduits to heat sinks should be considered, and, based on the resulting analysis, design methodologies for interplane signaling would be accordingly modified. Relocating buffers from hot spot locations to more thermally stable areas, for example, is a 2-D circuit technique [426, 427] that is also applicable to 3-D circuits [248, 428–430]. Many of these algorithms and methodologies are also applicable to the clock signal.

Standardization protocols to properly integrate disparate technologies from different IC fabrication houses should be investigated. These standards address the physical constraints

between signaling ports necessary for 3-D systems integration, and set guidelines to determine the optimal number of ports necessary to satisfy the interplane signaling bandwidth requirements of each device plane. Standardized port circuitry guarantee interchangeability and interoperability among planes of different technologies. These circuits include voltage detection circuits that automatically determine the voltage levels between the different device planes and control circuitry to set the power supply voltage based on the detected voltages.

13.2 Summary

As 3-D integrated circuits have evolved over the past ten years, the number of open research topics for circuit level design techniques and methodologies have significantly expanded. Future research initiatives related to these circuit techniques focus on the interface between device planes. More specifically, three possible research topics are suggested: 1) interface circuits and methodologies for interplane synchronization, 2) interface circuits and methodologies for interplane power delivery, and 3) interface circuits and methodologies for interplane signaling.

Bibliography

[1] J. Bardeen and W. H. Brattain, "The Transistor, a Semiconductor Triode," *Physical Review*, Vol. 74, No. 2, pp. 230 - 231, July 1948.

[2] W. Shockley, "The Theory of PN Junctions in Semiconductos and PN Junction Transistors," *Bell System Technical Journal*, Vol. 28, No. 3, pp. 435 - 489, July 1949.

[3] W. F. Brinkman, D. E. Haggan, and W. W. Troutman, "A History of the Invention of the Transistor and Where It Will Lead Us," *IEEE Journal of Solid-State Circuits*, Vol. 32, No. 12, pp. 1858 - 1865, December 1997.

[4] C. Weiner, "How the Transistor Emerged," *IEEE Spectrum*, Vol. 10, No. 1, pp. 24 - 33, January 1973.

[5] C. Pirtle, *Engineering the World: Stories from the First 75 Years of Texas Instruments*, Southern Methodist University Press, 2005.

[6] T. R. Reid, *The Chip: How Two Americans Invented the Microchip and Launched a Revolution*, Random House, Inc., 2001.

[7] L. Berlin, *Man Behind the Microchip: Robert Noyce and the Invention of Silicon Valley*, Oxford University Press, 2005.

[8] N. Noyce, "Microelectronics," *Scientific American*, Vol. 237, No. 3, pp. 62-69, September 1977.

[9] P. H. Abelson and A. L. Hammond, "The Electronics Revolution," *Science*, Vol. 195, No. 4283, pp. 1087 - 1091, March 1977.

[10] G. E. Moore, "The Future of Integrated Electronics," *Fairchild Semiconductor Internal Publication*, pp. 1-18, 1964.

[11] G. E. Moore, "Cramming More Components onto Integrated Circuits," *Electronics Magazine*, Vol. 38, No. 8, pp. 114-117, April 1965.

[12] G. E. Moore, "Cramming More Components onto Integrated Circuits," *Proceedings of the IEEE*, Vol. 86, No. 1, pp. 82-85, January 1998.

[13] G. E. Moore, "Progress in Digital Integrated Electronics," *Understanding Moore's Law: Four Decades of Innovation*, Chapter 6, pp. 61–66, 2006.

[14] C. D. Brady, "Integrated Circuit Packages for the AGC," *NASA Technical Report*, pp. 1–5, June 1964.

[15] L. Xue, C. Liu, H.-S. Kim, S. Kim, and S. Tiwari, "Three-Dimensional Integration: Technology, Use, and Issues for Mixed-Signal Applications," *IEEE Transactions On Electron Devices*, Vol. 50, No. 3, pp. 601-608, March 2003.

[16] W. R. Davis, J. Wilson, S. Mick, J. Xu, H. Hua, C. Mineo, A. M. Sule, M. Steer, and P. D. Franzon, "Demystifying 3D ICs: The Pros and Cons of Going Vertical," *IEEE Design and Test of Computers*, Vol. 22, No. 6, pp. 498–510, 2005.

[17] V. F. Pavlidis and E. G. Friedman, *Three-Dimensional Integrated Circuit Design*, Morgan Kaufmann, 2009.

[18] E. C. Hall, "Integrated Circuits in the Apollo Guidance Computer," *Apollo The International Magazine Of Art And Antiques*, pp. 1–24, 1963.

[19] E. C. Hall, "A Case History of the AGC Integrated Logic Circuit," *Apollo The International Magazine Of Art And Antiques*, pp. 1–40, December 1965.

[20] L. D. Hanley, J. Partridge, and E. C. Hall, "The Application of Failure Analysis in Procuring and Screening of Integrated Circuits," *Proceedings of the Fourth Annual Symposium on the Physics of Failure in Electronics*, pp. 95–139, November 1965.

[21] R. N. Noyce and M. E. Hoff Jr., "A History of Microprocessor Development at Intel," *IEEE Micro*, Vol. 1, No. 1, pp. 8–21, February 1981.

[22] *4004: Single Chip 4-Bit P-Channel Microprocessor*, Intel Corporation, March 1987.

[23] *Intel 4004, the First CPU, is 40 Years Old Today*, Extremetech, November 2011. [Online]. Available: http://www.extremetech.com/computing/105029-intel-4004-the-first-cpu-is-40-years-old-today

[24] *Happy 40th Birthday, Intel 4004! The First of the Bricks that Built the IT World*, The Register, November 2011. [Online]. Available: http://www.theregister.co.uk/2011/11/15/the_first_forty_years_of_intel_microprocessors/print.html

[25] *Intel 4004 Image*, The Register. [Online]. Available: http://regmedia.co.uk/2011/11/14/4004_die_large.jpg

[26] *8008: 8-Bit Parallel Central Processor Unit*, Intel Corporation, November 1972.

[27] *Intel 8008 Image*, The Register. [Online]. Available: http://regmedia.co.uk/2011/11/14/8008_die_large.jpg

[28] *8080A/8080A-1/8080A-2: 8-Bit N-Channel Microprocessor*, Intel Corporation, November 1986.

[29] *8085AH/8085AH-2/8085AH-1: 8-Bit HMOS Microprocessors*, Intel Corporation, September 1987.

[30] *8086: 16-Bit HMOS Microprocessors*, Intel Corporation, September 1990.

[31] *80C286/883: High Performance Microprocessor with Memory Management and Protection*, Intel Corporation, September 1988.

[32] *Intel386 EX Embedded Microprocessor*, Intel Corporation, October 1998.

[33] *Intel 80386 Image*, The Register. [Online]. Available: http://regmedia.co.uk/2011/11/14/80386_die_large.jpg

[34] *Embedded Intel486 SX Processor*, Intel Corporation, August 2004.

[35] *Pentium Processor at iCOMP Intex 610-75MHz*, Intel Corporation, June 1997.

[36] *Pentium Pro Processor with 1 MB L2 Cache at 200 MHz*, Intel Corporation, January 1998.

[37] *Pentium II Processor at 233MHz, 266MHz, 300MHz and 333MHz*, Intel Corporation, January 1998.

[38] *Pentium III Processor for the SC242 at 450MHz to 1.0GHz*, Intel Corporation, July 2002.

[39] *Pentium IV Processor in the 423-Pin Package at 1.30, 1.40, 1.50, 1.60, 1.70, 1.80, 1.90, 2GHz*, Intel Corporation, August 2001.

[40] *Pentium IV Processor on 90nm Process*, Intel Corporation, February 2005.

[41] T. López, R. Elferich, and E. Alarcón, *Voltage Regulators for Next Generation Microprocessors*, Springer, 2010.

[42] *Intel Core i5-600, i3-500 Desktop Processor Series, Intel Pentium Desktop Processor 6000 Series*, Intel Corporation, January 2011.

[43] *Intel Core i7-900 Desktop Processor Extreme Edition Series on 32-nm Process*, Intel Corporation, June 2010.

[44] *Intel Pentium II Image*, The Register. [Online]. Available: http://regmedia.co.uk/2011/11/14/pentium_ii_die_large.jpg

[45] *IBM Announces Supercomputer to Propel Sciences Forward; Blue Gene/Q Tackles Major Challenges of the Day, Delivers up to 100 Petaflops at Peak Performance*, International Business Machine, November 2011. [Online]. Available: http://www-03.ibm.com/press/us/en/pressrelease/35990.wss

[46] *IBM System Blue Gene/Q Data Sheet*, International Business Machine, November 2011. [Online]. Available: public.dhe.ibm.com/common/ssi/ecm/en/dcd12345usen/DCD12345USEN.PDF

[47] *IBM's BlueGene/Q Super Chip Grows 18th Core; It's Nice to Have a Spare*, The Register, August 2011. [Online]. Available: http://www.theregister.co.uk/2011/08/22/ibm_bluegene_q_chip/

[48] J.-Q. Lu, "3-D Hyperintegration and Packaging Technologies for Micro-Nano Systems," *Proceedings of the IEEE*, Vol. 97, No. 1, pp. 18–30, January 2009.

[49] C. R. King, D. Sekar, M. S. Bakir, B. Dang, J. Pikarsky, and J. D. Meindl, "3D Stacking of Chips with Electrical and Microfluidic I/O Interconnects," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1–7, May 2008.

[50] T. Zhang, R. Micheloni, G. Zhang, Z. R. Huang, and J. J.-Q. Lu, "3-D Data Storage, Power Delivery, and RF/Optical Transceiver—Case Studies of 3-D Integration from System Design Perspectives," *Proceedings of the IEEE*, Vol. 97, No. 1, pp. 161–174, January 2009.

[51] G. H. Ebel, J. A. Jeffery, and J. P. Farrell, "Wirebonding Reliability Techniques and Analysis," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. CHMT-5, No. 4, pp. 441–445, December 1982.

[52] *Technical Support - Professional Services; Chip on Board Assembly Process*, Beijing Unite Rayleigh Co., Ltd, 2010. [Online]. Available: http://www.ur-bj.com/tech/cob.html

[53] F. Wang, Y. Chen, and L. Han, "Ultrasonic Vibration at Thermosonic Flip-Chip Bonding Interface," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, Vol. 1, No. 6, pp. 852–858, June 2011.

[54] S.-Y. Kang, P. M. Williams, and Y.-C. Lee, "Modeling and Experimental Studies on Thermosonic Flip-Chip Bonding," *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B*, Vol. 18, No. 4, pp. 728–733, November 1995.

[55] Q. Tan, W. Zhang, B. Schaible, L. J. Bond, T.-H. Ju, and Y.-C. Lee, "Thermosonic Flip-Chip Bonding Using Longitudinal Ultrasonic Vibration," *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B*, Vol. 21, No. 1, pp. 53–58, February 1998.

[56] *Flip Chip Bonding*, Microbonding S.A., August 2005. [Online]. Available: http://www.microbonding.com/gb/fc_gb.htm

[57] S. A. Gee and C. G. M. van Kessel, "Stiffness and Yielding in the PLCC "J" Lead," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. CHMT-10, No. 3, pp. 379–390, September 1987.

[58] R. Kotlowitz, "Comparative Compliance of Representative Lead Designs for Surface-Mounted Components," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. 12, No. 4, pp. 431–448, December 1989.

[59] C. G. Schmidt, J. W. Simons, C. H. Kanazawa, and D. C. Erlich, "Thermal Fatigue Behavior of J-Lead Solder Joints," *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part A*, Vol. 18, No. 3, pp. 611–617, September 1995.

[60] *Advanced Packaging*, Fujitsu Semiconductor Limited. [Online]. Available: http://www.fujitsu.com/us/semiconductors/sms/packaging/index.html

[61] R. J. Drost, R. D. Hopkins, R. Ho, and I. E. Sutherland, "Proximity Communication," *IEEE Journal of Solid-State Circuits*, Vol. 39, No. 9, pp. 1529–1535, September 2004.

[62] J. Shi, D. Popovic, N. Nettleton, T. Sze, D. Douglas, H. Thacker, J. Cunningham, K. Furuta, R. Kojima, K. Hirose, and K. Hwang, "Direct Chip Powering and Enhancement of Proximity Communication through Anisotropic Conductive Adhesive Chip-to-Chip Bonding," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 363–368, June 2010.

[63] A. Chow, D. Hopkins, R. Drost, and R. Ho, "Enabling Technologies for Multi-Chip Integration Using Proximity Communication," *Proceedings of the IEEE International Symposium on VLSI Design, Automation and Test*, pp. 39–42, April 2009.

[64] A. V. Krishnamoorthy, J. E. Cunningham, X. Zheng, I. Shubin, J. Simons, D. Feng, H. Liang, C.-C. Kung, and M. Asghari, "Optical Proximity Communication With Passively Aligned Silicon Photonic Chips," *IEEE Journal of Quantum Electronics*, Vol. 45, No. 4, pp. 409–414, April 2009.

[65] A. Majumdar, J. E. Cunningham, and A. V. Krishnamoorthy, "Alignment and Performance Considerations for Capacitive, Inductive, and Optical Proximity Communication," *IEEE Transactions on Advanced Packaging*, Vol. 33, No. 3, pp. 690–701, August 2010.

[66] Q. Liu, P. Fay, and G. H. Bernstein, "A Novel Scheme for Wide Bandwidth Chip-to-Chip Communications," *Journal of Microelectronics and Electronic Packaging*, Vol. 4, No. 1, pp. 1–7, 2007.

[67] L. Smith, *Package on Package (PoP) Applications, Requirements, Infrastructure, and Technologies*, Amkor Technology, Inc, March 2011. [Online]. Available: http://www.smta.org/chapters/files/Oregon_SMTA_OR_March_16_LSMIT.pdf

[68] V. Solberg, *Addressing Product Miniaturization Challenges for Medical Electronics*, ECN Magazine, May 2011. [Online]. Available: http://www.ecnmag.com/Articles/2011/05/Hardware-Corner/Addressing-Product-Miniaturization-Challenges/

[69] G. T. Goeloe, E. W. Maby, D. J. Silversmith, R. W. Mountain, and D. A. Antoniadis, "Vertical Single-Gate CMOS Inverters on Laser-Processed Multilayer Substrates," *Proceedings of the IEEE International Electron Devices Meeting*, Vol. 27, pp. 554-556, Decemeber 1981.

[70] J. F. Gibbons and K. F. Lee, "One-Gate-Wide CMOS Inverter on Laser-Recrystallized Polysilicon," *IEEE Electron Device Letters*, Vol. EDL-1, No. 6, pp. 117-118, June 1980.

[71] R. P. Zingg and B. Hofflinger, "Stacked CMOS Inverter with Symmetric Device Performance," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 909–911, December 1989.

[72] P. Batude, M. Vinet, A. Pouydebasque, C. L. Royer, B. Previtali, C. Tabone, J. Hartmann, L. Sanchez, L. Baud, V. Carron, A. Toffoli, F. Allain, V. Mazzocchi, D. Lafond, S. Deleonibus, and O. Faynot, "3D Monolithic Integration," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2233–2236, May 2011.

[73] J. W. Joyner, P. Zarkesh-Ha, J. A. Davis, and J. D. Meindl, "A Three-Dimensional Stochastic Wire-Length Distribution for Variable Separation of Strata," *Proceedings of the IEEE International Interconnect Technology Conference*, pp. 126-128, June 2000.

[74] N. Saputra, M. Danesh, A. Baiano, R. Ishihara, J. R. Long, N. Karaki, and S. Inoue, "An Assessment of μ-Czochralski, Single-Grain Silicon Thin-Film Transistor

Technology for Large-Area, Sensor and 3-D Electronic Integration," *IEEE Journal of Solid-State Circuits*, Vol. 43, No. 7, pp. 1563–1576, July 2008.

[75] J. Derakhshandeh, N. Golshani, R. Ishihara, M. Reza, T. Mofrad, M. Robertson, T. Morrison, and C. I. M. Beenakker, "Monolithic 3-D Integration of SRAM and Image Sensor Using Two Layers of Single-Grain Silicon," *IEEE Transactions on Electron Devices*, Vol. 58, No. 11, pp. 3954–3961, November 2011.

[76] M. Koyanagi, H. Kurino, K. W. Lee, K. Sakuma, N. Miyakawa, and H. Itani, "Future System-on-Silicon LSI Chips," *IEEE Micro*, Vol. 18, No. 4, pp. 17-22, July/August 1998.

[77] V. K. Jain, S. Bhanja, G. H. Chapman, and L. Doddannagari, "A Highly Reconfigurable Computing Array: DSP Plane of a 3D Heterogeneous SoC," *Proceedings of the IEEE International Systems on Chip Conference*, pp. 243-246, September 2005.

[78] A. Papanikolaou, D. Soudris, and R. Radojcic, Eds., *Three-Dimensional System Integration: IC Stacking Process and Design*, Springer Science + Business Media, Inc., 2011.

[79] K.-W. Lee, A. Noriki, K. Kiyoyama, and T. Fukushima, "Three-Dimensional Hybrid Integration Technology of CMOS, MEMS, and Photonics Circuits for Optoelectronic Heterogeneous Integrated Systems," *IEEE Transactions on Electron*, Vol. 58, No. 3, pp. 748–757, March 2011.

[80] M. S. Bakir and J. D. Meindl, *Integrated Interconnect Technologies for 3-D Nanoelectronic Systems*, Artech House, Inc., 2009.

[81] R. H. Dennard, F. H. Gaensslen, V. L. Rideout, E. Bassous, and A. R. Leblanc, "Design Of Ion-Implanted MOSFET's with Very Small Physical Dimensions," *Proceedings of the IEEE*, Vol. 87, No. 4, pp. 668–678, April 1999.

[82] *MITLL Low-Power FDSOI CMOS Process Design Guide*, MIT Lincoln Laboratories, September 2008.

[83] J. A. Burns, B. F. Aull, C. K. Chen, C.-L. Chen, C. L. Keast, J. M. Knecht, V. Suntharalingam, K. Warner, P. W. Wyatt, and D.-R. W. Yost, "A Wafer-Scale 3-D Circuit Integration Technology," *IEEE Transactions on Electron Devices*, Vol. 53, No. 10, pp. 2507-2515, October 2006.

[84] S. William, "Semiconductive Wafer and Method of Making the Same," July 1962.

[85] M. G. Smith and S. Emanuel, "Methods of Making Thru-Connections in Semiconductor Wafers," September 1967.

[86] P. S. Andry, C. K. Tsang, B. C. Webb, E. J. Sprogis, S. L. Wright, B. Dang, and D. G. Manzer, "Fabrication and Characterization of Robust Through-Silicon Vias for Silicon-Carrier Applications," *IBM Journal of Research and Development*, Vol. 52, No. 6, pp. 571–581, November 2008.

[87] S. F. Al-sarawi, D. Abbott, and P. D. Franzon, "A Review of 3-D Packaging Technology," *IEEE Transactions on Components, Packaging, and Manufacturing Technology–Part B*, Vol. 21, No. 1, pp. 2-14, February 2006.

[88] P. Leduc, L. Di Cioccio, B. Charlet, M. Rousseau, M. Assous, D. Bouchu, A. Roule, M. Zussy, P. Gueguen, A. Roman, O. Rozeau, M. Heitzmann, J.-P. Nieto, L. Vandroux, P.-H. Haumesser, R. Quenouillere, A. Toffoli, P. Sixt, S. Maitrejean, L. Clavelier, and N. Sillon, "Enabling Technologies for 3D Chip Stacking," *Proceedings of the IEEE International Symposium on VLSI Technology, Systems and Applications*, pp. 76-78, April 2008.

[89] *MITLL Low-Power FDSOI CMOS Process Design Guide*, MIT Lincoln Laboratories, June 2006.

[90] F. von Trapp, "3D Integration: A Progress Report," *Semiconductor Equipment and Materials International*, Technical Report, 2009.

[91] Z. Xu and J.-Q. Lu, "Through-Silicon-Via Fabrication Technologies, Passives Extraction, and Electrical Modeling for 3-D Integration/Packaging," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 26, No. 1, pp. 23-34, February 2013.

[92] M. Koyanagi, T. Nakamura, Y. Yamada, H. Kikuchi, T. Fukushima, T. Tanaka, and H. Kurino, "Three-Dimensional Integration Technology Based on Wafer Bonding With Vertical Buried Interconnections," *IEEE Transactions on Electron Devices*, Vol. 53, No. 11, pp. 2799-2808, November 2006.

[93] M. Koyanagi, T. Fukushima, and T. Tanaka, "High-Density Through Silicon Vias for 3-D LSIs," *Proceedings of the IEEE*, Vol. 97, No. 1, pp. 49–59, January 2009.

[94] M. Kawano, N. Takahashi, Y. Kurita, K. Soejima, M. Komuro, and S. Matsui, "Three-Dimensional Packaging Technology for Stacked DRAM With 3-Gb/s Data

Transfer," *IEEE Transactions on Electron Devices*, Vol. 55, No. 7, pp. 1614–1620, July 2008.

[95] J. Bea, T. Fukushima, T. Tanaka, and M. Koyanagi, "Evaluatoin of Cu Diffusion From Cu Through-Silicon Via (TSV) in Three-Dimensional LSI by Transient Capacitance Measurement," *IEEE Transactions on Electron Devices Letters*, Vol. 32, No. 7, pp. 940-942, July 2011.

[96] K. A. Jenkins and C. S. Patel, "Copper-Filled Through Wafer Vias with Very Low Inductance," *Proceedings of the IEEE International Interconnect Technology Conference*, pp. 144-146, June 2005.

[97] C. Okoro, R. Agarwal, P. Limaye, B. Vandevelde, D. Vandepitte, and E. Beyne, "Insertion Bonding: A Novel Cu-Cu Bonding Approach for 3D Integration," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1370–1375, June 2010.

[98] Y. Liang and Y. Li, "Closed-Form Expressions for the Resistance and the Inductance of Different Profiles of Through-Silicon Vias," *IEEE Transactions on Electron Device Letters*, Vol. 32, No. 3, pp. 393-395, March 2011.

[99] *On the Road towards TSV Manufacturing*, Semi - European 3D TSV Summit, January 2013, courtesy of Fraunhofer-IZM. [Online]. Available: http://www.semi.org/eu/node/8206

[100] *Applications - Through Silicon Via (TSV)*, Tango Systems, Inc. [Online]. Available: http://www.tangosystemsinc.com/apps_TSV.html

[101] J. U. Knickerbocker, P. S. Andry, B. Dang, R. R. Horton, M. J. Interrante, C. S. Patel, R. J. Polastre, K. Sakuma, R. Sirdeshmukh, E. J. Sprogis, S. M. Sri-Jayantha, a. M. Stephens, a. W. Topol, C. K. Tsang, B. C. Webb, and S. L. Wright, "Three-Dimensional Silicon Integration," *IBM Journal of Research and Development*, Vol. 52, No. 6, pp. 553–569, November 2008.

[102] J. Knickerbocker, P. Andry, B. Dang, R. Horton, C. S. Patel, R. Polastre, K. Sakuma, E. Sprogis, C. Tsang, B. Webb, and S. Wright, "3D Silicon Integration," *Proceedings of the IEEE International Electronic Components and Technology Conference*, pp. 538–543, May 2008.

[103] P. Dixit and J. Miao, "Fabrication of High Aspect Ratio 35 μm Pitch Interconnects for Next Generation 3-D WLP by Through-Wafer Copper Electroplating," *Proceedings of the IEEE International Electronic Components and Technology Conference*, pp. 388–393, July 2006.

[104] Y. Civale, B. Majeed, D. S. Tezcan, P. Soussan, and E. Beyne, "Spin-on Dielectric Liner TSV for 3D Wafer Level Packaging Applications," *Proceeding of the IEEE International Interconnect Technology Conference*, pp. 3–5, June 2010.

[105] M. Koyanagi, T. Nakamura, Y. Yamada, H. Kikuchi, T. Fukushima, T. Tanaka, and H. Kurino, "Three-Dimensional Integration Technology Based on Wafer Bonding With Vertical Buried Interconnections," *IEEE Transactions on Electron Devices*, Vol. 53, No. 11, pp. 2799–2808, November 2006.

[106] K. Y. Au, S. L. Kriangsak, X. R. Zhang, W. H. Zhu, and C. H. Toh, "3D Chip Stacking & Reliability Using TSV-Micro C4 Solder Interconnection," *Proceedings of the IEEE International Electronic Components and Technology Conference*, pp. 1376–1384, June 2010.

[107] N. Sillon, a. Astier, H. Boutry, L. Di Cioccio, D. Henry, and P. Leduc, "Enabling Technologies for 3D Integration: From Packaging Miniaturization to Advanced Stacked ICs," *Proceedings of the IEEE International Electron Devices Meeting*, No. 1, pp. 1–4, December 2008.

[108] J. Burns, B. Aull, C. Chen, C. Keast, J. Knecht, V. Suntharalingam, K. Warner, P. Wyatt, and D.-R. Yost, "A Wafer-Scale 3-D Circuit Integration Technology," *IEEE Transactions on Electron Devices*, Vol. 53, No. 10, pp. 2507–2516, October 2006.

[109] H. Kitada, N. Maeda, K. Fujimoto, Y. Mizushima, Y. Nakata, T. Nakamura, W. Lee, Y. S. Kwon, and T. Ohba, "Development of Low Temperature Dielectrics Down to 150°C for Multiple TSVs Structure with Wafer-on-Wafer (WOW) Technology," *Proceedings of the IEEE International Interconnect Technology Conference*, pp. 1-3, June 2009.

[110] Y. P. R. Lamy, K. B. Jinesh, F. Roozeboom, D. J. Gravesteijn, and W. F. A. Besling, "RF Characterization and Analytical Modelling of Through Silicon Vias and Coplanar Waveguides for 3D Integration," *IEEE Transactions on Advanced Packaging*, Vol. 33, No. 4, pp. 1072-1079, November 2010.

[111] S. Arkalgud, "Stacking the Chips," *SEMATECH/ISMI Symposium*, 2009.

[112] G. Katti, M. Stucchi, K. De Meyer, and W. Dehaene, "Electrical Modeling and Characterization of Through Silicon Via for Three-Dimensional ICs," *IEEE Transactions on Electron Devices*, Vol. 57, No. 1, pp. 256–262, January 2010.

[113] *ANSYS Q3D Extractor*, ANSYS. [Online]. Available: http://www.ansys.com/Products/Simulation+Technology/Electromagnetics/High-Performance+Electronic+Design/ANSYS+Q3D+Extractor

[114] *ANSYS HFSS*, ANSYS. [Online]. Available: http://www.ansys.com/Products/Simulation+Technology/Electromagnetics/High-Performance+Electronic+Design/ANSYS+HFSS

[115] M. Ravenstahl and M. Kopp, *Application Brief: ANSYS HFSS for ECAD*, ANSYS, 2013.

[116] *ANSYS HFSS Brochure*, ANSYS, 2011.

[117] *EIP Electromagnetic Field Solver Suite of Tools*, IBM, March 2010. [Online]. Available: http://www.alphaworks.ibm.com/tech/eip

[118] M. Kamon, M. J. Tsuk, and J. K. White, "FastHenry: A Multipole-Accelerated 3-D Inductance Extraction Program," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 42, No. 9, pp. 1750-1758, September 1994.

[119] K. Nabors and J. White, "FastCap: A Multipole-Accelerated 3-D Capacitance Extraction Program," *IEEE Transactions on Computer-Aided Design of Intergrated Circuits and Systems*, Vol. 10, No. 11, pp. 1447-1459, November 1991.

[120] *EMPro 3D EM Simulation Software*, Agilent Technologies. [Online]. Available: http://www.home.agilent.com/en/pc-1297143/empro-3d-em-simulation-software

[121] *Agilent EEs of EDA EMPro*, Agilent Technologies, February 2012.

[122] *Sonnet Suites: High Frequency Electromagnetic Software*, Sonnet. [Online]. Available: http://www.sonnetsoftware.com/products/sonnet-suites

[123] *Going With the Flow: Sonnet Professional Suite Release 14*, Sonnet, 2012.

[124] *3D Electromagnetic Simulation Software*, CST. [Online]. Available: http://www.cst.com

[125] S. W. Ho, S. W. Yoon, Q. Zhou, K. Pasad, V. Kripesh, and J. H. Lau, "High RF Performance TSV Silicon Carrier for High Frequency Application," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1946–1952, May 2008.

[126] J. H. Wu, *Through-Substrate Interconnects for 3-D Integration and RF Systems*, Ph.D. Dissertation, Massachusetts Institute of Technology, October 2006.

[127] C. Xu, R. Suaya, and K. Banerjee, "Compact Modeling and Analysis of Through-Si-Via-Induced Electrical Noise Coupling in Three-Dimensional ICs," *IEEE Transactions on Electron Devices*, Vol. 58, No. 11, pp. 4024-4034, November 2011.

[128] P. Schneider, S. Reitz, A. Wilde, G. Elst, and P. Schwarz, "Towards a Methodology for Analysis of Interconnect Structures for 3D-Integration of Micro Systems," *Analog Integrated Circuits and Signal Processing*, Vol. 57, No. 3, pp. 205–211, February 2008.

[129] I. Savidis, S. M. Alam, A. Jain, S. Pozder, R. E. Jones, and R. Chatterjee, "Electrical Modeling and Characterization of Through-Silicon Vias (TSVs) for 3-D Integrated Circuits," *Microelectronics Journal*, Vol. 41, No. 1, pp. 9–16, January 2010.

[130] G. Katti, M. Stucchi, J. Van Olmen, K. De Meyer, and W. Dehaene, "Through-Silicon-Via Capacitance Reduction Technique to Benefit 3-D IC Performance," *IEEE Electron Device Letters*, Vol. 31, No. 6, pp. 549–551, June 2010.

[131] R. Weerasekera, M. Grange, D. Pamunuwa, H. Tenhunen, and L.-R. Zheng, "Compact Modelling of Through-Silicon Vias (TSVs) in Three-Dimensional (3-D) Integrated Circuits," *Proceedings of the IEEE International Conference on 3D System Integration*, pp. 1–8, September 2009.

[132] J. Kim, J. S. Pak, J. Cho, E. Song, J. Cho, J. Lee, H. Lee, K. Park, S. Yang, M.-S. Suh, K.-Y. Byun, and J. Kim, "High-Frequency Scalable Electrical Model and Analysis of a Through Silicon Via (TSV)," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, Vol. 1, No. 2, pp. 181–195, February 2011.

[133] Z. Xu and J.-Q. Lu, "High-Speed Design and Broadband Modeling of Through-Strata-Vias (TSVs) in 3D Integration," *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, Vol. 1, No. 2, pp. 154-162, February 2011.

[134] I. Savidis and E. G. Friedman, "Closed-Form Expressions of 3-D Via Resistance, Inductance, and Capacitance," *IEEE Transactions on Electron Devices*, Vol. 56, No. 9, pp. 1873-1881, September 2009.

[135] K. Salah, H. Ragai, Y. Ismail, and A. El Rouby, "Equivalent Lumped Element Models for Various n-Port Through Silicon Vias Networks," *Proceedings of the IEEE International Asia and South Pacific Design Automation Conference*, pp. 176-183, January 2011.

[136] Z. Guo and G. G. Pan, "On Simplified Fast Modal Analysis for Through Silicon Vias in Layered Media Based Upon Full-Wave Solutions," *IEEE Transactions on Advanced Packaging*, Vol. 33, No. 2, pp. 517–523, May 2010.

[137] T. H. Lee, *Planar Microwave Engineering*, Cambridge University Press, 2004.

[138] K. Banerjee, S. J. Souri, and P. Kapur, "3-D ICs: A Novel Chip Design for Improving Deep-Submicrometer Interconnect Performance and Systems-on-Chip Integration," *Proceedings of the IEEE*, Vol. 89, No. 5, pp. 602–633, May 2001.

[139] J. E. Sergent and A. Krum, Eds., *Thermal Management Handbook for Electronic Assemblies*, McGraw-Hill, 1998.

[140] M. R. Ward, *Electrical Engineering Science*, McGraw-Hill, 1971.

[141] R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Kose, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors*, *Second Edition*, Springer Science+Business Media, 2011.

[142] E. B. Rosa, "The Self and Mutual Inductances of Linear Conductors," *Bulletin of the Bureau of Standards*, Vol. 4, No. 2, pp. 301-344, 1908.

[143] A. Ruehli, "Inductance Calculations in a Complex Integrated Circuit Environment," *IBM Journal of Research and Development*, Vol. 16, No. 5, pp. 470-481, September 1972.

[144] A. Mezhiba and E. Friedman, "Inductive Properties of High-Performance Power Distribution Grids," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 10, No. 6, pp. 762–776, December 2002.

[145] M. W. Beattie and L. T. Pileggi, "On-Chip Induction Modeling: Basics and Advanced Methods," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 10, No. 6, pp. 712-729, December 2002.

[146] M. Beattie and L. Pileggi, "Inductance 101: Modeling and Extraction," *Proceedings of the IEEE Design Automation Conference*, pp. 323-328, June 2001.

[147] R. Pucel, "Gallium Arsenide Technology," D. Feny, Ed., Chapter 6, p. 216, 1985.

[148] M. E. Goldfarb and R. A. Pucel, "Modeling Via Hole Grounds in Microstrip," *IEEE Microwave and Guided Wave Letters*, Vol. 1, No. 6, pp. 135-137, June 1991.

[149] C. Xu, V. Kourkoulos, R. Suaya, and K. Banerjee, "A Fully Analytical Model for the Series Impedance of Through-Silicon Vias With Consideration of Substrate Effects and Coupling With Horizontal Interconnects," *IEEE Transactions on Electron Devices*, Vol. 58, No. 10, pp. 3529-3540, October 2011.

[150] S. M. Alam, R. E. Jones, S. Rauf, and R. Chatterjee, "Inter-Strata Connection Characteristics and Signal Transmission in Three-dimensional (3D) Integration Technology," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 580–585, March 2007.

[151] S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, *Third Edition*, John Wiley & Sons Press, 2006.

[152] Y. Tsividis and C. McAndrew, *Operation and Modeling of the MOS Transistor*, *Third Edition*, Oxford University Press, 2010.

[153] C. Jueping, J. Peng, Y. Lei, H. Yue, and L. Zan, "Through-Silicon Via (TSV) Capacitance Modeling for 3D NoC Energy Consumption Estimation," *Proceedings of the IEEE International Conference on Solid-State and Integrated Circuit Technology*, pp. 815-817, November 2010.

[154] F. T. Ulaby, *Fundamentals of Applied Electromagnetics*, Prentice Hall, 2004.

[155] B. Curran, I. Ndip, S. Guttovski, and H. Reichl, "The Impacts of Dimensions and Return Current Path Geometry on Coupling in Single Ended Through Silicon Vias," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1092–1097, May 2009.

[156] N. N. Rao, *Elements of Engineering Electromagnetics*, *Fifth Edition*, Prentice Hall, 2000.

[157] J. Kim, P. J. S., J. Cho, J. Lee, H. Lee, K. Park, and J. Kim, "Modeling and Analysis of Differential Signal Through Silicon Via (TSV) in 3D IC," *Proceedings of the IEEE Components, Packaging, and Manufacturing Technology Symposium Japan*, pp. 1-4, August 2010.

[158] H.-Y. Liao and H.-K. Chiou, "RF Model and Verification of Through-Silicon Vias in Fully Integrated SiGe Power Amplifier," *IEEE Electron Device Letters*, Vol. 32, No. 6, pp. 809-811, June 2011.

[159] K. Salah, A. El Rouby, H. Ragai, K. Amin, and Y. Ismail, "Compact Lumped Element Model for TSV in 3D-ICs," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2321–2324, May 2011.

[160] C. Bermond, L. Cadix, A. Farcy, T. Lacrevaz, P. Leduc, and B. Flechet, "High Frequency Characterization and Modeling of High Density TSV in 3D Integrated Circuits," *Proceedings of the IEEE Workshop on Signal Propagation on Interconnects*, pp. 1–4, May 2009.

[161] M. Popovich, A. V. Mezhiba, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors*, Springer-Verlag, 2008.

[162] E. Salman and E. G. Friedman, *High Performance Integrated Circuit Design*, McGraw-Hill, 2012.

[163] R. R. Tummala, A. Hughes, and A. Poole, *Fundamentals of Microsystems Packaging*, McGraw-Hill, 2001.

[164] G. R. Blackwell, *The Electronic Packaging Handbook*, CRC Press, 2000.

[165] P. Zhou, K. Sridharan, and S. S. Sapatnekar, "Optimizing Decoupling Capacitors in 3D Circuits for Power Grid Integrity," *IEEE Design and Test of Computers*, Vol. 26, No. 5, pp. 15–25, September 2009.

[166] K. Kim, J. S. Pak, H. Lee, and J. Kim, "Effects of On-Chip Decoupling Capacitors and Silicon Substrate on Power Distribution Networks in TSV-based 3D-ICs," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 690-697, May 2012.

[167] N. H. Khan, S. Reda, and S. Hassoun, "Early Estimation of TSV Area for Power Delivery in 3-D Integrated Circuits," *Proceedings of the IEEE International 3D Systems Integration Conference*, pp. 1-6, November 2010.

[168] Q. Wu and T. Zhang, "Design Techniques to Facilitate Processor Power Delivery in 3-D Processor-DRAM Integrated Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 19, No. 9, pp. 1655-1666, September 2011.

[169] V. F. Pavlidis and G. De Micheli, "Power Distribution Paths in 3-D ICs," *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI*, pp. 263-268, May 2009.

[170] S. M. Satheesh and E. Salman, "Power Distribution in TSV-Based 3-D Processor-Memory Stacks," *IEEE Transactions on Emerging and Selected Topics in Circuits and Systems*, Vol. 2, No. 4, pp. 692-703, December 2012.

[171] J. Xie and M. Swaminathan, "DC IR Drop Solver for Large Scale 3D Power Delivery Networks," *Proceedings of the IEEE Electrical Performance of Electronic Packaging and Systems*, pp. 217-220, October 2010.

[172] X. Hu, P. Du, J. F. Buckwalter, and C.-K. Cheng, "Modeling and Analysis of Power Distribution Networks in 3-D ICs," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 21, No. 2, pp. 354-366, February 2013.

[173] S. Pant and E. Chiprout, "Power Grid Physics and Implications for CAD," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 199-204, June 2006.

[174] X. Hu, P. Du, and C. K. Cheng, "Exploring the Rogue Wave Phenomenon in 3D Power Distribution Networks," *Proceedings of the IEEE Electrical Performance of Electronic Packaging and Systems*, pp. 57-60, October 2010.

[175] Z. Z. Oo, L. En-Xiao, J. R. Cubillo, and L. Er-Ping, "Systematic Analysis for Static and Dynamic Drops in Power Supply Grids of 3-D Integrated Circuits," *Proceedings of the IEEE Asia-Pacific Symposium on Electromagnetic Compatibility*, pp. 49-52, May 2012.

[176] A. Shayan, X. Hu, H. Peng, M. Popovich, W. Zhang, C.-K. Cheng, L. Chua-Eoan, and X. Chen, "3D Power Distribution Network Co-Design for Nanoscale Stacked Silicon IC," *Proceedings of the IEEE Electrical Performance of Electronic Packaging*, pp. 11-14, October 2008.

[177] G. Huang, M. Bakir, A. Naeemi, H. Chen, and J. D. Meindl, "Power Delivery for 3D Chip Stacks: Physical Modeling and Design Implication," *Proceedings of the IEEE Electrical Performance of Electronic Packaging*, pp. 205–208, October 2007.

[178] W. Ahmad, Q. Chen, L.-R. Zheng, and H. Tenhunen, "Modeling of Peak-to-Peak Switching Noise Along a Vertical Chain of Power Distribution TSV Pairs in a 3D Stack of ICs Interconnected Through TSVs," *Proceedings of the NORCHIP Conference*, pp. 1-6, November 2010.

[179] S. Kose and E. G. Friedman, "Fast Algorithms for Power Grid Analysis Based on Effective Resistance," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 3661–3664, May 2010.

[180] G. Venezian, "On the Resistance between Two Points on a Grid," *American Journal of Physics*, Vol. 62, No. 11, pp. 1000–1004, November 1994.

[181] S. Kose and E. G. Friedman, "Effective Resistance of a Two Layer Mesh," *IEEE Transactions on Circuits and Systems II: Express Briefs*, Vol. 58, No. 11, pp. 739–743, November 2011.

[182] S. Kose and E. G. Friedman, "Efficient Algorithms for Fast IR Drop Analysis Exploiting Locality," *Integration, The VLSI Journal*, Vol. 45, No. 2, pp. 149–161, March 2012.

[183] N. H. Khan, S. M. Alam, and S. Hassoun, "Power Delivery Design for 3-D ICs Using Different Through-Silicon Via (TSV) Technologies," *IEEE Transactions on*

Very Large Scale Integration (VLSI) Systems, Vol. 19, No. 4, pp. 647–658, April 2011.

[184] T. Okoshi, Y. Uehara, and T. Takeuchi, "The Segmentation Method–An Approach to the Analysis of Microwave Planar Circuits," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 24, No. 10, pp. 662-668, October 1976.

[185] J. Kim, W. Lee, Y. Shim, J. Shim, K. Kim, J. S. Pak, and J. Kim, "Chip-Package Hierarchical Power Distribution Network Modeling and Analysis Based on a Segmentation Method," *IEEE Transactions on Advanced Packaging*, Vol. 33, No. 3, pp. 647-659, August 2010.

[186] I. Savidis, S. Kose, and E. G. Friedman, "Power Noise in TSV-Based 3-D Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 48, No. 2, pp. 587-597, February 2013.

[187] G. Huang, D. C. Sekar, A. Naeemi, K. Shakeri, and J. D. Meindl, "Compact Physical Models for Power Supply Noise and Chip/Package Co-Design of Gigascale Integration," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1659-1666, May 2007.

[188] W. Zhang, R. Zhang, R. Shi, H. Peng, Z. Zhu, L. Chua-Eoan, R. Murgai, T. Shibuya, N. Ito, and C. K. Cheng, "Fast Power Network Analysis with Multiple Clock Domains," *Proceedings of the IEEE International Conference on Computer Design*, pp. 456-463, October 2007.

[189] H. Yu, J. Ho, and L. He, "Allocating Power Ground Vias in 3D ICs for Simultaneous Power and Thermal Integrity," *ACM Transactions on Design Automation of Electronic Systems*, Vol. 14, No. 3, pp. 1-31, June 2009.

[190] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, Society for Industrial and Applied Mathematics, 1995.

[191] A. D. Polyanin, *Handbook of Linear Partial Differential Equations for Engineers and Scientists*, CRC Press, 2002.

[192] S. Kose, S. Tam, S. Pinzon, B. McDermott, and E. G. Friedman, "Active Filter Based Hybrid On-Chip DC-DC Converters for Point-of-Load Voltage Regulation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 21, No. 4, pp. 680-691, April 2013.

[193] V. Kursun and E. G. Friedman, *Multi-Voltage CMOS Circuit Design*, Wiley & Sons, 2006.

[194] E. Bonizzoni, F. Borghetti, P. Malcovati, and F. Maloberti, "A 200-mA, 93% Peak Power Efficiency, Single-Inductor, Dual-Output DCDC Buck Converter," *Analog Integrated Circuits and Signal Processing*, Vol. 62, No. 2, pp. 121–129, June 2009.

[195] J. Rosenfeld and E. G. Friedman, "A Distributed Filter Within a Switching Converter for Application to 3-D Integrated Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems Systems*, Vol. 2, pp. 1–11, 2010.

[196] J. Rosenfeld and E. G. Friedman, "Linear and Switch-Mode Conversion in 3-D Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 19, No. 11, pp. 2095-2108, November 2011.

[197] K. Onizuka, K. Inagaki, H. Kawaguchi, M. Takamiya, and T. Sakurai, "Stacked-Chip Implementation of On-Chip Buck Converter for Distributed Power Supply System in SiPs," *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 11, pp. 2404–2410, November 2007.

[198] J. Rosenfeld and E. G. Friedman, "On-Chip DC-DC Converters for Three-Dimensional ICs," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, No. 1, pp. 759–764, March 2009.

[199] K. Ishida, K. Takemura, K. Baba, M. Takamiya, and T. Sakurai, "3D Stacked Buck Converter with 15 μm Thick Spiral Inductor on Silicon Interposer for Fine-Grain Power-Supply Voltage Control in SiP's," *Proceedings of the IEEE International 3-D Systems Integration Conference*, pp. 1-4, November 2010.

[200] G. Schrom, P. Hazucha, J.-H. Hahn, V. Kursun, D. Gardner, S. Narendra, T. Karnik, and V. De, "Feasibility of Monolithic and 3D-Stacked DC-DC Converters for Microprocessors in 90nm Technology Generation," *Proceedings of the ACM International Symposium on Low Power Electronics and Design*, pp. 263–268, August 2004.

[201] J. Sun, D. Giuliano, S. Devarajan, J.-Q. Lu, T. P. Chow, and R. J. Gutmann, "Fully Monolithic Cellular Buck Converter Design for 3-D Power Delivery," *IEEE*

Transactions on Very Large Scale Integration (VLSI) Systems, Vol. 17, No. 3, pp. 447–451, March 2009.

[202] M.-J. Wang, C.-Y. Hung, C.-L. Kao, P.-N. Lee, C.-H. Chen, C.-P. Hung, and H.-M. Tong, "TSV Technology for 2.5D IC Solution," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 284-288, May 2012.

[203] C. Zhang and G. Sun, "Fabrication Cost Analysis for 2D, 2.5D, and 3D IC Designs," *Proceedings of the IEEE 3D Systems Integration Conference*, pp. 1-4, January 2012.

[204] Y.-K. Ho and Y.-W. Chang, "Multiple Chip Planning for Chip-Interposer Codesign," *Proceedings of the ACM/EDAC/IEEE Design Automation Conference*, pp. 1-6, May 2013.

[205] K. Onizuka, K. Inagaki, H. Kawaguchi, M. Takamiya, and T. Sakurai, "Stacked-Chip Implementation of On-Chip Buck Converter for Distributed Power Supply System in SiPs," *IEEE Journal of Solid-State Circuits*, Vol. 42, No. 11, pp. 2404–2410, November 2007.

[206] *Predictive Technology Model*, Nanoscale Integration and Modeling Group, Arizona State University, Tempe. [Online]. Available: http://www.eas.asu.edu/~ptm

[207] I. Savidis and E. G. Friedman, "Electrical Modeling and Characterization of 3-D Vias," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 784–787, May 2008.

[208] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*, Springer, 2004.

[209] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu, "On Thermal Effects in Deep Sub-Micron VLSI Interconnects," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 885-890, June 1999.

[210] K. Banerjee, M. Pedram, and A. H. Ajami, "Analysis and Optimization of Thermal Issues in High-Performance VLSI," *Proceedings of the IEEE International Symposium on Physical Design*, pp. 230-237, April 2001.

[211] J. R. Black, "Electromigration — A Brief Survey and Some Recent Results," *IEEE Transactions on Electron Devices*, Vol. ED-16, No. 4, pp. 338-347, April 1969.

[212] K. Banerjee, A. Amerasekera, N. Cheung, and C. Hu, "High-Current Failure Model for VLSI Interconnects Under Short-Pulse Stress Conditions," *IEEE Transactions on Electron Device Letters*, Vol. 18, No. 9, pp. 405-407, September 1997.

[213] S. Im and K. Banerjee, "Full Chip Thermal Analysis of Planar (2-D) and Vertically Integrated (3-D) High Performance ICs," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 727-730, December 2000.

[214] A. Zjajo, N. Meijs, and R. Leuken, "Thermal Analysis of 3D Integrated Circuits Based on Discontinuous Galerkin Finite Element Method," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 117-122, March 2012.

[215] C. Torregiani, H. Oprins, B. Vandevelde, E. Beyne, and I. De Wolf, "Compact Thermal Modeling of Hot Spots in Advanced 3D-Stacked ICs," *Proceedings of the IEEE Electronics Packaging Technology Conference*, pp. 131-136, December 2009.

[216] C. Torregiani, H. Oprins, B. Vandevelde, E. Beyne, and I. De Wolf, "Analytical Heat Transfer Model for Thermal Through-Silicon Vias," *Proceedings of the IEEE Design, Automation & Test in Europe Conference and Exhibition*, pp. 1-6, March 2011.

[217] A. Jain, R. E. Jones, R. Chatterjee, and S. Pozder, "Analytical and Numerical Modeling of the Thermal Performance of Three-Dimensional Integrated Circuits," *IEEE Transactions on Components and Packaging Technologies*, Vol. 33, No. 1, pp. 56-63, March 2010.

[218] K. Banerjee, A. Mehrotra, A. Sangiovanni-Vincentelli, and C. Hu, "On Thermal Effects in Deep Sub-Micron VLSI Interconnects," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 885–891, June 1999.

[219] M. B. Kleiner, S. A. Kuhn, P. Ramm, and W. Weber, "Thermal Analysis of Vertically Integrated Circuits," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 487–490, December 1995.

[220] C. C. Liu, A. K. Datta, and S. Tiwari, "Heating Effects of Clock Drivers in Bulk, SOI, and 3-D CMOS," *IEEE Electron Device Letters*, Vol. 23, No. 12, pp. 716–728, December 2002.

[221] M. N. Özisik, *Heat Transfer: A Basic Approach*, McGraw-Hill, 1985.

[222] E. R. Eckert and D. Eckert, *Analysis of Heat and Mass Transfer*, CRC Press, 1986.

[223] S. Sapatnekar, "Addressing Thermal and Power Delivery Bottlenecks in 3D Circuits," *Proceedings of the IEEE Asia and South Pacific Design Automation Conference*, pp. 423-428, January 2009.

[224] D. Oh, C. C. P. Chen, and Y. H. Hu, "Efficient Thermal Simulation for 3-D IC With Thermal Through-Silicon Vias," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 31, No. 11, pp. 1767–1771, November 2012.

[225] A. Rahman, A. Fan, and R. Reif, "Thermal Analysis of Three-Dimensional (3-D) Integrated Circuits (ICs)," *Proceedings of the IEEE International Interconnect Technology Conference*, pp. 157-159, June 2001.

[226] S. Das, *Design Automation and Analysis of Three-Dimensional Integrated Circuits*, Ph.D. Dissertation, Massachusetts Institute of Technology, May 2004.

[227] A. Jain, R. E. Jones, R. Chatterjee, and S. Pozder, "Analytical and Numerical Modeling of the Thermal Performance of Three-Dimensional Integrated Circuits," *IEEE Transactions on Components and Packaging Technologies*, Vol. 33, No. 1, pp. 56–63, March 2010.

[228] S. Das, A. Fan, K.-N. Chen, C. S. Tan, N. Checka, and R. Reif, "Technology, Performance, and Computer-Aided Design of Three-Dimensional Integrated Circuits," *Proceedings of the IEEE International Symposium on Physical Design*, pp. 108–115, 2004.

[229] J. Palacín, M. Sallera, J. Samitier, and S. Marco, "Dynamic Compact Thermal Models With Multiple Power Sources: Application to an Ultrathin Chip Stacking Technology," *IEEE Transactions on Advanced Packaging*, Vol. 28, No. 4, pp. 694-703, November 2005.

[230] P. Wilkerson, M. Furmanczyk, and M. Turowski, "Fast, Automated Thermal Simulation of Three-Dimensional Integrated Circuits," *Proceedings of the Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 706-713, June 2004.

[231] P. Wilkerson, M. Furmanczyk, and M. Turowski, "Compact Thermal Model Analysis for 3-D Integrated Circuits," *Proceedings of the IEEE International Conference on Mixed Design of Integrated Circuits and Systems*, pp. 277-282, June 2004.

[232] G. Digele, S. Lindenkreuz, and E. Kasper, "Fully Coupled Dynamic Electro-Thermal Simulation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 5, No. 3, pp. 250-257, September 1997.

[233] G. Digele, S. Lindenkreuz, and E. Kasper, "Electro-Thermal Circuit Simulation Using Simulator Coupling," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 5, No. 3, pp. 277-282, September 1997.

[234] B. Goplen and S. Sapatnekar, "Placement of Thermal Vias in 3-D ICs Using Various Thermal Objectives," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 4, pp. 692–709, April 2006.

[235] S. Lee, T. F. Lemczyk, and M. M. Yovanovich, "Analysis of Thermal Vias in High Density Interconnect Technology," *Proceedings of the IEEE Semiconductor Thermal Measurement and Management Symposium*, pp. 55-61, February 1992.

[236] R. S. Li, "Optimization of Thermal Via Design Parameters Based on an Analytical Thermal Resistance Model," *Proceedings of the Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 475-480, June 1998.

[237] T.-Y. Chiang, K. Banerjee, and K. C. Saraswat, "Analytical Thermal Model for Multilevel VLSI Interconnects Incorporating Via Effect," *IEEE Transactions on Electron Device Letters*, Vol. 23, No. 1, pp. 31-33, January 2002.

[238] B. Goplen and S. Sapatnekar, "Placement of 3D ICs with Thermal and Interlayer Via Considerations," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 626–631, June 2007.

[239] J. Li and H. Miyashita, "Efficient Thermal Via Planning for Placement of 3D Integrated Circuits," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 145–148, May 2007.

[240] J.-T. Yan, Y.-C. Chang, and Z.-W. Chen, "Thermal Via Planning for Temperature Reduction in 3D ICs," *Proceedings of the IEEE International SOC Conference*, pp. 392-395, September 2010.

[241] T.-Y. Chiang, K. Banerjee, and K. C. Saraswat, "Effect of Via Separation and Low-k Dielectric Materials on the Thermal Characteristics of Cu Interconnects," *Proceedings of the IEEE International Electron Devices Meeting Technical Digest*, pp. 261-264, December 2000.

[242] T.-Y. Chiang, S. J. Souri, C. O. Chui, and K. C. Saraswat, "Thermal Analysis of Heterogeneous 3-D ICs with Various Integration Scenarios," *Proceedings of the IEEE International Electron Devices Meeting Technical Digest*, pp. 681-684, December 2001.

[243] J. Cong and Y. Zhang, "Thermal Via Planning for 3-D ICs," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 745-752, November 2005.

[244] E. Wong and S. K. Lim, "Whitespace Redistribution for Thermal Via Insertion in 3D Stacked ICs," *Proceedings of the IEEE International Conference on Computer Design*, pp. 267–272, October 2007.

[245] A. Jain, S. M. Alam, S. Pozder, and R. E. Jones, "Thermal-Electrical Co-Optimisation of Floorplanning of Three-Dimensional Integrated Circuits Under Manufacturing and Physical Design Constraints," *Institution of Engineering and Technology Computers and Digital Techniques*, Vol. 5, No. 3, pp. 169–178, May 2011.

[246] X. Li, Y. Ma, X. Hong, and S. Dong, "Thermal-Aware Incremental Floorplanning for 3D ICs," *Proceedings of the IEEE International Conference on ASIC*, pp. 1092-1095, October 2007.

[247] J. Cong, G. Luo, J. Wei, and Y. Zhang, "Thermal-Aware 3D IC Placement Via Transformation," *Proceedings of the IEEE Asia and South Pacific Design Automation Conference*, pp. 780–785, January 2007.

[248] J. Cong, J. Wei, and Y. Zhang, "A Thermal-Driven Floorplanning Algorithm for 3D ICs," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 306-313, November 2004.

[249] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, Vol. 220, No. 4598, pp. 671-680, May 1983.

[250] H. Eisenmann and F. M. Johannes, "Generic Global Placement and Floorplanning," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 269-274, June 1998.

[251] A. B. Kahng and Q. Wang, "An Analytic Placer for Mixed-Size Placement and Timing-Driven Placement," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 265-272, November 2004.

[252] F. Kreith, *Principles of Heat Transfer*, *Third Edition*, Harper and Row, 1973.

[253] Y. A. Cengel, *Heat Transfer: A Practical Approach*, McGraw-Hill, 1998.

[254] D. B. Tuckerman and R. F. W. Pease, "High-Performance Heat Sinking for VLSI," *IEEE Transactions on Electron Device Letters*, Vol. 2, No. 5, pp. 126-129, May 1981.

[255] V. K. Samalam, "Convective Heat Transfer in Microchannels," *Journal of Electronic Materials*, Vol. 18, No. 5, pp. 611-618, September 1989.

[256] M. S. Bakir, C. King, D. Sekar, H. Thacker, B. Dang, G. Huang, A. Naeemi, and J. D. Meindl, "3D Heterogeneous Integrated Systems: Liquid Cooling, Power Delivery, and Implementation," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 663–670, September 2008.

[257] C. R. King, J. Zaveri, M. S. Bakir, and J. D. Meindl, "Electrical and Fluidic C4 Interconnections for Inter-Layer Liquid Cooling of 3D ICs," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1674–1681, June 2010.

[258] J.-M. Koo, S. Im, L. Jiang, and K. E. Goodson, "Integrated Microchannel Cooling for Three-Dimensional Electronic Circuit Architectures," *ASME Journal of Heat Transfer*, Vol. 127, No. 1, pp. 49-58, January 2005.

[259] B. Dang, M. S. Bakir, D. C. Sekar, C. R. King Jr., and J. D. Meindl, "Integrated Microfluidic Cooling and Interconnects for 2D and 3D Chips," *IEEE Transactions on Advanced Packaging*, Vol. 33, No. 1, pp. 79–87, February 2010.

[260] B. Shi, A. Srivastava, and P. Wang, "Non-Uniform Micro-Channel Design for Stacked 3D-ICs," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 658-663, June 2011.

[261] Y. J. Kim, Y. K. Joshi, A. G. Fedorov, Y.-J. Lee, and S.-K. Lim, "Thermal Characterization of Interlayer Microfluidic Cooling of Three-Dimensional Integrated Circuits with Nonuniform Heat Flux," *ASME Journal of Heat Transfer*, Vol. 132, No. 4, pp. 0410091 - 0410099, April 2010.

[262] A. Sridhar, A. Vincenzi, M. Ruggiero, T. Brunschwiler, and D. Atienza, "3D-ICE : Fast Compact Transient Thermal Modeling for 3D ICs with Inter-Tier Liquid Cooling," *Proceedings of the IEEE International Conference on Computer-Aided Design*, pp. 463–470, November 2010.

[263] H. Mizunuma, Y.-C. Lu, and C.-L. Yang, "Thermal Modeling and Analysis for 3-D ICs with Integrated Microchannel Cooling," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 30, No. 9, pp. 1293–1306, September 2011.

[264] A. Sridhar, A. Vincenzi, M. Ruggiero, T. Brunschwiler, and D. Atienza, "Compact Transient Thermal Model for 3D ICs with Liguid Cooling via Enhanced Heat Transfer Cavity Geometries," *Proceedings of the IEEE International Workshop on Thermal Investigation of ICs and Systems*, pp. 105-110, October 2010.

[265] R. W. Knight, D. J. Hall, J. S. Goodling, and R. C. Jaeger, "Heat Sink Optimization with Application to Microchannels," *IEEE Transactions on Components, Hybrids, and Manufacturing Technology*, Vol. 15, No. 5, pp. 832-842, October 1992.

[266] L. Zhang, J.-M. Koo, L. Jiang, M. Asheghi, K. E. Goodson, J. G. Santiago, and T. W. Kenny, "Measurements and Modeling of Two-Phase Flow in Microchannels with Nearly Constant Heat Flux Boundary Conditions," *Journal of Microelectromechanical Systems*, Vol. 11, No. 1, pp. 12-19, February 2002.

[267] D. Sekar, C. King, B. Dang, T. Spencer, H. Thacker, P. Joseph, M. S. Bakir, and J. D. Meindl, "A 3D-IC Technology with Integrated Microchannel Cooling," *Proceedings of the IEEE International Interconnect Technology Conference*, pp. 13-15, June 2008.

[268] K. Nose and T. Sakurai, "Optimization of VDD and VTH for Low-Power and High-Speed Applications," *Proceedings of the IEEE Asia South Pacific Design Automation Conference*, pp. 469-474, January 2000.

[269] *MITLL Low-Power FDSOI CMOS Process Application Notes*, MIT Lincoln Laboratories, June 2006.

[270] T. Sakurai and K. Tamaru, "Simple Formulas for Two- and Three-Dimensional Capacitances," *IEEE Transactions on Electron Devices*, Vol. ED-30, No. 2, pp. 183-185, February 1983.

[271] M. I. Elmasry, "Capacitance Calculations in MOSFET VLSI," *IEEE Electron Device Letters*, Vol. EDL-3, No. 1, pp. 6-7, January 1982.

[272] A. Rahman, J. Trezza, B. New, and S. Trimberger, "Die Stacking Technology for Terabit Chip-to-Chip Communications," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 587-590, September 2006.

[273] L. L. W. Leung and K. J. Chen, "Microwave Characterization and Modeling of High Aspect Ratio Through-Wafer Interconnect Vias in Silicon Substrates," *IEEE Transactions on Microwave Theory and Techniques*, Vol. 53, No. 8, pp. 2472-2480, August 2005.

[274] F. M. Finkbeiner, C. Adams, E. Apodaca, J. A. Chervenak, J. Fischer, N. Doan, M. J. Li, C. K. Stahle, R. P. Brekosky, S. R. Bandler, E. Figueroa-Feliciano, M. A. Lindeman, R. L. Kelley, T. Saab, and D. J. Talley, "Development of Ultra-Low Impedance Through-Wafer Micro-Vias," *Nuclear Instruments and Methods in Physics Research Section A*, Vol. 520, No. 1-3, pp. 463-465, March 2004.

[275] E. M. Chow, V. Chandrasekaran, A. Partridge, T. Nishida, M. Sheplak, C. F. Quate, and T. W. Kenny, "Process Compatible Polysilicon-Based Electrical Through-Wafer Interconnects in Silicon Substrates," *Journal of Microelectromechanical Systems*, Vol. 11, No. 6, pp. 631-640, December 2002.

[276] I. Luusua, K. Henttinen, P. Pekko, T. Vehmas, and H. Luoto, "Through-Wafer Polysilicon Interconnect Fabrication with In-Situ Boron Doping," *Micro- and Nanosystems – Materials and Devices*, Vol. 872, pp. 77-81, March 2005.

[277] R. Weerasekera, D. Pamunuwa, M. Grange, H. Tenhunen, and L.-R. Zheng, "Closed-Form Equations for Through-Silicon Via (TSV) Parasitics in 3-D Integrated Circuits

(ICs)," *Proceedings of the IEEE Design, Automation & Test in Europe Conference*, pp. 1-3, April 2009.

[278] A. Y. Weldezion, R. Weerasekera, D. Pamunuwa, L.-R. Zheng, and H. Tenhunen, "Bandwidth Optimization for Through Silicon Via (TSV) Bundles in 3D Integrated Circuits," *Proceedings of the IEEE Design, Automation & Test in Europe Conference*, pp. 1-4, April 2009.

[279] R. Weerasekera, *System Interconnection Design Trade-offs in Three-Dimensional Integrated Circuits*, Ph.D. Dissertation, KTH School of Information and Communication Technologies, Stockholm, Sweden, December 2008.

[280] *Ansoft Quick 3-D*, "http://www.ansoft.com/products/si/q3d_extractor/".

[281] D. L. Logan, *A First Course in the Finite Element Method*, *Fourth Edition*, Nelson, A Division of Thomson Canada Limited, 2007.

[282] A. R. Hall, *Generalized Method of Moments*, Oxford University Press, 2005.

[283] E. K. Miller, L. Medgyesi-Mitschang, and E. H. Newman, Eds., *Computational Electromagnetics: Frequency-Domain Method of Moments*, IEEE Press, 1992.

[284] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, *Second Edition*, Cambridge University Press, 2004.

[285] Y. I. Ismail and E. G. Friedman, "On the Extraction of On-Chip Inductance," *Journal of Circuits, Systems and Computers*, Vol. 12, No. 1, pp. 31-40, February 2003.

[286] E. G. Friedman, Ed., *Clock Distribution Networks in VLSI Circuits and Systems*, IEEE Press, 1995.

[287] E. G. Friedman, "Clock Distribution Networks in Synchronous Digital Integrated Circuits," *Proceedings of the IEEE*, Vol. 89, No. 5, pp. 665-692, May 2001.

[288] W. Bailey and B. J. Benschneider, "Clocking Design and Analysis for a 600-MHz Alpha Microprocessor," *IEEE Journal of Solid State Circuits*, Vol. 22, No. 11, pp. 1627-1633, November 1998.

[289] T. Xanthopoulos, D. W. Bailey, A. K. Gangwar, M. K. Gowan, A. K. Jain, and B. K. Prewitt, "The Design and Analysis of the Clock Distribution Network for a 1.2 GHz Alpha Microprocessor," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 402-403, February 2001.

[290] V. F. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Networks in 3-D Integrated Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 19, No. 12, pp. 2256-2266, December 2011.

[291] V. F. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Networks for 3-D Integrated Circuits," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 651-654, September 2008.

[292] V. F. Pavlidis, I. Savidis, and E. G. Friedman, "Clock Distribution Architectures for 3-D SOI Integrated Circuits," *Proceedings of the IEEE International Silicon-on-Insulator Conference*, pp. 111-112, October 2008.

[293] D. L. Lewis and H.-H. S. Lee, "A Scan-Island Based Design Enabling Pre-bond Testability in Die-Stacked Microprocessor," *Proceedings of the IEEE International Test Conference*, pp. 21.2.1-21.2.8, October 2007.

[294] J. Minz, X. Zhao, and S. K. Lim, "Buffered Clock Tree Synthesis for 3D ICs Under Thermal Variations," *Proceedings of the IEEE International Asia and South Pacific Design Automation Conference*, pp. 504-509, January 2008.

[295] M. Mondal, A. J. Ricketts, S. Kirolos, T. Ragheb, G. Link, N. Viyaykrishnan, and Y. Massoud, "Thermally Robust Clocking Schemes for 3D Integrated Circuits," *Proceedings of the IEEE International Conference on Design, Automation and Test in Europe*, pp. 1206-1211, April 2007.

[296] V. Arunachalam and W. Burleson, "Low-Power Clock Distribution in a Multilayer Core 3D Microprocessor," *Proceedings of the ACM/IEEE International Great Lakes Symposium on VLSI*, pp. 429-434, May 2008.

[297] V. F. Pavlidis, *Interconnect-Based Design Methodologies for Three-Dimensional Integrated Circuits*, Ph.D. Dissertation, University of Rochester, July 2008.

[298] *Available online: http://www.ece.ncsu.edu/erl/3DIC/pub*.

[299] N. Hedenstiema and K. O. Jeppson, "CMOS Circuit Speed and Buffer Optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 6, No. 2, pp. 270-281, March 1987.

[300] B. S. Cherkauer and E. G. Friedman, "A Unified Design Methodology for CMOS Tapered Buffers," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 3, No. 1, pp. 99-111, March 1995.

[301] H. B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI*, Addison-Wesley, 1990.

[302] I. Savidis, V. F. Pavlidis, and E. G. Friedman, "Clock Distribution Models of 3-D Integrated Systems," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2225 - 2228, May 2011.

[303] S. Song and A. Glasser, "Power Distribution Techniques for VLSI Circuits," *IEEE Journal of Solid-State Circuits*, Vol. SC-21, No. 1, pp. 150–156, February 1986.

[304] K. T. Tang and E. G. Friedman, "Simultaneous Switching Noise in On-Chip CMOS Power Distribution Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 10, No. 4, pp. 487-493, August 2002.

[305] K. T. Tang and E. G. Friedman, "Incorporating Voltage Fluctuations of the Power Distribution Network into the Transient Analysis of CMOS Logic Gates," *Analog Integrated Circuits and Signal Processing*, Vol. 31, No. 3, pp. 249-259, June 2002.

[306] M. Popovich and E. G. Friedman, "Decoupling Capacitors for Multi-Voltage Power Distribution Systems," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 3, pp. 217-228, March 2006.

[307] M. Popovich, E. G. Friedman, M. Sotman, and A. Kolodny, "On-Chip Power Distribution Grids with Multiple Supply Voltages for High Performance Integrated Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 16, No. 7, pp. 908-921, July 2008.

[308] S. Zhao, K. Roy, and C.-K. Koh, "Decoupling Capacitance Allocation and its Application to Power-Supply Noise-Aware Floorplanning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 21, No. 1, pp. 81-92, January 2002.

[309] A. Mukheijee and M. Marek-Sadowska, "Clock and Power Gating with Timing Closure," *IEEE Transactions on Design and Test of Computers*, Vol. 20, No. 3, pp. 32-39, May 2003.

[310] J. Rosenfeld and E. G. Friedman, "A Distributed Filter Within a Switching Converter for Application to 3-D Integrated Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems Systems*, Vol. 19, No. 6, pp. 1075–1085, June 2011.

[311] M. Nagata, T. Okumoto, and K. Taki, "A Built-in Technique for Probing Power Supply and Ground Noise Distribution within Large-Scale Digital Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 40, No. 4, pp. 813–819, April 2005.

[312] M. W. Newman, S. Muthukumar, M. Schuelein, T. Dambrauskas, P. A. Dunaway, J. M. Jordan, S. Kulkarni, C. D. Linde, T. A. Opheim, R. A. Stingel, W. Worwag,

L. A. Topic, and J. M. Swan, "Fabrication and Electrical Characterization of 3D Vertical Interconnects," *Proceedings of the IEEE International Electronic Components and Technology Conference*, pp. 394–398, July 2006.

[313] K. Sugahara, T. Nishimura, S. Kusunoki, Y. Akasaka, and H. Nakata, "SOI/SOI/Bulk-Si Triple-Level Structure for Three-Dimensional Devices," *IEEE Electron Device Letters*, Vol. EDL-7, No. 3, pp. 193-194, March 1986.

[314] A. Fan, A. Rahman, and R. Reif, "Copper Wafer Bonding," *Electrochemical and Solid-State Letters*, Vol. 2, No. 10, pp. 534-536, October 1999.

[315] M. B. Kleiner, S. A. Kuhn, P. Ramn, and W. Weber, "Thermal Analysis of Vertically Integrated Circuits," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 487-490, December 1995.

[316] *MITLL Low-Power FDSOI CMOS Process Design Guide*, MIT Lincoln Laboratories, September 2008.

[317] Tezzaron Semiconductor, "Tezzaron Multi-Project Wafer Design Guide," pp. 1–29, February 2010.

[318] *3-D Stacked DRAM/Bi-STAR Overview*, Tezzaron Semiconductor, May 2011. [Online]. Available: http://www.tezzaron.com/memory/Overview_3D_DRAM.htm

[319] *Bi-STAR Technology*, Tezzaron Semiconductor, May 2011. [Online]. Available: http://www.tezzaron.com/technology/Bi_STAR.htm

[320] *FaStack Memory*, Tezzaron Semiconductor, May 2011. [Online]. Available: http://www.tezzaron.com/memory/FaStack_memory.html

[321] R. Patti, "Three-Dimensional Integrated Circuits and the Future of System-on-Chip Designs," *Proceedings of the IEEE*, Vol. 94, No. 6, pp. 1214–1224, June 2006.

[322] R. Patti, "Homogeneous 3-D Integration," *Three Dimensional System Integration*, A. Papanikolaou, D. Soudris, and R. Radojcic, Eds., Chapter 4, pp. 51-71, Springer-Verlag, 2011.

[323] Tezzaron Semiconductor, "Bond Interface Rules," pp. 1–4, July 2007.

[324] Tezzaron Semiconductor, "SuperContact Rules," pp. 1–3, January 2008.

[325] E. Culurciello and P. Weerakoon, "Three-Dimensional Photodetectors in 3-D Silicon-On-Insulator Technology," *IEEE Electron Device Letters*, Vol. 28, No. 2, pp. 117–119, February 2007.

[326] N. McArdle, M. Naruse, H. Toyoda, Y. Kobayashi, and M. Ishikawa, "Reconfigurable Optical Interconnections for Parallel Computing," *Proceedings of the IEEE*, Vol. 88, No. 6, pp. 829–837, June 2000.

[327] J. W. Goodman, F. J. Leonberger, S.-Y. Kung, and R. A. Athale, "Optical Interconnetions for VLSI Systems," *Proceedings of the IEEE*, Vol. 72, No. 7, pp. 850-866, July 1984.

[328] D. A. B. Miller, "Optical Interconnects to Silicon," *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 6, No. 6, pp. 1312-1317, November/December 2000.

[329] A. Shacham, K. Bergman, and L. P. Carloni, "Photonic Networks-on-Chip for Future Generations of Chip Multiprocessors," *IEEE Transactions on Computers*, Vol. 57, No. 9, pp. 1246-1260, September 2008.

[330] A. Shacham and K. Bergman, "Building Ultralow-Latency Interconnection Networks Using Photonic Integration," *IEEE Micro*, Vol. 27, No. 4, pp. 6-20, July 2007.

[331] R. Soref and B. Bennett, "Electrooptical Effects in Silicon," *IEEE Journal of Quantum Electronics*, Vol. 23, No. 1, pp. 123-129, January 1987.

[332] L. Liao, D. Samara-Rubio, M. Morse, A. Liu, D. Hodge, D. Rubin, U. Keil, and T. Franck, "High Speed Silicon Mach-Zehnder Modulator," *Optics Express*, Vol. 13, No. 8, pp. 3129-3135, April 2005.

[333] K. J. Vahala, "Optical Microcavities," *Nature*, Vol. 424, No. 6950, pp. 839-846, August 2003.

[334] M. Haurylau, G. Chen, J. Zhang, N. A. Nelson, D. H. Albonesi, E. G. Friedman, and P. M. Fauchet, "On-Chip Optical Interconnect Roadmap: Challenges and Critical Directions," *IEEE Journal on Selected Topics in Quantum Electronics*, Vol. 12, No. 6, pp. 1699-1705, November/December 2006.

[335] J. Jahns, "Planar Packaging of Free-Space Optical Interconnections," *Proceedings of the IEEE*, Vol. 82, No. 11, pp. 1623-1631, November 1994.

[336] D. V. Plant, M. B. Venditti, E. Laprise, J. Faucher, K. Razavi, M. Chateauneuf, A. G. Kirk, and J. S. Ahearn, "256-Channel Bidirectional Optical Interconnect Using VCSELs and Photodiodes on CMOS," *IEEE Journal of Lightwave Technology*, Vol. 19, No. 8, pp. 1093-1103, August 2001.

[337] M. J. McFadden, M. Iqbal, T. Dillon, R. Nair, T. Gu, D. W. Prather, and M. W. Haney, "Multiscale Free-Space Optical Interconnects for Intrachip Global Communications: Motivation, Analysis, and Experimental Validation," *Journal of Applied Optics*, Vol. 45, No. 25, pp. 6358-6366, September 2006.

[338] C. Debaes, M. Vervaeke, V. Baukens, H. Ottevaere, P. Vynck, P. Tuteleers, B. Volckaerts, W. Meeus, M. Brunfaut, J. Van Campenhout, A. Hermanne, and H. Thienpont, "Low-Cost Microoptical Modules for MCM Level Optical Interconnetions," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 9, No. 2, pp. 518-530, March/April 2003.

[339] J. Xue, M. Huang, H. Wu, E. Friedman, G. Wicks, D. Moore, A. Garg, B. Ciftcioglu, J. Hu, S. Wang, I. Savidis, M. Jain, R. Berman, and P. Liu, "An Intra-Chip Free-Space Optical Interconnect," *Proceedings of the Annual International Symposium on Computer Architecture*, pp. 94–105, June 2010.

[340] B. Ciftcioglu, R. Berman, J. Zhang, Z. Darling, S. Wang, J. Hu, J. Xue, A. Garg, M. Jain, I. Savidis, D. Moore, M. Huang, E. G. Friedman, and G. Wicks, "A 3-D Integrated Intrachip Free-Space Optical Interconnect for Many-Core Chips," *IEEE Photonics Technology Letters*, Vol. 23, No. 3, pp. 164–166, February 2011.

[341] B. Ciftcioglu, R. Berman, S. Wang, J. Hu, I. Savidis, M. Jain, D. Moore, M. Huang, E. G. Friedman, G. Wicks, and H. Wu, "3-D Integrated Heterogeneous Intra-Chip Free-Space Optical Interconnect," *Optics Express*, Vol. 20, No. 4, pp. 4331–4345, February 2012.

[342] R. Michalzik and K. J. Ebeling, "Operating Principles of VCSELs," *Vertical-Cavity Surface-Emitting Laser Devices*, H. Li and K. Iga, Eds., Chapter 3, pp. 53-98, Springer-Verlag, 2003.

[343] T. Egawa and T. Jimbo, "Structure of AlAs/GaAs Distributed Bragg Relector Grown on Si Substrate by Metalorganic Chemical Vapor Deposition," *Journal of Applied Physics*, Vol. 77, No. 8, pp. 2329-2330, April 1995.

[344] J. Wang, I. Savidis, and E. G. Friedman, "Thermal Analysis of Oxide-Confined VCSEL Arrays," *Microelectronics Journal*, Vol. 42, No. 5, pp. 820–825, May 2011.

[345] K. D. Choquette, K. M. Geib, H. C. Chui, B. E. Hammons, H. Q. Hou, T. J. Drummond, and R. Hull, "Selective Oxidation of Buried AlGaAs Versus AlAs Layers," *Applied Physics Letters*, Vol. 69, No. 10, pp. 1385-1387, September 1996.

[346] S. Oktyabrsky, M. Yakimov, V. Tokranov, J. van Eisden, and A. Katsnelson, "Oxidation Lift-Off Technology," *Proceedings of SPIE*, pp. 149-157, August 2005.

[347] B. Ciftcioglu, J. Zhang, R. Sobolewski, and H. Wu, "An 850-nm Normal Incidence Germanium Metal-Semiconductor-Metal Photodetector with 13-GHz Bandwidth and 8- μA Dark Current," *IEEE Photonics Technology Letters*, Vol. 22, No. 24, pp. 1850-1852, December 2010.

[348] B. Ciftcioglu, J. Zhang, L. Zhang, J. R. Marciante, J. D. Zuegel, R. Sobolewski, and H. Wu, "3-GHz Silicon Photodiodes Integrated in a 0.18-μm CMOS Technology," *IEEE Photonics Technology Letters*, Vol. 20, No. 24, pp. 2069–2071, December 2008.

[349] M. D. Austin, H. Ge, W. Wu, M. Li, Z. Yu, D. Wasserman, S. A. Lyon, and S. Y. Chou, "Fabrication of 5 nm Linewidth and 14 nm Pitch Features by Nanoimprint Lithography," *Applied Physics Letters*, Vol. 84, No. 26, pp. 5299-5301, June 2004.

[350] *http://www.tezzaron.com/*, Tezzaron Semiconductor.

[351] F. O'Mahony, J. Kennedy, J. E. Jaussi, G. Balamurugan, M. Mansuri, C. Roberts, S. Shehar, R. Mooney, and B. Casper, "A 47x10 Gb/s 1.4 mW/(Gb/s) Parallel Interface in 45 nm CMOS," *Proceedings of the IEEE Solid-State Circuits Conference*, pp. 156-158, February 2010.

[352] T. Geurts, W. Rens, J. Crols, S. Kashiwakura, and Y. Segawa, "A 2.5 Gbps - 3.125 Gbps Multi-Core Serial-Link Transceiver in 0.13 μm CMOS," *Proceedings of the IEEE European Solid-State Circuits Conference*, pp. 487-490, September 2004.

[353] J. Hu, *Energy-Efficient, Wideband Transceiver Architectures and Circuits for High-Speed Communications and Interconnects*, Ph.D. Dissertation, University of Rochester, July 2012.

[354] *RISC32E Processor Core Datasheet*, Zhejiang University, July 2007.

[355] C. Prott, F. Romer, E. O. Ataro, J. Daleiden, S. Irmer, A. Tarraf, and H. Hillmer, "Modeling of Ultrawidely Tunable Vertical Cavity Air-Gap Filters and VCSELs," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 9, No. 3, pp. 918-928, May/June 2003.

[356] N. Mahmood, B. M. A. Rahman, and K. T. V. Grattan, "Accurate Three-Dimensional Modal Solutions for Optical Resonators with Periodic Layered Structure by Using the Finite Element Method," *Journal of Lightwave Technology*, Vol. 16, No. 1, pp. 156-161, jan 1998.

[357] H. Lee, Y. Song, Y. Lee, and J. Yu, "Thermal Analysis of Asymmetric Intracavity-Contacted Oxide-Aperture VCSELs for Efficient Heat Dissipation," *Solid-State Electronics: An Internation Journal*, Vol. 53, No. 10, pp. 1086-1091, 2009.

[358] T. Ouchi, "Thermal Analysis of Thin-Film Vertical-Cavity Surface-Emitting Lasers Using Finite Element Method," *Japanese Journal of Applied Physics*, Vol. 41, No. 8, pp. 5181-5186, 2002.

[359] P. S. Ivanov, M. Dragas, M. Cryan, and J. M. Rorison, "Theoretical Investigation of Transverse Optical Modes in Photonic-Crystal Waveguides Imbedded into Proton-Implanted and Oxide-Confined Vertical-Cavity Surface-Emitting Lasers," *Journal of the Optical Society of America B*, Vol. 22, No. 10, pp. 2270–2276, October 2005.

[360] W. Hofmann, M. Gorblich, M. Ortsiefer, G. Bohm, and M.-C. Amann, "Long-Wavelength (λ = 1.55 μm) Monolithic VCSEL Array with > 3W CW Output Power," *Electronics Letters*, Vol. 43, No. 19, pp. 1025-1026, September 2007.

[361] F. Lemmi, C. L. Chua, and L. Jeng-Ping, "Hybrid Integration of GaAs-Based VCSEL Array with Amorphous Silicon Sensor," *IEEE Electron Device Letters*, Vol. 25, No. 6, pp. 378-380, June 2004.

[362] L. Piskorski, M. Wasiak, R. P. Sarzala, and W. Nakwaski, "Tuning Effects in Optimisation of GaAs-Based InGaAs/GaAs Quantum-Dot VCSELs," *Optics Communications*, Vol. 281, No. 11, pp. 3163-3170, June 2008.

[363] H. Soda, K. Iga, C. Kitahara, and Y. Suematsu, "GaInAsP/InP Surface Emitting Injection Lasers," *Japanese Journal of Applied Physics*, Vol. 18, No. 12, pp. 2329-2330, December 1979.

[364] D. L. Boiko, G. Guerrero, and E. Kapon, "Thermoelectrical Model for Vertical Cavity Surface Emitting Lasers and Arrays," *Journal of Applied Physics*, Vol. 100, No. 10, pp. 103102-1 - 103102-9, November 2006.

[365] F. M. di Sopra, M. Brunner, H.-P. Gauggel, H. P. Zappe, M. Moser, R. Hovel, and E. Kapon, "Continuous-Wave Operation of Phase-Coupled Vertical-Cavity Surface-Emitting Laser Arrays," *Applied Physics Letters*, Vol. 77, No. 15, pp. 2283-2285, October 2000.

[366] J. W. Scott, R. S. Geels, S. W. Corzine, and L. A. Coldren, "Modeling Temperature Effects and Spatial Hole Burning to Optimize Vertical-Cavity Surface-Emitting

Laser Performance," *IEEE Journal of Quantum Electronics*, Vol. 29, No. 5, pp. 1295-1308, May 1993.

[367] G. R. Hadley, K. L. Lear, M. E. Warren, K. D. Choquette, J. W. Scott, and S. W. Corzine, "Comprehensive Numerical Modeling of Vertical-Cavity Surface-Emitting Lasers," *IEEE Journal of Quantum Electronics*, Vol. 32, No. 4, pp. 607-616, April 1996.

[368] B. Klein, L. Register, M. Grupen, and K. Hess, "Numerical Simulation of Vertical Cavity Surface Emitting Lasers," *Optical Express*, Vol. 2, No. 4, pp. 163-168, February 1998.

[369] P. V. Mena, J. J. Morikuni, S.-M. Kang, A. V. Harton, and K. W. Wyatt, "A Comprehensive Circuit-Level Model of Vertical-Cavity Surface-Emitting Lasers," *Journal of Lightwave Technology*, Vol. 17, No. 12, pp. 2612-2632, December 1999.

[370] J. Mulet and S. Balle, "Spatio-Temporal Modeling of the Optical Properties of VCSELs in the Presence of Polarization Effects," *IEEE Journal of Quantum Electronics*, Vol. 38, No. 3, pp. 291-305, March 2002.

[371] M. Streiff, A. Witzig, M. Pfeiffer, P. Royo, and W. Fichtner, "A Comprehensive VCSEL Device Simulator," *IEEE of Journal Selected Topics in Quantum Electronics*, Vol. 9, No. 3, pp. 879-891, May/June 2003.

[372] M. X. Jungo, D. Erni, and W. Bachtold, "VISTAS: A Comprehensive System-Oriented Spatiotemporal VCSEL Model," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 9, No. 3, pp. 939-948, May/June 2003.

[373] J. S. Gustavsson, A. Haglund, J. Bengtsson, P. Modh, and A. Larsson, "Dynamic Behavior of Fundamental-Mode Stabilized VCSELs Using a Shallow Surface Relief," *IEEE Journal of Quantum Electronics*, Vol. 40, No. 6, pp. 607-619, June 2004.

[374] R. P. Sarzala, P. Mendla, M. Wasiak, P. Mackowiak, W. Nakwaski, and M. Bugajski, "Three-Dimensional Comprehensive Self-Consistent Simulation of a Room-Temperature Continuous-Wave Operation of GaAs-Based 1.3-μm Quantum-Dot (InGa)As/GaAs Vertical-Cavity Surface-Emitting Lasers," *Proceedings of the International Conference on Transparent Optical Networks*, Vol. 2, pp. 95-98, June/July 2003.

[375] P. Debernardi, "HOT-VELM: A Comprehensive and Efficient Code for Fully Vectorial and 3-D Hot-Cavity VCSEL Simulation," *IEEE Journal of Quantum Electronics*, Vol. 45, No. 8, pp. 979-992, August 2009.

[376] M. Osinski and W. Nakwaski, "Thermal Analysis of Closely-Packed Two-Dimensional Etched-Well Surface-Emitting Laser Arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 1, No. 2, pp. 681-696, June 1995.

[377] S. W. Corzine, R. H. Yan, and L. A. Coldren, "A Tanh Substitution Technique for the Analysis of Abrupt and Graded Interface Multilayer Dielectric Stacks," *IEEE Journal of Quantum Electronics*, Vol. 27, No. 9, pp. 2086-2090, September 1991.

[378] G. M. Yang, M. H. MacDougal, V. Pudikov, and P. D. Dapkus, "Influence of Mirror

Reflectivity on Laser Performance of Very-Low-Threshold Vertical-Cavity Surface-Emitting Lasers," *IEEE Photonics Technology Letters*, Vol. 7, No. 11, pp. 1228-1230, November 1995.

[379] T. A. DeTemple and C. M. Herzinger, "On the Semiconductor Laser Logarithmic Gain-Current Density Relation," *IEEE Journal of Quantum Electronics*, Vol. 29, No. 5, pp. 1246-1252, May 1993.

[380] C. Zhu, Y.-G. Zhang, A.-Z. Li, and Z.-B. Tian, "Analysis of Key Parameters Affecting the Thermal Behavior and Performance of Quantum Cascade Lasers," *Journal of Applied Physics*, Vol. 100, No. 5, pp. 053105-1 - 053105-6, September 2006.

[381] R. Mehandru, G. Dang, S. Kim, F. Ren, W. S. Hobson, J. Lopata, S. J. Pearton, W. Chang, and H. Shen, "Finite Difference Analysis of Thermal Characteristics of CW Operation 850 nm Lateral Current Injection and Implant-Apertured VCSEL with Flip-Chip Bond Design," *Solid-State Electronics: An Internation Journal*, Vol. 46, No. 5, pp. 699-704, May 2002.

[382] A. Sarua, J. Hangfeng, K. P. Hilton, D. J. Wallis, M. J. Uren, T. Martin, and M. Kuball, "Thermal Boundary Resistance Between GaN and Substrate in AlGaN/GaN Electronic Devices," *IEEE Transactions on Electron Devices*, Vol. 54, No. 12, pp. 3152-3158, December 2007.

[383] J. A. Quintana and J. R. Rodriguez, "Interfacial Effects on the Thermal Conductivity of a-Ge Thin Films Grown on Si Substrates," *Journal of Applied Physics*, Vol. 104, No. 7, pp. 074903-1 - 074903-4, October 2008.

[384] R. Mackenzie, J. Lim, S. Bull, S. Sujecki, and E. Larkins, "Inclusion of Thermal Boundary Resistance in the Simulation of High-Power 980nm Ridge Waveguide Lasers," *Optical and Quantum Electronics*, Vol. 40, pp. 373-377, 2008.

[385] M. Osinski and W. W. Nakwaski, "Effective Thermal Conductivity Analysis of 1.55 μm InGaAsP/InP Vertical-Cavity Top-Surface-Emitting Microlasers," *Electronics Letters*, Vol. 29, No. 11, pp. 1015-1016, May 1993.

[386] A. Balandin and K. L. Wang, "Significant Decrease of the Lattice Thermal Conductivity Due to Phonon Confinement in a Free-Standing Semiconductor Quantum Well," *Physical Review B*, Vol. 58, No. 3, pp. 1544-1549, Jul 1998.

[387] A. Khitun, A. Balandin, J. L. Liu, and K. L. Wang, "In-Plane Lattice Thermal Conductivity of a Quantum-Dot Superlattice," *Journal of Applied Physics*, Vol. 88, No. 2, pp. 696-699, July 2000.

[388] W. Nakwaski, "Thermal Conductivity of Binary, Ternary, and Quaternary III-V Compounds," *Journal of Applied Physics*, Vol. 64, No. 1, pp. 159-166, July 1988.

[389] M. Guden and J. Piprek, "Material Parameters of Quaternary III-V Semiconductors for Multilayer Mirrors at 1.5 μm Wavelength," *Modelling and Simulation in Materials Science and Engineering*, Vol. 4, No. 4, pp. 349-357, April 1996.

[390] L. Piskorski, R. P. Sarzala, and W. Nakwaski, "Self-Consistent Model of 650 nm GaInP/AlGaInP Quantum-Well Vertical-Cavity Surface-Emitting Diode Lasers," *Semiconductor Science and Technology*, Vol. 22, No. 6, pp. 593-600, June 2007.

[391] C. Hu, M. Kiene, and P. S. Ho, "Thermal Conductivity and Interfacial Thermal Resistance of Polymeric Low k Films," *Applied Physics Letters*, Vol. 79, No. 25, pp. 4121-4123, December 2001.

[392] R. Rossbach, T. Ballmann, R. Butendeich, H. Schweizer, F. Scholz, and M. Jetter, "Red VCSEL for High-Temperature Applications," *Journal of Crystal Growth*, Vol. 272, No. 1-4, pp. 549-554, December 2004.

[393] A. K. Coskun, J. L. Ayala, D. Atienza, T. S. Rosing, and Y. Leblebici, "Dynamic Thermal Management in 3D Multicore Architectures," *Proceedings of the IEEE Design, Automation & Test in Europe Conference and Exhibition*, pp. 1–6, April 2009.

[394] K. Puttaswamy and G. H. Loh, "Scalability of 3D-Integrated Arithmetic Units in High-Performance Microprocessors," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 622–625, June 2007.

[395] S. Im and K. Banerjee, "Full Chip Thermal Analysis of Planar (2-D) and Vertically Integrated (3-D) High Performance ICs," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 727–730, December 2000.

[396] M.-C. Hsieh, C.-K. Yu, and C. H. Rd, "Thermo-Mechanical Simulations for 4-Layer Stacked IC Packages," *Proceedings of the International Conference on Thermal, Mechanical and Multi-Physics Simulation and Experiments in Microelectronics and Micro-Systems*, pp. 1–7, April 2008.

[397] L. Jiang, S. Kolluri, B. J. Rubin, H. Smith, E. G. Colgan, M. R. Scheuermann, J. A. Wakil, A. Deutsch, and J. Gill, "Thermal Modeling of On-Chip Interconnects and

3D Packaging Using EM Tools," *Proceedings of the IEEE Electrical Performance of Electronic Packaging*, pp. 279–282, October 2008.

[398] A. Jain, R. E. Jones, and S. Pozder, "Thermal Modeling and Design of 3D Integrated Circuits," *Proceedings of the Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pp. 1139–1145, May 2008.

[399] W. Huang, M. R. Stan, K. Skadron, K. Sankaranarayanan, S. Ghosh, and S. Velusamy, "Compact Thermal Modeling for Temperature-Aware Design," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 878–883, June 2004.

[400] W. Huang, S. Ghosh, S. Velusamy, K. Sankaranarayanan, K. Skadron, and M. R. Stan, "HotSpot: A Compact Thermal Modeling Methodology for Early-Stage VLSI Design," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 5, pp. 501–513, May 2006.

[401] K. Balakrishnan, V. Nanda, and S. Easwar, "Wire Congestion and Thermal Aware 3D Global Placement," *Proceedings of the IEEE Asia and South Pacific Design Automation Conference*, pp. 1131–1134, January 2005.

[402] K. Puttaswamy and G. H. Loh, "Thermal Herding: Microarchitecture Techniques for Controlling Hotspots in High-Performance 3D-Integrated Processors," *Proceedings of the IEEE International Symposium on High Performance Computer Architecture*, pp. 193–204, February 2007.

[403] J.-M. Koo, S. Im, L. Jiang, and K. E. Goodson, "Thermal Characterization of Interlayer Microfluidic Cooling of Three-Dimensional Integrated Circuits With

Nonuniform Heat Flux," *Journal of Heat Transfer*, Vol. 132, No. 4, pp. 041009-1–041009-9, April 2010.

[404] C. R. King, D. Sekar, M. S. Bakir, B. Dang, J. Pikarsky, and J. D. Meindl, "3D Stacking of Chips with Electrical and Microfluidic I/O Interconnects," pp. 1–7, May 2008.

[405] Y. J. Kim, Y. K. Joshi, A. G. Fedorov, Y.-J. Lee, and S.-K. Lim, "Thermal Characterization of Interlayer Microfluidic Cooling of Three-Dimensional Integrated Circuits With Nonuniform Heat Flux," *Journal of Heat Transfer*, Vol. 132, No. 4, pp. 041009-1 – 041009-9, April 2010.

[406] M. M. Sabry, D. Atienza, and A. K. Coskun, "Thermal Analysis and Active Cooling Management for 3D MPSoCs," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 2237–2240, May 2011.

[407] H. Oprins, V. O. Cherman, B. Vandevelde, G. Van der Plas, P. Marchal, and E. Beyne, "Numerical and Experimental Characterization of the Thermal Behavior of a Packaged DRAM-on-Logic Stack," *Proceedings of the IEEE Electronic Components and Technology Conference*, pp. 1081-1088, June 2012.

[408] M. El-Moursy and E. Friedman, "Power Characteristics of Inductive Interconnect," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 12, pp. 1295–1306, December 2004.

[409] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Exploiting the On-Chip Inductance in High-Speed Clock Distribution Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 9, No. 6, pp. 963–973, December 2001.

[410] V. Adler and E. G. Friedman, "Uniform Repeater Insertion in *RC* Trees," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 47, No. 10, pp. 1515–1523, October 2000.

[411] Y. I. Ismail and E. G. Friedman, "Effects of Inductance on the Propagation Delay and Repeater Insertion in VLSI Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 8, No. 2, pp. 195–206, April 2000.

[412] W. Lu, P. Xie, and C. M. Lieber, "Nanowire Transistor Performance Limits and Applications," *IEEE Transactions on Electron Devices*, Vol. 55, No. 11, pp. 2859–2876, November 2008.

[413] K. Novoselov, "Graphene Electronics, Unzipped," *IEEE Spectrum*, Vol. 47, No. 11, pp. 28–33, November 2010.

[414] E. Salman, H. Akin, O. Gursoy, A. Ergintav, I. Tekin, A. Bozkurt, and Y. Gurbuz, "Design of a 3.2 mW PLL Based Clock and Data Recovery Circuit in 90 nm CMOS Technology," *Proceedings of the Mediterranean Microwave Symposium*, September 2005.

[415] I. S. Kourtev, B. Taskin, and E. G. Friedman, Eds., *Timing Optimization Through Clock Skew Scheduling*, *Second Edition*, Springer Science+Business Media, 2009.

[416] I. S. Kourtev and E. G. Friedman, "Synthesis of Clock Tree Topologies to Implement Nonzero Clock Skew Schedule," *IEE Proceedings-Circuits, Devices and Systems*, Vol. 146, No. 6, pp. 321–326, December 1999.

[417] I. Savidis, E. G. Friedman, V. F. Pavlidis, and G. De Micheli, "Clock and Power Distribution Networks for 3-D Integrated Circuits," *Proceedings of the Workshop on 3D Integration, Design, Automation & Test in Europe Conference*, pp. 1-2, March 2009.

[418] I. Savidis, S. Kose, and E. G. Friedman, "Power Grid Noise in TSV-Based 3-D Integrated Systems," *Proceedings of the Government Microcircuit Applications & Critical Technology Conference*, pp. 129–132, March 2011.

[419] V. Kursun, S. Narendra, V. K. De, and E. G. Friedman, "Analysis of Buck Converters for On-Chip Integration With a Dual Supply Voltage Microprocessor," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 11, No. 3, pp. 514–522, June 2003.

[420] V. Kursun and E. G. Friedman, *Multi-Voltage CMOS Circuit Design*, John Wiley & Sons Press, 2006.

[421] E. G. Friedman, "Small Area Power Converter for Application to Distributed On-Chip Power Delivery," *Proceedings of the International Workshop on Power Supply On Chip*, p. 41, October 2010.

[422] S. M. Alam, R. E. Jones, S. Pozder, R. Chatterjee, S. Rauf, and A. Jain, "Interstratum Connection Design Considerations for Cost-Effective 3-D System

Integration," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 18, No. 3, pp. 450–460, March 2010.

[423] V. F. Pavlidis and E. G. Friedman, "3-D Topologies for Networks-on-Chip," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 15, No. 10, pp. 1081–1090, October 2007.

[424] W. J. Dally and B. Towles, *Principles and Practices of Interconnection Networks*, Morgan Kaufmann Publishers, 2004.

[425] A. Jantsch and H. Tenhunen, *Networks on Chip*, Springer Science + Business Media, Inc., 2004.

[426] K. Haghdad, M. Anis, and Y. Ismail, "Floorplanning for Low Power IC Design Considering Temperature Variations," *Microelectronics Journal*, Vol. 42, No. 1, pp. 89–95, January 2011.

[427] J.-M. Lin and Y.-W. Chang, "TCG: A Transitive Closure Graph-Based Representation for General Floorplans," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 13, No. 2, pp. 288–292, February 2005.

[428] W.-L. Hung, G. Link, Y. Xie, N. Vijaykrishnan, and M. Irwin, "Interconnect and Thermal-Aware Floorplanning for 3D Microprocessors," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 98–104, March 2006.

[429] Z. Li, "Efficient Thermal Via Planning Approach and Its Application in 3-D Floorplanning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. E38-A, No. 4, pp. 53–658, April 2007.

[430] S. Dong, R. Wang, F. Guo, J. Yuan, and X. Hong, "Floorplanning by a Revised 3-D Corner Block List with sub-C^+-Tree," *Proceedings of the Joint Conference on Information Sciences*, pp. 3–6, October 2006.