Physical Resource Allocation for On-Chip Power Delivery Systems

by

Renatas Jakushokas

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
2011

Dedication

This work is dedicated to my wife, Victoria Jakushokas, and my son, Daniel Noam Jakushokas.

Curriculum Vitae

Renatas Jakushokas was born in Kaunas, Lithuania. He received the B.Sc. degree (with honors) in electrical engineering from Ort-Braude College, Karmiel, Israel in 2005, and the M.S. degree in electrical and computer engineering from the University of Rochester, Rochester, New York in 2007. He is currently completing the Ph.D. degree in electrical engineering at the University of Rochester.

He was previously an intern at Intrinsix Corporation, Fairport, New York in 2006, working on Sigma Delta ADCs. In the summer of 2007, he interned with Eastman Kodak Company, Rochester, New York, where he designed a high speed, precision comparator for high performance ADCs. During the summer of 2008, he interned

at Freescale Semiconductor Corporation, Tempe, Arizona, where he worked on developing a noise coupling estimation calculator, permitting the efficient evaluation of diverse substrate isolation techniques. His research interests are in the areas of power distribution, noise evaluation, signal integrity, substrate modeling/analysis, and optimization techniques for high performance integrated circuit design.

Publications

Authored Book

1. R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Kose, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors, Second Edition,* Springer Publishers, 2011.

Journal Papers

2. R. Jakushokas and E. G. Friedman, "Multi-Layer Interdigitated Power Distribution Networks," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* (in press).
3. R. Jakushokas and E. G. Friedman, "Resource Based Optimization for Simultaneous Shield and Repeater Insertion," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 18, No. 5, pp. 742-749, May 2010.
4. E. Salman, R. Jakushokas, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Methodology for Efficient Substrate Noise Analysis in Large Scale Mixed-Signal Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems,* Vol. 17, No. 10, pp. 1405-1418, October 2009.
5. R. Jakushokas and E. G. Friedman, "Inductance Model of Interdigitated Power and Ground Distribution Networks," *IEEE Transactions on Circuits and Systems II: Express Briefs,* Vol. 56, No. 7, pp. 585-589, July 2009.

Conference Papers

6. R. Jakushokas, E. Salman, E. G. Friedman, R. M. Secareanu, O. L. Hartin, and C. L. Recker, "Compact Substrate Models for Efficient Noise Coupling and Signal Isolation Analysis," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 2346 - 2349, May/June 2010.
7. R. Jakushokas and E. G. Friedman, "Globally Integrated Power and Clock Distribution Network," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 1751 - 1754, May/June 2010.
8. R. Jakushokas and E. G. Friedman, "Methodology for Multi-Layer Interdigitated Power and Ground Network Design," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 3208 - 3211, May/June 2010.
9. R. Jakushokas and E. G. Friedman, "Line Width Optimization for Interdigitated Power/Ground Networks," *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI,* pp. 329 - 334, May 2010.
10. R. Jakushokas and E. G. Friedman, "Minimizing Noise via Shield and Repeater Insertion," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 2265 - 2268, May 2009.
11. E. Salman, R. Jakushokas, E. G. Friedman, R. M. Secareanu, and O L. Hartin, "Contact Merging Algorithm for Efficient Substrate Noise Analysis in Large Scale Circuits," *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI,* pp. 9 -14, May 2009.
12. R. Jakushokas and E. G. Friedman, "Simultaneous Shield and Repeater Insertion," *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI,* pp. 15 - 19, May 2009.
13. E. Salman, R. Jakushokas, E. G. Friedman, R. M. Secareanu, and O. L. Hartin, "Input Port Reduction for Efficient Substrate Extraction in Large Scale IC's," *Proceedings of the IEEE International Symposium on Circuits and Systems,* pp. 376 - 379, May 2008.

Acknowledgments

My sincere gratitude is given to my advisor and mentor, Professor Eby G. Friedman for his continuous support, encouragement, and direction during the course of my entire Ph.D. study. His knowledge, experience, insight, and enthusiasm instilled in me a desire to originate and pursue my research studies. It was a great pleasure for me to accomplish my thesis work under his supervision. What I know today about the process of research, I learned from Professor Eby G. Friedman.

I would like to thank Professor Roman Sobolewski, Professor Zeljko Ignjatovic, and Professor Peter Salzman for serving on my proposal and defense committees, and Professor Michael L. Scott for chairing the defense committee. Your insightful comments, critical reviews, and helpful advice enhanced my research and filled important gaps in my thesis dissertation. I want to also express my warm thanks to Dr. Radu M. Secareanu, Dr. Olin L. Hartin, and Dr. Cynthia L. Recker from Freescale Semiconductor, Dr. Mike Guidash, John Campton, and Dr. Weize Xu from Eastman Kodak, and Dr. Muchadit Kozak and Gene Petilli from Intrinsix Corportion for your guidance during three insightful summer internships.

The diversity of professional and cultural backgrounds within the High Performance Integrated Circuit Laboratory made my experience highly interesting and joyful. I would like to thank previous and current members of the laboratory: Dr. Guoqing Chen, Dr. Mikhail Popovich, Dr. Vasilis F. Pavlidis, Prof. Emre Salman, Dr. Jonathan Rosenfeld, Ioannis Savidis, Selcuk Kose, Inna Vaisband, and Ravi Patel. The long discussions on a variety of topics have expanded the scope of my professional and personal knowledge. I thank RuthAnn Williams for her endless effort on my behalf, allowing me to focus my time on my research. I would also like to acknowledge the staff in the Department of Electrical and Computer Engineering for their assistance and support.

I express my thanks to all of my friends, who made my life in Rochester full of joy and happiness. I will newer forget these amazing days. Most important, I owe a special thanks to my parents, my wife's parents, my wife Victoria, and my son Daniel for their love, unconditional support, and endless encouragements. Without them, I would never be in the place I am now.

Abstract

The technology behind integrated circuits is growing rapidly with billions of devices integrated on the same die. These devices operate at several gigahertz and require tens of watts, with voltage levels below a volt. Highly complicated on-chip networks manage and support the operation of these billions of devices. Resources, such as metal, power, and area, are however limited; these resources must be efficiently utilized. The increase in the number of metal layers within an integrated circuit does not keep up with device scaling, creating challenges in global signaling, synchronization, and power delivery. The objective is to address design, analysis, and optimization challenges for highly complicated structures. Power distribution networks, global signal networks, and monolithic substrate are considered in this dissertation.

An effective impedance model of a monolithic substrate is developed within this dissertation, achieving high accuracy in estimating power/ground noise characteristics. A methodology for simultaneously inserting shields and repeaters is described,

optimizing multiple resources for global signal interconnects. A closed-form model of the self- and mutual inductance of an interdigitated power and ground distribution network is described, providing less than 5% error for a typical power distribution network. The optimal width of the metal lines that minimizes the impedance of the power distribution network is determined, significantly enhancing the performance of an integrated circuit. A design methodology is also described for a multi-layer power distribution network, achieving enhanced reliability by equalizing the current density over multiple metal layers. Furthermore, a novel link breaking methodology for a mesh structured power distribution network is introduced, reducing coupling noise while improving the maximum operating frequency, on average, by 12%. Finally, a globally integrated power and clock distribution network is presented which utilizes a single network to distribute both global signals; thereby reducing the metal requirement.

The performance of integrated circuits is highly affected by the power delivery system. The primary focus of this dissertation is the development of design and analysis methodologies for on-chip power delivery systems. Integrated circuits developed with these novel design methodologies will provide higher performance, while simultaneously consuming less power, area, and metal.

Contents

List of Tables

List of Figures

Foreword

The author, R. Jakushokas, has developed all of the design and analysis methodologies, derived the closed-form models, and evaluated all of the results described in this Ph.D. dissertation. The introduction (Chapter 1) and background chapters (Chapters 2 and 3) are based on the academic literature, presented by other researchers. Other contributions from colleagues are listed below.

Chapter 4: R. Jakushokas is the primary author of this chapter, which has been submitted for publication in the *IEEE Transactions on Circuits and Systems II: Express Briefs*. The development and evaluation of the research have been performed with co-authors E. Salman and E. G. Friedman. This research has also been conducted in collaboration with co-authors R. M. Secareanu, O. L. Hartin, and C. L. Recker from Freescale Semiconductor Corporation. Part of this work has also been published in the *Proceedings of the IEEE International Symposium on Circuits and Systems* with the aforementioned co-authors.

Chapter 5: R. Jakushokas is the primary author of this chapter, which has been published in the *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*.

The development and evaluation of the research have been performed with the co-author of the paper, Eby G. Friedman. Part of this work has also been published in the *Proceedings of the IEEE International Symposium on Circuits and Systems* and the *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI* with the aforementioned co-author.

Chapter 6: R. Jakushokas is the primary author of this chapter, which has been published in the *IEEE Transactions on Circuits and Systems II: Express Briefs*. The development and evaluation of the research have been performed with the co-author, E. G. Friedman. Part of this work has also been published in the *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI* with the aforementioned co-author.

Chapter 7: R. Jakushokas is the primary author of this chapter, which has been published in the *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*. The development and evaluation of the research have been performed with the co-author, E. G. Friedman. Part of this work has also been published in the *Proceedings of the IEEE International Symposium on Circuits and Systems* with the aforementioned co-author.

Chapter 8: R. Jakushokas is the primary author of this chapter, which has been submitted for publication in the *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*. The development and evaluation of the research have been performed with the co-author, E. G. Friedman.

Chapter 9: R. Jakushokas is the primary author of this chapter, which has been published in the *Proceedings of the IEEE International Symposium on Circuits and Systems*. The development and evaluation of the research have been performed with the co-author, E. G. Friedman.

Chapter 1

Introduction

Although the battery, light bulb, and electricity itself were invented much earlier, only after discovery of the electron in 1897 by Joseph J. Thomson [1] did the era of electronics begin. This discovery led to the understanding of current flow, charge storage, and related phenomena. At the time, Joseph J. Thomson worked at Cambridge University in the United Kingdom and was experimenting with cathode tubes, placing these tubes within electric and magnetic fields. The cathode rays bent, suggesting that the rays are composed of small particles. Joseph J. Thomson thought that his invention would never leave the science labs. However, as all of us know today, his invention completely changed our lives. Joseph J. Thomson received the Nobel prize for his work in 1906 [2].

> *"Excuse me... how can you discover a particle so small that nobody has ever seen one?" – J. J. Thomas, 1987 [3].*

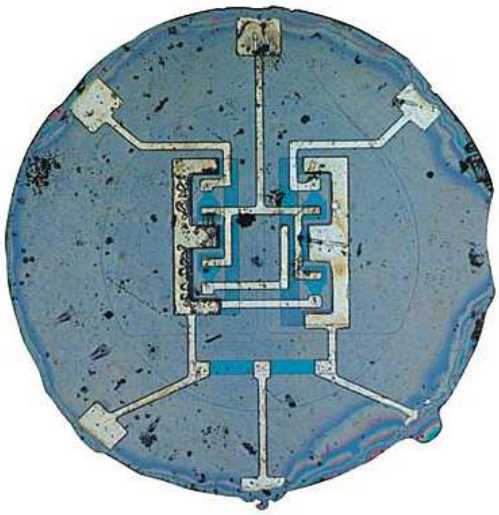
From discovery of the electron to today's modern computers, satellite radios, and cell phones, there have been many fundamental discoveries. A primary invention of seminal importance is the transistor, a multi-terminal device capable of producing a large change in output voltage with a small change in input voltage. This fundamental function is called amplification. The first transistor, depicted in Fig. 1.1(a), was fabricated by Bell Laboratories, specifically by Williams Shockley, John Bardeen, and Walter Brattain in 1947. All three scientists received the Nobel prize for physics in 1956 [2].

> *"This circuit was actually spoken over and by switching the device in and out a distinct gain in speech level could be heard and seen on the scope presentation with no noticeable change in quality." – W. Brattain, 1947 [4].*

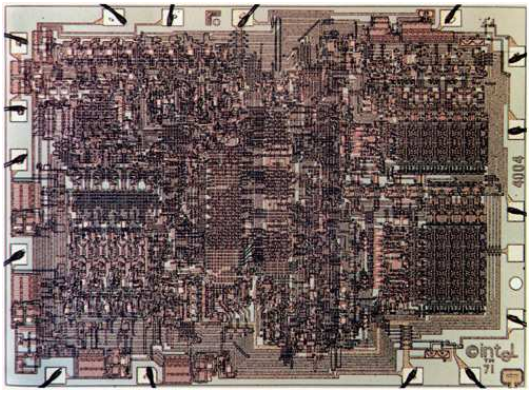
The first planar transistor, shown in Fig. 1.1(b), was fabricated by Fairchild Semiconductor Company in 1959, where Robert Noyce worked, initiating the integrated circuit (IC) era. The first commercially available microprocessor, illustrated in Figs. 1.1(c) and 1.1(d), was fabricated by Intel in 1971.

> *"This invention relates to electrical circuit structures incorporating semiconductor devices. Its principal objects are these: to provide improved device-and-lead structures for making electrical connections to the various semiconductor regions; to make unitary circuit structures more compact*

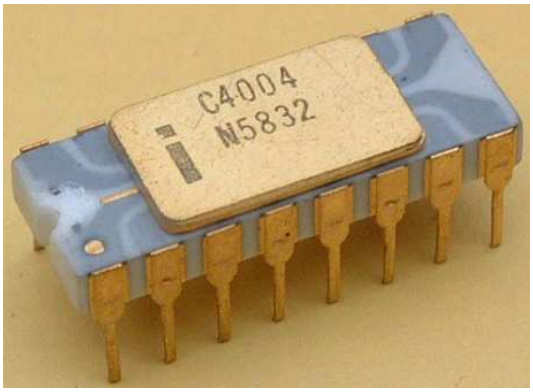
(a)



(b)



(c)



(d)

Figure 1.1: Fundamental inventions in the IC era. (a) First transistor fabricated by Bells Laboratories in 1947; (b) First planar transistor fabricated by Fairchild Semiconductor Company in 1959; (c) Layout view and (d) Package of the first commercially available microprocessor fabricated by Intel in 1971.

and more easily fabricated in small sizes than has heretofore been feasible; and to facilitate the inclusion of numerous semiconductor devices within a single body of material." – R. Noyce, 1959 [5].

Since the first IC, the demand for integrated circuits has been increasing. The push for greater on-chip functionality still continues today. According to Intel's co-founder, George Moore, IC complexity will double every IC generation, known as Moore's law. Microprocessors continue to maintain this growth. The number of transistors per single Intel microprocessor is illustrated in Fig. 1.2.

"At present [1965], it [minimum cost] is reached when 50 components are used per circuit. ...by 1975, the number of components per integrated circuit for minimum cost will be 65,000." – G. Moore, 1965 [6].

1.1 Global Interconnect: Friend or Foe

With increasing numbers of components on a single IC, the number of interconnects is also growing. There are three basic types of interconnects: local, semi-global, and global. Wires connecting two gates in the same circuit are typically referred to as local interconnects. Global interconnects are long interconnects, connecting different parts of an IC, for purposes such as distributing clock signals or delivering power.

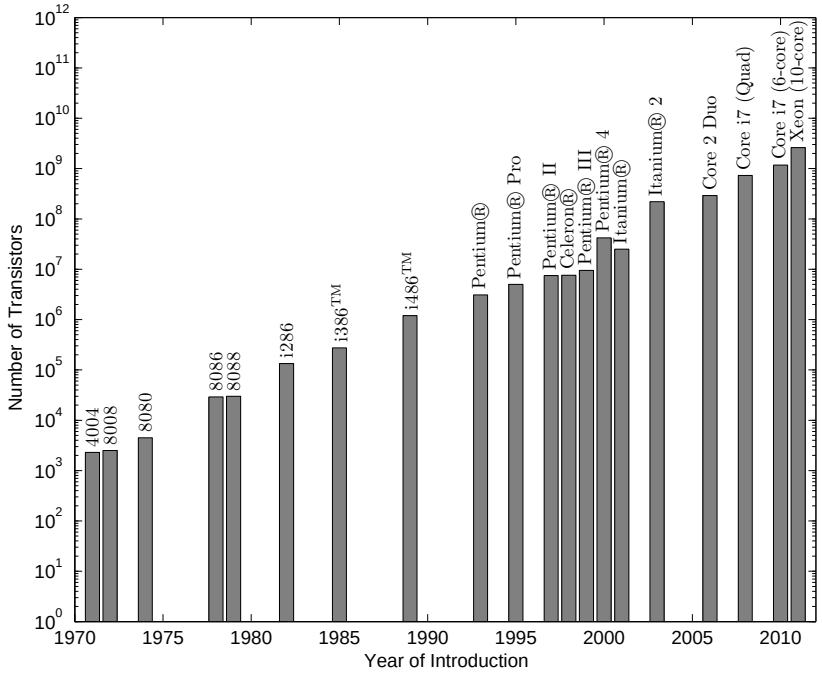


Figure 1.2: Transistor count for Intel microprocessors.

Semi-global interconnects are wires of medium length, between the length of local and global interconnects. The following discussion focuses on the global interconnects.

All kinds of global interconnect are used on-chip. A buss interconnect structure utilizes global signal interconnects to provide a common communication protocol within a system. A synchronous clock signal is distributed over an entire IC using a complex interconnect network composed of multiple global interconnects. Power and ground supplies are delivered to almost every transistor in an IC, utilizing global interconnects. The global interconnect is an absolutely essential element of an integrated

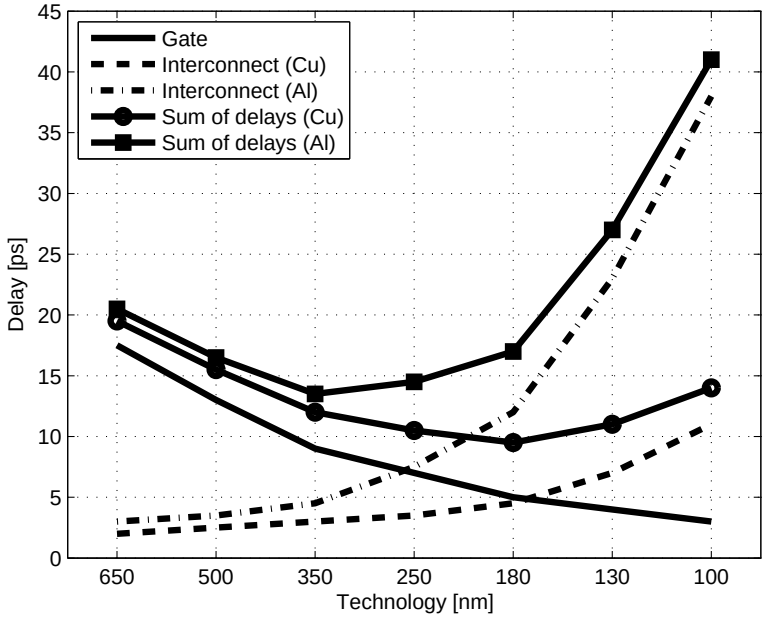


Figure 1.3: Interconnect vs. gate delay for different technologies [7]

circuit.

In advanced technologies, the interconnect delay is often greater than the gate delay, making the interconnect a primary performance bottleneck [7]. This phenomenon is illustrated in Fig. 1.3 for several technology nodes. Interconnect delay can be accurately modeled by dividing the interconnect into shorter segments. The tradeoff is between reducing the computational complexity with a smaller number of segments, and increasing the delay accuracy with a larger number of segments. Representing the interconnect as a single segment is referred to as a lumped interconnect model, while representing the interconnect as an infinite (or large) number of segments is

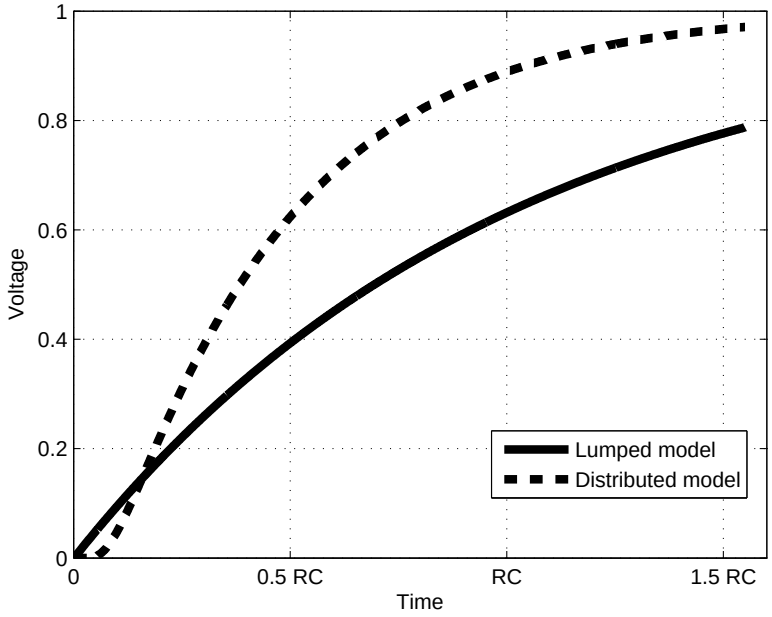


Figure 1.4: Step input voltage response of an interconnect using a lumped and distributed *RC* interconnect model. The accuracy of a signal delay is higher with a greater number of segments.

referred to as a distributed model. The distributed model with an infinite number of segments more accurately represents the actual signal delay. The accuracy decreases with a fewer segments. The step input voltage response of an interconnect using a lumped and distributed *RC* interconnect model is depicted in Fig 1.4.

The power consumed by the on-chip interconnects is typically between 20% and 60% of the total power dissipated by an IC [8, 9]. A focus on the global interconnects is therefore necessary. The wire pitch of the global interconnect, illustrated in Fig. 1.5,

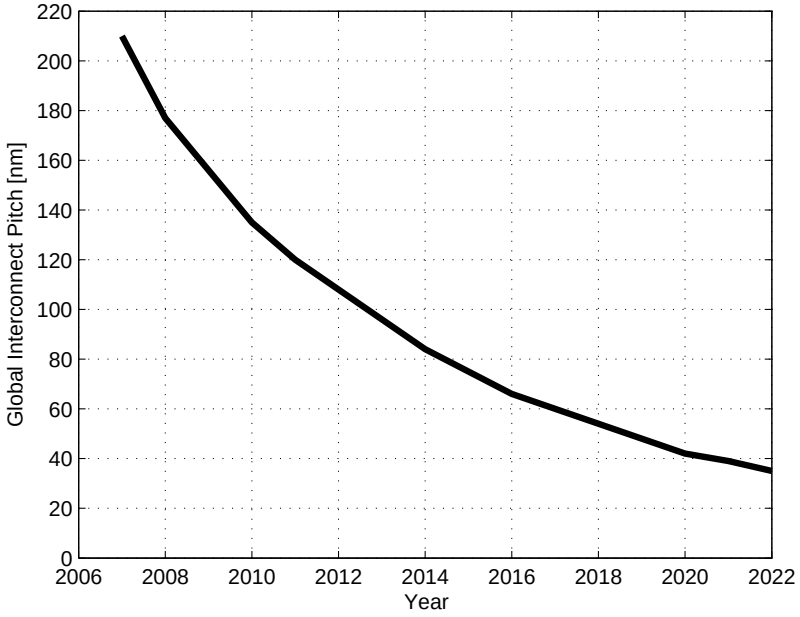


Figure 1.5: Minimum pitch of global interconnect [7].

scales drastically [7], further increasing the delay and power of an interconnect. Techniques to overcome these power and delay limitations are therefore required. Accurate and computationally efficient models are also necessary to characterize these global interconnect structures to determine the behavior of a system.

1.2 Dissertation Outline

This dissertation is organized as follows. Several design techniques that enhance circuit performance are summarized in Chapter 2. The focus of this chapter is to

describe these different design techniques at several levels of abstraction. Some commonly used design techniques, such as repeaters, shielding, and decoupling capacitors, are discussed in greater detail. Note that each design technique utilizes multiple and different design resources.

The analysis and optimization of integrated circuits rely on models at different abstraction levels. Modeling of power distribution networks and the monolithic substrate is discussed in Chapter 3. For power distribution networks, the focus is on commonly used network structures; specifically, mesh and interdigitated structures. A reduced order model for mesh structures is reviewed in this chapter. The primary advantage of an interdigitated structure is a reduction in the on-chip inductance; inductance model for interdigitated structures are therefore described. For the monolithic substrate, two noise estimation methods are reviewed. One method is based on solving differential equations utilizing either the boundary element method or finite difference method. Another method is based on compact models. These models consider a lightly doped substrate and an epi-type substrate with a grounded back side metal.

In Chapter 4, current propagation within a lightly doped substrate is approximated with a half-ellipse to efficiently estimate substrate resistances. In contrast to existing work, the proposed model utilizes only one fitting parameter. Compact models are also developed to determine the isolation efficiency of several commonly

used structures, such as a guard ring and triple well. The accuracy of these models is verified by comparing the models with a commercial substrate extraction tool that utilizes a boundary element method. These models are used to compare several isolation structures within an industrial mixed-signal circuit with a lightly doped substrate.

Two techniques, shield and repeater insertion, are simultaneously investigated in Chapter 5. To optimize resources, the relationship among noise, power, and delay is developed. Coupling noise as a function of power dissipation is shown to behave parabolically. Based on this parabolic behavior, the minimum noise can be established. The resulting design expressions are compared with SPICE, exhibiting good agreement. A design case is compared with only shielding and only repeater insertion, suggesting that simultaneous shield and repeater insertion can achieve enhanced performance.

Higher operating frequencies have increased the importance of inductance in power and ground networks. A closed-form expression is presented in Chapter 6 to accurately estimate the effective inductance of a single layer within an interdigitated power and ground distribution network. Due to the large number of power and ground lines in power networks, excessive time is required to calculate the inductance using 3-D simulation tools. The proposed expression is compared with previous models and FastHenry, exhibiting accurate and computationally efficient results. The inductance

of a single layer within an interdigitated power and ground distribution network is bounded for any number of parallel lines. The error of the proposed expression decreases rapidly with increasing numbers of pairs within the network. The upper bound for the error of the proposed model is also determined. Based on the resistive and inductive (both self- and mutual) impedance, a closed-form expression for the optimal power and ground wire width is described, producing the minimum impedance for a single metal layer. The optimal wire width is determined under different physical network dimensions and signal frequencies, suggesting useful trends for interdigitated power and ground networks.

An interdigitated power and ground distribution network for a multi-layer metal system is described in Chapter 7. Higher frequencies are increasing the importance of the inductance, while larger current loads increase the current densities, making electromigration an important design issue. In Chapter 7, methods for optimizing a multi-layer interdigitated power and ground network are presented. The effect of electromigration is considered, permitting the appropriate number of metal layers to be determined. The tradeoff between the network impedance and current density is investigated. Based on 65 nm, 45 nm, and 32 nm CMOS technologies, the optimal width as a function of metal layer is determined for different frequencies. In addition, the effect of routability, frequency range, and decoupling capacitors on multi-layer interdigitated power and ground distribution networks are discussed.

In Chapter 8, a link breaking methodology is discussed to reduce voltage degradation within a mesh structured power distribution network. The resulting power distribution network is a combination of a single power distribution network to lower the network impedance, and multiple networks to reduce noise coupling among the circuits. Since the sensitivity to supply voltage variations within a power distribution network can vary among different circuits, the proposed methodology reduces the voltage drop at the more sensitive circuits, while penalizing the less sensitive circuits. Each circuit within an IC can behave as an aggressor as well as a victim. The methodology therefore utilizes two matrices, describing, respectively, the aggressiveness and sensitivity of the circuits. This methodology is evaluated for multiple case studies, reducing the voltage drop in the sensitive circuits (the critical paths). The higher supply voltage produces a faster delay within the critical paths. Based on these case studies, an average enhancement of 12% in the maximum operating frequency is achieved utilizing this link breaking methodology.

The global networks within a conventional integrated circuits (IC) consists of three major types: power, ground, and clock distribution networks. These three networks consume most of the metal resources in the highest metal layers. The signals traversing the power and clock distribution networks are fundamentally different in terms of signal frequency and current flow. Combining the power and clock network into a globally integrated network is therefore possible. In Chapter 9, this general concept

of a globally integrated power and clock (GIPAC) system is reviewed. The circuitry supporting this GIPAC system is also presented. Simulation results presented in this chapter are based on a 90 nm CMOS technology and demonstrate the potential of GIPAC.

Future research directions are summarized in Chapter 10. The GIPAC configuration discussed in Chapter 9 is proposed for further investigation, permitting multiple power supply voltages and multiple clock signals with different operating frequencies to be distributed while only utilizing a single global network. Multiple voltages can be generated using low-dropout DC-to-DC converters. Band-pass filters can be used to extract multiple clock signals at different operating frequencies. An improved power delivery scheme is proposed for investigation that delivers a high voltage, low current signal. Conversion to a low voltage and high current signal is performed on-chip by utilizing an on-chip transformer. Three dimensional integration is expected to benefit from such a power delivery scheme since the higher quality transformer can use specialized technology and integrated with other technologies within a common package. An adaptive link breaking methodology for power network optimization is proposed, expanding the link breaking methodology described in Chapter 8. Based on different noise characteristics, multiple power distribution networks can be connected to reduce the impedance of the network. A changing power network impedance can be used

to reduce side-channel attacks during a power analysis. Voltage drops, electromagnetic interference, and heat distribution are significantly affected by the impedance of the power distribution network. Continuously changing the network impedance is expected to serve as a cryptographic mechanism, increasing the security of integrated circuits.

Finally, the conclusions of this dissertation are described in Chapter 11. In this dissertation, supporting structures, such as global signaling, synchronization, and power delivery, are discussed. A primary focus is the development of accurate models of these structures to estimate the performance and noise characteristics of integrated circuits. The primary tradeoff of these models is between computational complexity and accuracy.

Chapter 2

Resources in Integrated Circuits

A large variety of different resources are available in integrated circuits (IC). The most common resources are metal area, silicon area, and a power budget. Since the primary goal of an integrated circuit is achieving a target performance, a tradeoff between the resources and performance is the primary focus. Resources are typically limited. It is therefore important to effectively allocate these resources among different networks, blocks, and circuits. To properly distribute resources within an IC, numerous design techniques have been developed. Tradeoffs among different resources are discussed in this chapter for each of these design techniques.

With increasing scaling and higher transistor speed, the surrounding components (such as global interconnects) should also increase in performance. More efficient power supply networks, clock distribution networks, and signal interconnects are therefore required. The metal system within an IC is illustrated in Fig. 2.1. The transistors are located at the bottom of this structure. The top metal layers include

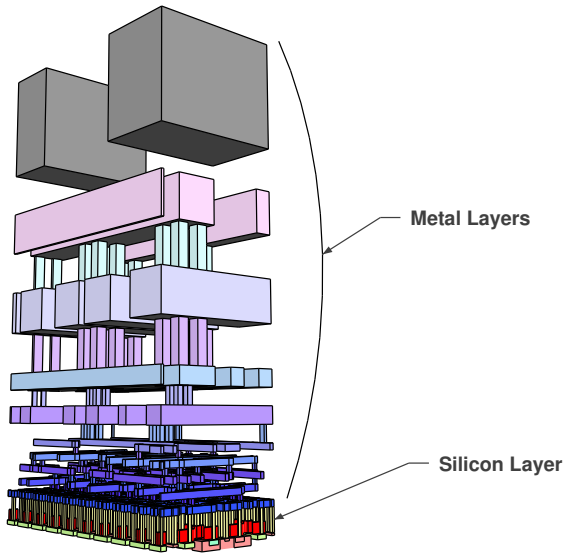


Figure 2.1: IC metal system. The bottom layer is the silicon layer where the transistors are located. The metal layers above the silicon area provide supporting interconnect structures, such as the signal interconnects, power supply network, and clock distribution network.

the power supply, clock network, and signal interconnects to connect the transistors. Since metal area and input/output (I/O) pads are limited, the metal area should be efficiently allocated among these networks.

Power is a limited resource within an IC due to the growing desire for portable applications. In addition, the power dissipated by the supportive network degrades the performance of an IC due to heat and noise. Physical structures and characteristics of the network affect the power dissipation. For clock distribution networks, the primary power dissipation component is the capacitance (of the network and load), which is charged and discharged at each clock cycle. For the supply network, the

voltage is typically close to DC; therefore, the capacitive component of the network is not the primary source of power dissipation. The resistive component, however, is significant since large currents pass within the network, dissipating power as heat. The power dissipated by interconnects is typically a combination of the clock and power supply network characteristics.

Metal resources are also limited in ICs. With each new technology, the size of the IC scales significantly, permitting greater functionality on a single IC. The number of metal layers, however, increases slowly due to mechanical problems, such as material stability, greater physical stress, and polishing difficulties [10]. With additional metal layers, the entire structure is less reliable [11] due to the effect of wafer bow [12], limiting the maximum number of available on-chip metal layers. More metal layers increase mechanical pressure on the bottom lattice of an IC. Furthermore, each additional metal layer requires additional processing steps (additional mask layers), increasing production cost. The optimal sharing and distribution of the metal resource among the networks is therefore an important issue.

2.1 Design Techniques

IC development can be functionally separated into two major layers, namely, the design layer and supportive layer, as listed in Table. 2.1. The design layer includes the architecture, circuit, and interconnect. The power supply system, clock distribution

Table 2.1: Two primary IC development layers, where each layer is further separated into three categories. Each design technique typically targets a single category.

Design Layer	Architecture	Circuit	Interconnect
Supportive Layer	Power Supply System	Clock Network	Substrate

network, and substrate are related to the supportive layer. A large number of design techniques are described in the literature to increase performance or optimize a particular resource. These design technique can be grouped into one of the categories described above. The techniques included within the *architecture* category are arithmetic based voltage scaling [13], arithmetic transformation [14, 15], retiming [16], and wave pipelining [17]. The techniques associated with the *circuit* category are multi-threshold design [18], adaptive body biasing [19], transistor sizing [20], transistor reordering [21], transistor stacking [22], and transistor tapering [23, 24]. Low swing interconnects [25, 26], cascaded buffers [27, 28], repeater insertion [29], shielding [30], differential signaling [31, 32], active regeneration [33, 34], interconnect tapering [35], intentional switching time delay [36], and bus swizling [37, 38] techniques are related to the *interconnect* category.

In the *supportive layer*, the individual techniques are as follows. Multiple supply voltages [39], decoupling capacitors [40], and different power grid structures [41,

42, 43, 44] are techniques for *power supply system* optimization. The characteristics of the *clock distribution network* may be enhanced using H-tree [45], resonant H-tree [46], mesh structures [47, 48], coupled standing wave oscillators [49], rotary traveling wave oscillators [50], resonant distributed differential oscillators [51], and clock gating technique [52]. Improvements in an integrated circuits at the *substrate* level may be achieved with guard rings [53], substrate contacts [54], and triple well [55] techniques.

The focus of this research is on two design categories - global interconnect for signaling and power delivery systems. Each of these categories can greatly enhance performance. Each category is described in greater details in the following subsections to provide enhanced understanding of the inherent design techniques and resource tradeoffs.

2.1.1 Interconnect from a Resource Perspective

The focus of this subsection is on the global interconnects since the power, delay, and other resources are dominated by the interconnect as compared to the logic gates. As mentioned earlier in this chapter, a large number of design techniques are described in the literature to provide enhanced performance of the signal interconnect. The discussion below describes three more common techniques used in the signal interconnect design process.

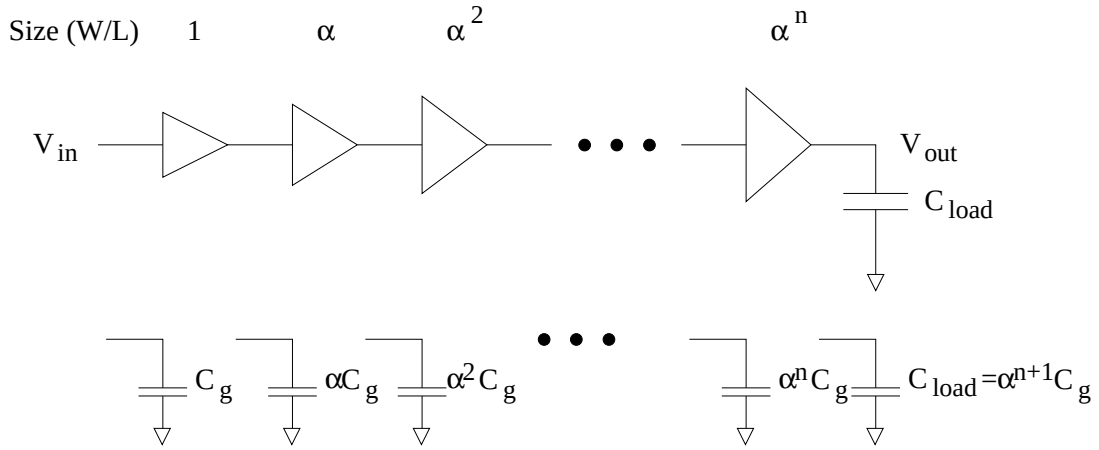


Figure 2.2: System of cascaded buffers. This technique is used to drive large capacitive loads.

a. Cascaded Buffers

In those cases where the load is primarily capacitive, a system of cascaded buffers may reduce the delay of a signal driving that load. A schematic model of a system of cascaded buffers is shown in Fig. 2.2. Geometrically increasing the size of each stage reduces the load on each stage. The area requirement is high due to the large area required for the buffers. To minimize the propagation delay, the following constraints must be satisfied,

$$\alpha\left(ln\alpha - 1\right) = \frac{C_d}{C_g}, \tag{2.1}$$

$$N = \frac{ln\left(\frac{C_d}{C_g}\right)}{ln\left(\alpha\right)}, \tag{2.2}$$

where N and α are, respectively, the number and tapering factor of the cascaded buffers. C_d and C_g are, respectively, the drain/source capacitance and gate capacitance of a minimum size buffer. The load capacitance is $\alpha^{N+1}C_g$.

Neglecting C_d, the optimal tapering factor α is $e \approx 2.7$ [56]. A tapering factor of 11.5 has been used to minimize the power dissipated by the cascaded buffers [28].

b. Repeater Insertion

Repeater insertion is a well known design technique to reduce the delay required to propagate a signal along a resistive line [29]. The technique is illustrated in Fig. 2.3. The objective is to divide the interconnect into smaller sections of equal distance, reducing the quadratic delay dependence on length to a linear dependency, thereby reducing the overall delay [57]. If the number of repeaters is too small, the delay of the interconnect will be dominant. If the number of repeaters is too large, the repeater gate delay dominates. The optimal number of repeaters that minimizes the overall delay has been described in numerous papers, such as [9, 29, 57].

An additional advantage of repeater insertion is reducing coupling noise between adjacent interconnects. The reduction in coupled noise is achieved by restoring the voltage level of the signal to V_{dd} or V_{ss}. It is impossible, however, to insert excessive repeaters due to delay, power and area constraints. The optimal number of inserted repeaters is based on the geometric parameters of the interconnect. The optimal

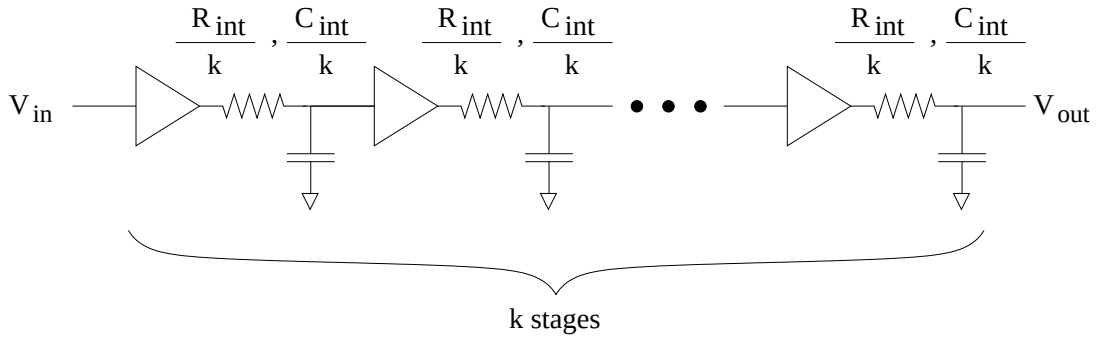


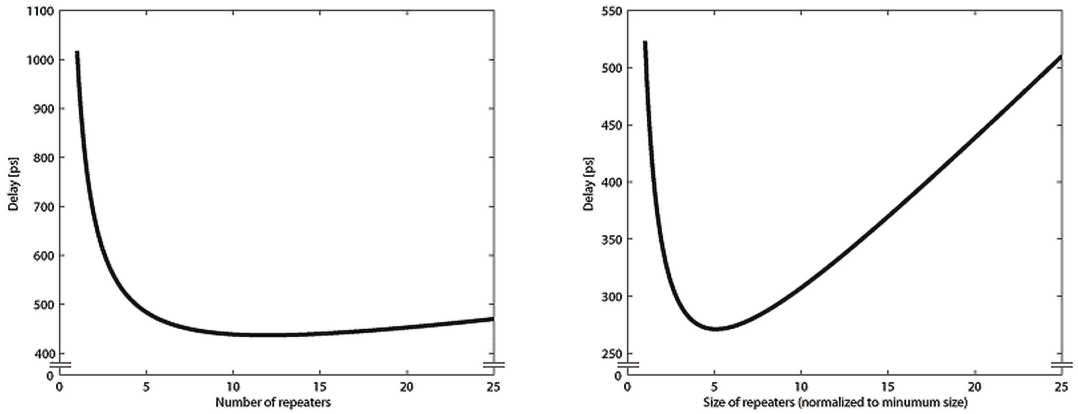
Figure 2.3: Repeater insertion technique. The technique used to drive long interconnects, primarily reducing the delay of the interconnect.

number of repeaters depends upon the resources being optimized. Based on a 45 nm CMOS technology [58], the effect of the number of inserted repeaters and the size of the repeaters on the signal delay is depicted in Figs. 2.4(a) and 2.4(b), respectively, for a 5 mm long line. The optimal number of repeaters to minimize the delay is [29]

$$k = \sqrt{\frac{R_{int}C_{int}}{2R_oC_o}}, \tag{2.3}$$

where R_{int} and C_{int} are the resistance and capacitance of the interconnect, respectively. R_o and C_o are the minimum resistance and capacitance of the repeater, respectively. The optimal size of the repeaters is

$$h = \sqrt{\frac{R_oC_{int}}{R_{int}C_o}}. \tag{2.4}$$



(a) (b)

Figure 2.4: Effect of repeater insertion on signal delay. The signal delay based on the (a) number and (b) size of the repeaters. Both graphs exhibit parabolic behavior, illustrating the optimal number and size of the repeaters to minimize the overall delay.

The repeaters are assumed to be uniformly distributed along the line in (2.3) and (2.4).

Other resources, such as power and area, monotonically increase with the number or size of the repeaters. Hence, to optimize the number of recourses, the power and area of the repeater system should be considered.

c. Shielding

The process of shielding inserts an additional line between a victim line and an aggressor line, as depicted in Fig. 2.5. This technique can be divided into two major categories: passive [59] and active shielding [60]. A passive shield line is connected to the power/ground network, filtering the noise induced by the aggressor on the victim

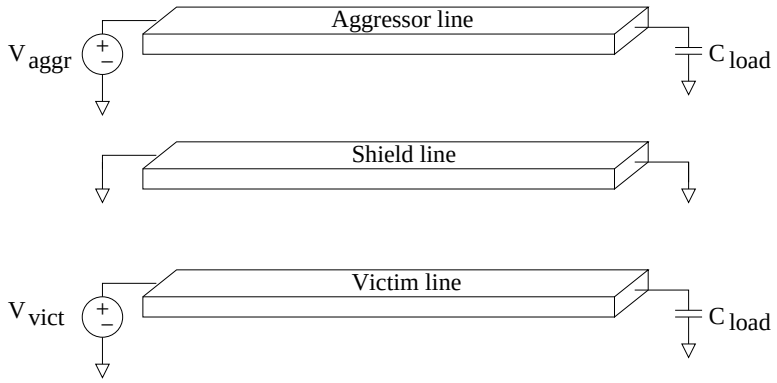


Figure 2.5: Passive shielding technique. Shielding the interconnect reduces the noise coupled from the neighboring circuits and interconnects. Additional metal area is required.

line. The technique is highly effective, although significant area is required. Active shielding switches the state on the shield line based on the aggressor activity, further lowering the coupling capacitance. Additional area for the active shielding circuit and shield line is required; however, the noise in the global bus can be further reduced by up to 25% [61].

Inserting a shield line between the aggressor and victim interconnect is often preferred over physically separating the aggressor and victim interconnects [62]. The shield lines are typically treated as a noiseless, ideal filter. However, the shield line is directly connected to the power/ground network, which is noisy; therefore, the assumption of a noiseless shield lines is inaccurate. Design guidelines for shielding in

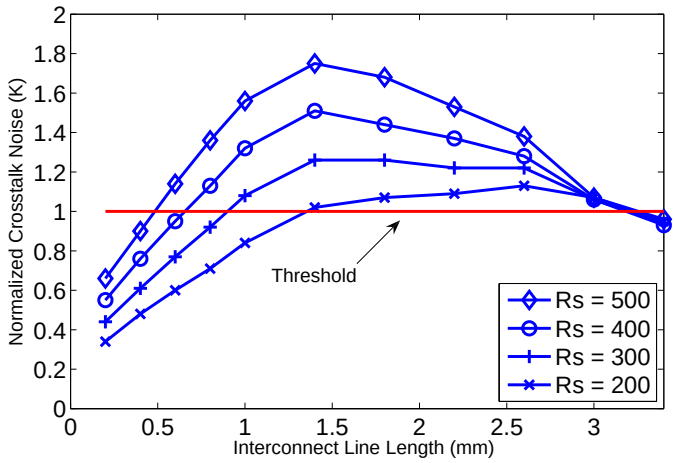


Figure 2.6: Crosstalk utilizing shielding or spacing is illustrated for different values of source resistance based on a 65 nm CMOS technology [63].

the presence of power/ground noise are presented in [63]. The parameter K is

$$K = \frac{V_{shielding}}{V_{spacing}}, \tag{2.5}$$

where $V_{shielding}$ and $V_{spacing}$ are the noise on the victim utilizing only shielding and only spacing, respectively. The overall area is maintained equal with both approaches. Based on a 65 nm CMOS technology [64], assuming a 20% noise peak on the power or ground lines, the preference for shielding or spacing is illustrated in Fig. 2.6 for different values of source resistance R_s. Shielding is preferred in the region where K is less than one. Only physical separation is advantageous when K is greater than one.

2.1.2 Power Supply System

Since every circuit, gate, or transistor requires energy to operate, power delivery is a primary issue in integrated circuits. Large currents pass through the power network, therefore, a small change in the impedance of the network can result in a large voltage drop, significantly decreasing the performance of the circuits. Techniques for designing the power network focus on the impedance and noise characteristics of the network. Similar to the signal interconnect, a large number of design techniques are available in the literature; however, only two common techniques, decoupling capacitor allocation and power grid design, are described below.

a. Decoupling Capacitor Allocation

The switching noise on the power and ground network is caused by transistor activity. This effect can be reduced by placing local capacitors along the network, called decoupling capacitors, as shown in Fig. 2.7. These capacitors act as local current sources, providing charge to the power/ground network. Placed in close proximity to the source of switching current, energy is delivered locally by the decoupling capacitor rather than from the power line, enhancing the characteristics of the power/ground network. The capacitors are typically charged during low switching activity. Since significant area is required for the decoupling capacitors, the location and size of the decoupling capacitors are critical to this technique [40, 65].

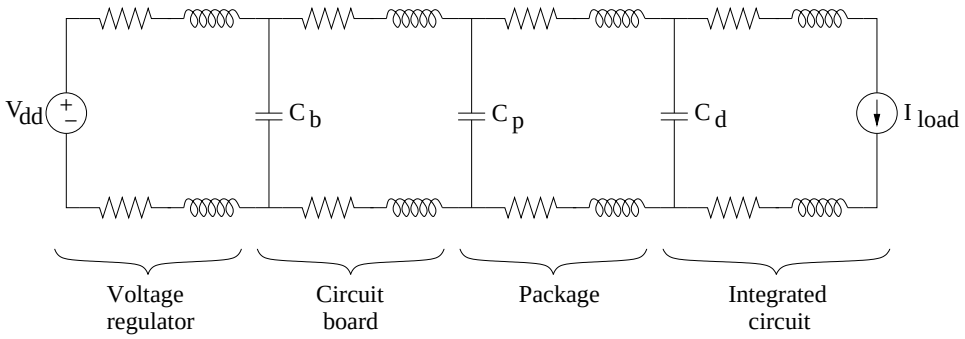


Figure 2.7: Power supply network with decoupling capacitors. The capacitors are allocated between power and ground at each level of the network. The size and location of the on-chip decoupling capacitors (shown as C_d) are critical to providing an efficient on-chip power grid.

Decoupling capacitors are typically placed within the empty area on an IC since significant area is required to implement these capacitors. In this case, the load (sink source) may be distant from the decoupling capacitor, degrading the performance of the power grid. A placement algorithm for decoupling capacitors has been developed, treating the capacitor as an ideal component [66]. Additional algorithms have been developed to optimally size and place decoupling capacitors based on a variety of different assumptions and constraints [40, 67, 68, 69].

b. Power Grid Design

The purpose of the power/ground distribution network is to efficiently distribute power and ground signals over the entire physical area of an IC. Two primary constraints are the impedance and current density of the network. The impedance of

the power/ground network must be sufficiently low to maintain a constant voltage across the network. Due to high frequency switching noise, the impedance of the power/ground distribution network includes both an inductive and capacitive component. The current density is also a significant constraint since the magnitude of the current increases with advancements in technology; therefore, electromigration [70] has become a fundamental issue in power/ground distribution networks. The physical layout can drastically change the impedance and current density of a power/ground network. Four global power distribution network structures are presented in this chapter.

Mesh Structure

In high complexity ICs, mesh structured power/ground distribution networks are typically used. The basic structure is illustrated in Fig. 2.8. The advantages of this structure are high reliability and easy of design. The physical layout of the network permits a local power and ground signal to be connected at various locations across the IC. The primary drawback is high utilization of metal resources, requiring up to 35% of the metal resources to be dedicated to the power/ground distribution network [71, 72]. The mesh structure also achieves lower impedance as compared to nonuniform power distribution structures.

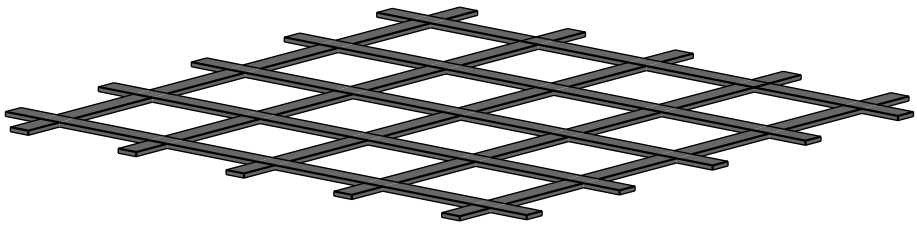


Figure 2.8: Mesh structured power distribution network.

Planar Structure

To further reduce the impedance of the power/ground distribution network, a planar distribution network has been proposed [73]. An entire metal layer is dedicated to the power or ground network, where the signals are routed thought small holes in the metal layer. High metallization usage is therefore required for a planar structured power/ground distribution network.

Cascaded Ring Structure

Another power/ground distribution scheme, a cascaded power/ground ring, has been proposed in [43]. The distribution network is illustrated in Fig. 2.9. The structure is implemented with rings, where rings with smaller radii are allocated at the lower metal layers, as compared to large radii at the higher metal layers. This structure is motivated by providing higher current density at the periphery as compared

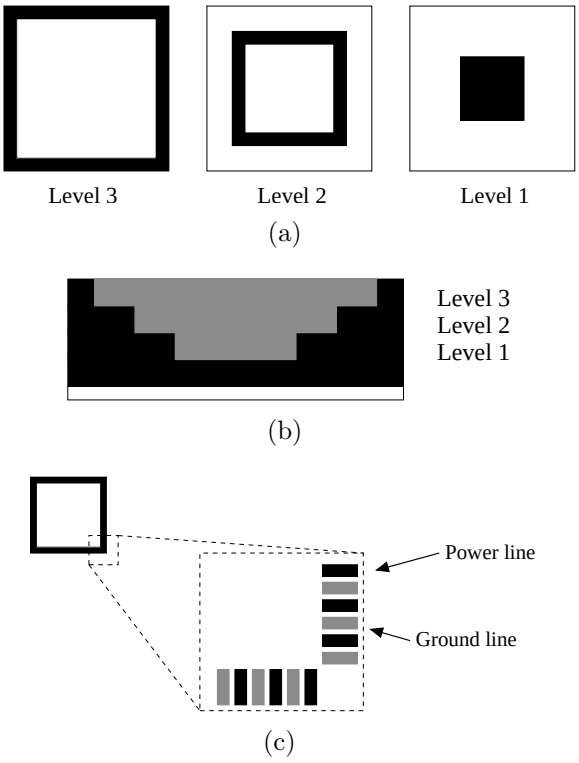


Figure 2.9: Cascaded power/ground ring structure used to provide a higher current density closer to the periphery of the IC. (a) Top view, (b) cross-section view, and (c) zoomed view of the cascaded power/ground ring structure.

to the center of the IC.

Interdigitated Structure With higher operating frequencies, the switching noise on the power/ground networks increases the importance of the inductive component of the network impedance. While the focus of these power/ground networks are primarily on the resistive component of the network impedance, an interdigitated structure [44], depicted in Fig. 2.10, is used to lower the inductive impedance of

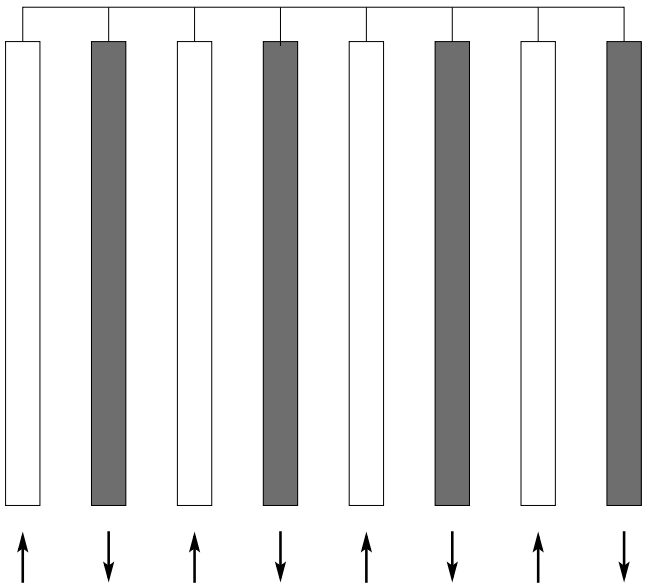


Figure 2.10: Interdigitated power and ground distribution network. The darker and lighter lines represent the power and ground lines, respectively. The structure is used to reduce the inductive component of the network impedance.

the network. By replacing the wide lines with narrow lines, maintaining the same area increases the mutual inductance and decreases the overall inductance of the network [72].

2.2 Optimization Process

Optimization is an extremely large and important field in mathematics. Linear programming was first motivated by problems related to training scheduling and operation planning in the U.S. military, where George Dantzig worked on these scheduling problems. George Dantzig is referred to as the father of linear programming [74].

Since optimization theory is a broad and deep topic, the field of optimization can be separated into many subfields. Integer programming, quadratic programming, linear programming, and nonlinear programming are well known optimization subfields [75, 76, 77]. An important optimization subfield is convex programming [78], where the function and constraints being optimized are convex functions. When classical optimization techniques are too expensive, a random search method is typically used, known as stochastic programming [79] and includes optimization methods such as Monte Carlo techniques [80].

A specific resource with a single variable may be optimally allocated by determining the minimum/maximum point, known as single variable optimization. The complexity however increases significantly with the number of resources that need to be optimized. A classical approach to the multiple resource optimization process typically utilizes an objective function, also known as a *cost function* [81, 82]. Utilizing two or more design techniques, a large number of resources can be optimized to increase the overall performance of an IC; however, greatly increasing the complexity of the cost function.

2.2.1 Multi-Objective Optimization

The general form of a multi-objective optimization problem is

$$minimize\ F = \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix},\ where\ n \geq 2\ and\ x \in \mathbb{R}. \tag{2.6}$$

Each function $f_i(x) : \mathbb{R} \rightarrow \mathbb{R}$, where $i = 1, ..., n$. In the IC design process, this function represents a specific resource such as power, area, or noise. The multi-objective optimization problem is typically solved by converting (2.6) into a single objective cost function [81, 82]. A cost function is typically a weighed sum of the resources or product of resources with power coefficients, such as

$$cost = \alpha_1 \cdot f_1(x) + \alpha_2 \cdot f_2(x) + \alpha_3 \cdot f_3(x) + \ldots + \alpha_n \cdot f_n(x), \tag{2.7}$$

or

$$cost = f_1^{\beta_1}(x) \cdot f_2^{\beta_2}(x) \cdot f_3^{\beta_3}(x) \cdot \ldots \cdot f_n^{\beta_n}(x), \tag{2.8}$$

where α and β quantify the importance of a particular resource. In [83], the function with $n = 2$, $f_1 = power$, $f_2 = delay$, and $\beta_1 = \beta_2 = 1$, referred to as a power-delay product, is used to optimize a system of tapered buffers. While normalization is

required for the resources in (2.7), (2.8) is more complicated. The key issue in this standard optimization process is the requirement of choosing the value of $\alpha_{1..n}$ and $\beta_{1..n}$ prior to the optimization process. For the same system, two different sets of coefficients may be chosen, producing different results. Additionally, the importance of certain resources may change. These aspects can constrain the standard optimization process. A general flow diagram of the cost function optimization process is shown in Fig. 2.11. In this optimization process, the initial objective is to determine the coefficients of the resources being optimized. Upon completion of this manual step, the automated optimization process is performed. A *cost function* based on a specific resource model is used to determine the optimal solution.

For an objective function, the optimal solution is the point where the cost function produces the minimum value. An analytic solution is typically not possible due to the complexity of the cost function. A variety of numerical methods can however be used to determine the optimum solution [76, 77, 84, 85].

2.3 Summary

A variety of design techniques are listed in this chapter for different IC abstraction layers. Design techniques for signal interconnects and power distribution networks, such as repeaters, shielding, and decoupling capacitors, are reviewed. Tradeoffs for each design technique are introduced. In the next chapter, common power grid design

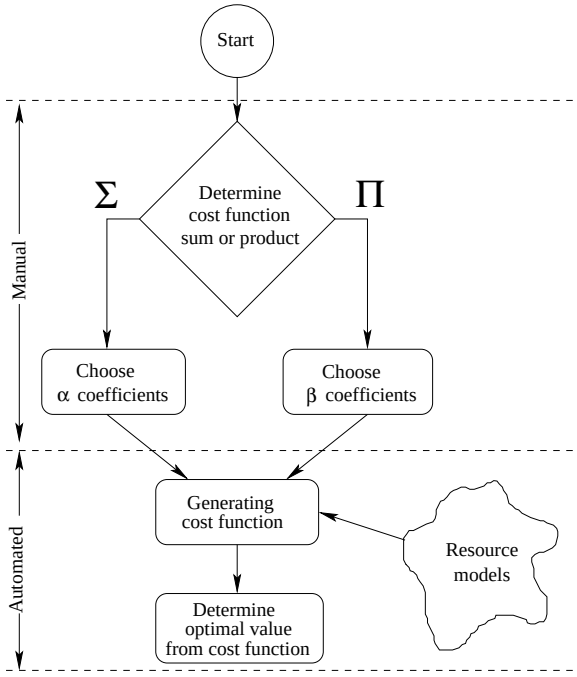


Figure 2.11: Cost function optimization process.

techniques are discussed.

The concept of a cost function is presented, permitting multiple resources to be simultaneously optimized. The primary disadvantage of the cost function approach is the requirement to choose weight coefficients prior to the optimization process.

Chapter 3

Modeling IC Structures

Modeling is an essential step in the analysis process. A variety of models are proposed in the literature for different system and circuit components. Complex models are used to accurately predict the behavior of a system. The computational complexity can however be high. Since optimization processes utilize models to estimate the behavior of a system, only those models with low computational complexity may be effective in a complex design flow. The tradeoff between accuracy and computational complexity is the primary tradeoff; each model should therefore be characterized according to these two aspects.

The complexity of optimizing highly complex integrated circuits is a well known issue since billions of transistors often need to be simultaneously simulated. To overcome this issue of computational complexity in the IC design process, simulation tools have been developed [86, 87] where the logic gates are modeled as simple resistors. These modeling techniques are used, for example, to optimize the transistor size.

In [88], the concept of micromodeling VLSI circuits is described. The digital gates are modeled at a lower level, where the resistive and capacitive components of the transistors are considered. Based on these models, techniques and algorithms have been proposed to optimize both delay and power.

The discussion in this chapter is focused on two IC structures: power networks in Section 3.1 and the monolithic substrate in Section 3.2. Due to the high complexity of these structures, focus is on computationally efficient closed-form expressions that express the behavior of these structures.

3.1 Modeling Power Networks

With increasing length and decreasing width of the interconnect, the wires can no longer be neglected when modeling a system for power and delay. Figures of merit characterizing the length of an interconnect have been developed [89] which consider the inductive interconnect characteristics. Since global interconnects dissipate large power and require long signal delay (and higher delay uncertainty), research has focused on the power, clock, and global signal networks. In this section, models of power/ground distribution networks are discussed, focusing on mesh and interdigitated structures.

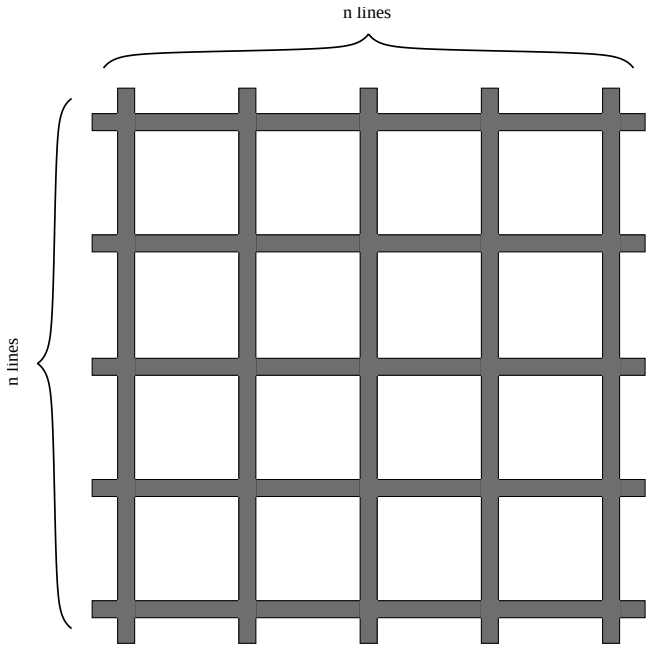


Figure 3.1: Mesh structured power distribution network. n^2 nodes are determined for $2n$ power lines.

3.1.1 Model of Mesh Power and Ground Distribution Networks

A mesh structure, as shown in Fig. 3.1, typically utilizes a single metal layer for distributing power and an additional metal layer for distributing ground; the two networks can therefore be modeled independently. The number of nodes in a mesh structured power supply network can be extremely high. For $2n$ power lines, n^2 nodes are required. Since the power supply network typically covers an entire IC, the number of power lines can be large; therefore, the number of nodes in the network

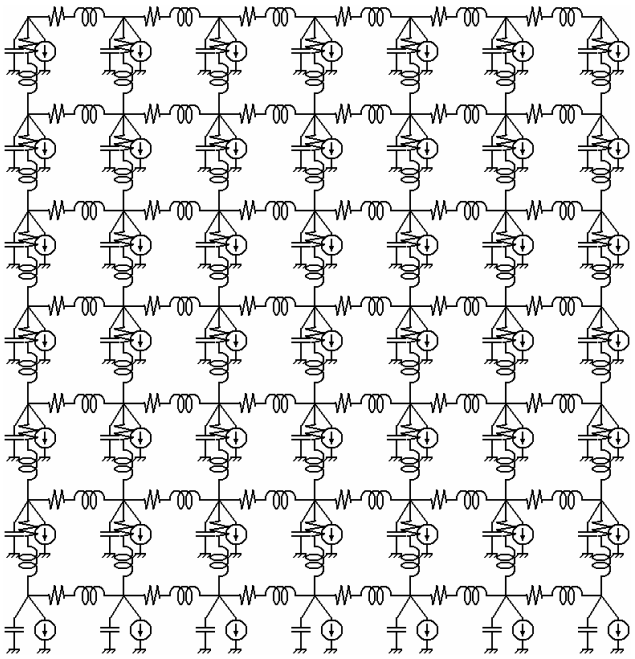


Figure 3.2: RLC grid model of mesh structured power network.

is often excessively high. Directly solving Kirchhoff's voltage or current equations is therefore a complicated and computationally inefficient method.

In [90], the approximate frequency response of an RLC mesh structured power supply network is described. This structure is depicted in Fig. 3.2. With this model, equal current is assumed to be sunk at every node [90]. This assumption is justified in large scale global power networks with a uniform physical layout. The power supply network is reduced to a single RLC interconnect line, converting the problem from a 2-D structure into a 1-D structure. The computational complexity is therefore

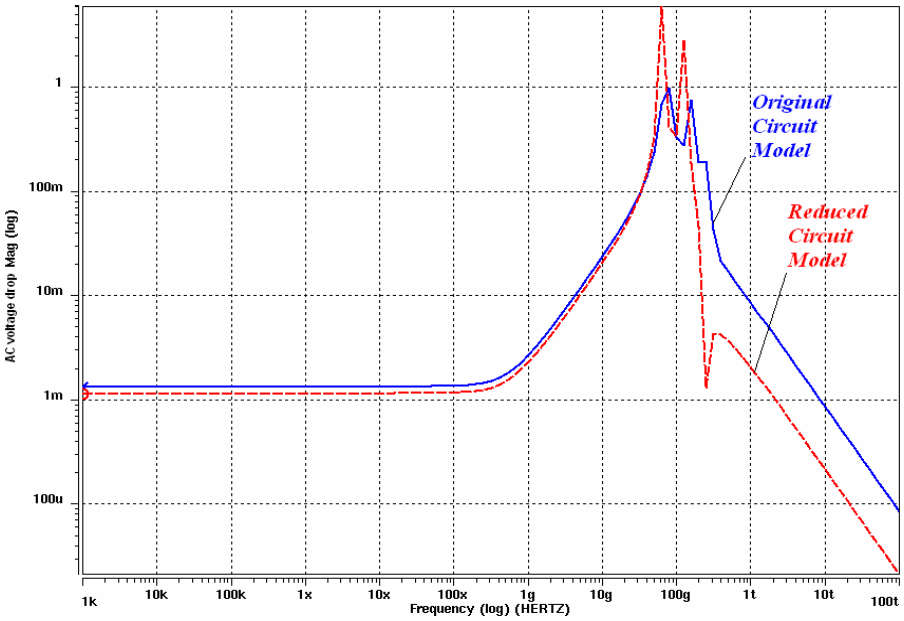


Figure 3.3: Frequency responde of original large mesh and reduced model of an RLC power grid [90].

decreased from $O(n^2)$ to $O(n)$. The frequency response of the original mesh structure and reduced model is illustrated in Fig. 3.3, indicating that a reduced model accurately predicts the frequency response of a complex mesh network. A reduced model can therefore achieve a comparable frequency response with low computational complexity.

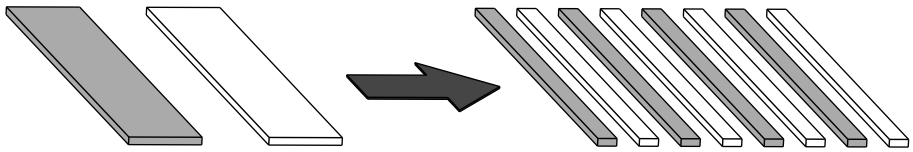


Figure 3.4: Replacing a few wide lines with many narrow lines reduces the inductance of a system of interdigitated interconnect.

3.1.2 Interdigitated Power and Ground Distribution Network Model

A primary advantage of an interdigitated structure is exploiting this structure to reduce the inductance of a system. The inductance is reduced by replacing a single wide wire by several narrow lines, as shown in Fig. 3.4. Coupling among the wires complicates the modeling procedure in an interdigitated structure, since a greater number of coupling components need to be calculated. In [91], the inductance of a pair of wires within an interdigitated structure is determined by treating the inductance as a local phenomena. This assumption is justified if current flowing within the neighboring power and ground wires has the same magnitude and opposite direction. The mutual inductance between two distant pairs can therefore be neglected. The

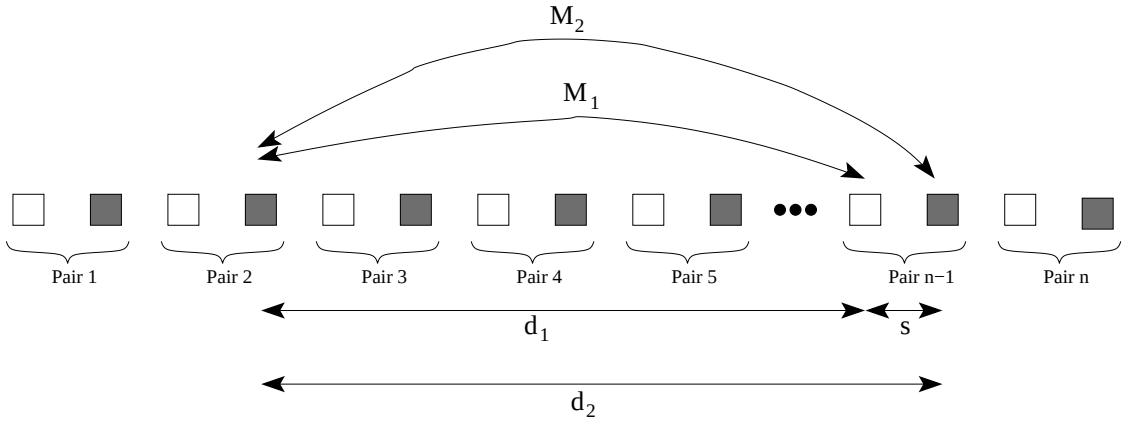


Figure 3.5: Mutual inductance between distant power/ground pairs. Since the direction of the current in the $(n-1)^{st}$ pair is typically opposite in power and ground lines, the sum of the two mutual inductance terms is $M_2 - M_1$.

mutual inductance terms between two distant pairs are shown in Fig. 3.5. Since distance $d_1 \gg s$ (refer to Fig. 3.5), distances d_1 and d_2 can be assumed equal, making the magnitude of the mutual inductances M_1 and M_2 approximately equal. Since the direction of the current is often opposite in power and ground wires, the sum of the two mutual terms is equal to zero ($M = M_2 - M_1$). M_1 and M_2 can only be assumed equal in distant pairs. By maintaining the assumption $d_1 \gg s$, the error is greater when determining the mutual inductance between close power/ground pairs.

To lower the computational complexity, the inductance can be determined by considering a single power/ground pair. The accuracy is enhanced by including additional power/ground pairs. The computational complexity, however, increases significantly. The required number of mutual terms for different number of power/ground pairs is

Table 3.1: Number of required mutual terms to estimate the inductance for different number of power/ground pairs. A large number of power/ground pairs increases the model accuracy with an increase in computational complexity.

Model	Number of power/ground pairs	Number of mutual terms
Mezhiba [91]	1 2 3 4	1 6 15 28
Grover [92]	n	$2n^2 - n$

listed in Table 3.1, suggesting that the computational complexity rapidly increases with additional power/ground pairs. In Table 3.1, the Grover model [92] refers to those calculations where all of the individual power/ground pairs are included. This model is used in 3-D multipole extraction tools, such as FastHenry [93].

3.2 Modeling the Monolithic Substrate

Substrate noise coupling continues to be a primary concern in mixed-signal circuits, such as a transceiver where the digital and analog/RF functions are placed on the same monolithic substrate. The demand for higher integration magnifies this issue due to the shorter distance between the analog and digital blocks, and the greater noise within the substrate.

A large variety of substrate noise isolation techniques have been proposed in the literature [94, 95, 96, 97]. These techniques require the monolithic substrate to be properly modeled. The substrate is analyzed based on one of two methodologies: (a) substrate extraction based on a 3-D RC mesh and (b) macromodels that model the effective impedance between multiple ports.

The substrate extraction methodology is typically based on discretizing the substrate into a 3-D mesh and solving the differential equations utilizing the finite difference method (FDM) or boundary element method (BEM). With FDM, the substrate is composed of multiple three-dimensional cubical structures. The accuracy of the method can be increased by decreasing the size of the basic cube; the computational complexity is however increased. The size of the cubes can also be varied throughout the substrate, reducing the size of the basic cube near the contacts or devices, while not significantly increasing the computational complexity of the process. Another approach for solving the set of differential equations characterizing the current flow within the substrate is BEM. Different from FDM, in BEM the substrate is discretized only near the boundaries, such as the contacts and active areas, reducing the extraction process from a 3-D problem to a 2-D problem. This change greatly reduces the computational complexity. Non-uniformities within the substrate are however neglected using BEM.

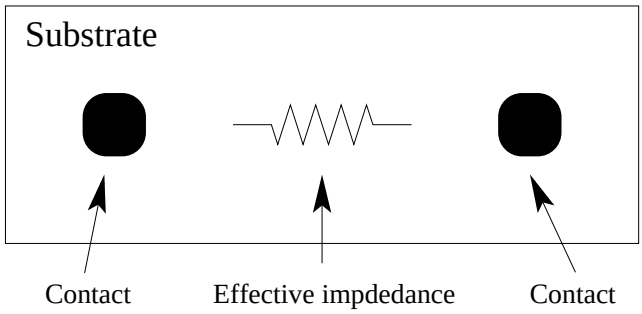


Figure 3.6: Top view of the effective impedance of the substrate between two contacts.

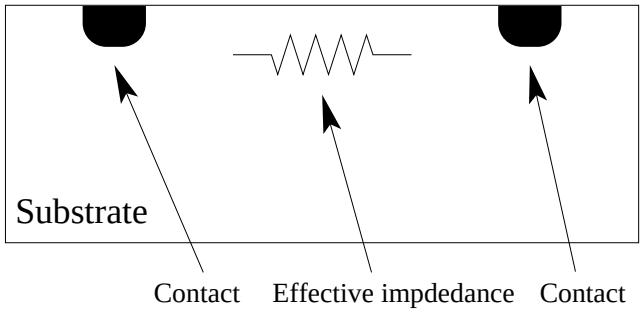


Figure 3.7: Cross-section of the effective impedance of the substrate between two contacts.

3.2.1 Compact Models

Substrate analysis can also be based on compact models, describing the effective impedance between two ports, as illustrated in Figs. 3.6 and 3.7. These compact models significantly reduce the computational complexity, which is essential when analyzing large scale systems.

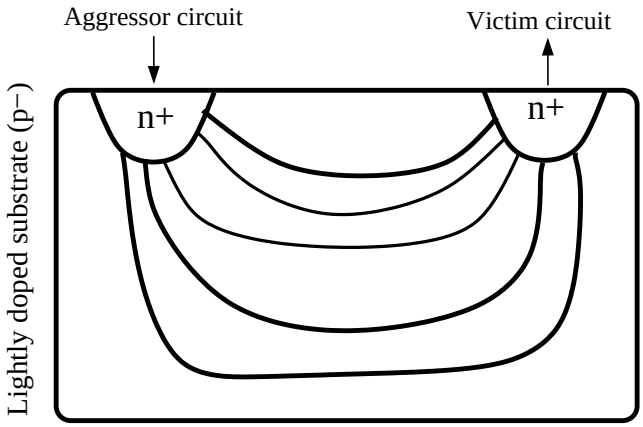


Figure 3.8: Current distribution between the input and output ports within a lightly doped substrate.

A compact model of the substrate is different for a lightly doped substrate and an epi-type substrate with a grounded back side metal. The current flow for those two scenarios is depicted in Figs. 3.8 and 3.9. In an epi-type doped substrate with a grounded back side metal, the effective impedance from the port to the back side metal should be highly accurate. In a lightly doped substrate, the emphasis is on the effective impedance between the two ports.

In [98], the effective resistance between a port and the back side metal is

$$Z = \frac{1}{K_1 \cdot A + K_2 \cdot P + K_3}, \tag{3.1}$$

where A and P are, respectively, the area and perimeter of a substrate contact (or port). The variables K_1, K_2, and K_3 are fitting parameters for a specific technology,

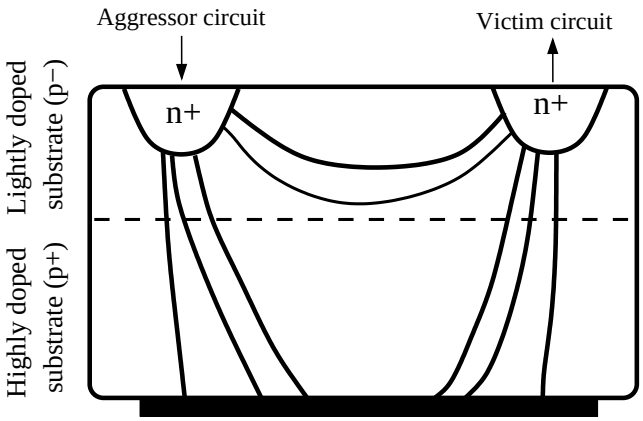


Figure 3.9: Current distribution between the input and output ports within an epi-type substrate with a grounded back side metal.

which are determined from test structures. The impedance between the contacts (or ports) is [99]

$$Z = \frac{1}{\alpha \cdot \exp\left(-\beta x\right)}, \tag{3.2}$$

where α and β are, respectively, fitting parameters. x is the separation between the two ports.

The effective resistance between two ports is proposed in [98] to be modeled as

$$R = K \cdot x^{p}, \tag{3.3}$$

where K and p are, respectively, fitting parameters. For those cases where two ports

are of different size, the effective resistance is

$$R = \frac{K \cdot x^p}{\sqrt{A_a} + \sqrt{A_b}}. \tag{3.4}$$

A_a and A_b are, respectively, the physical area of contacts a and b.

In [100], a compact model for the effective resistance between two contacts, a and b, is

$$R = \beta \cdot [\ln(x+1)]^{\alpha_1} \cdot (A_a + A_b)^{\alpha_2} \cdot (P_a + P_b)^{\alpha_3}, \tag{3.5}$$

where β, α_1, α_2, and α_3 are technology coefficients.

These compact models accurately predict noise propagation within the substrate for a particular technology. The large number of fitting coefficients however makes these models less attractive, since all of these coefficients need to be determined for every technology. Compact models with a fewer number of fitting coefficients would be more widely adopted and applied in noise analysis and optimization.

3.3 Summary

Modeling is applied within every level of the design process. Each model can be characterized in terms of accuracy and computational complexity. More complex models typically achieve higher accuracy, while suffering from excessive computational time. In this chapter, models of power distribution networks and the monolithic

substrate are discussed.

In Section 3.1, models of two primary power distribution networks are discussed. Since the on-chip inductance is increasingly important, models of power supply networks have become highly complex due to the large size and distributed nature of the network. For mesh structured networks, models for reducing a mesh to a single RLC interconnect line have been described, significantly reducing the complexity of the model [90]. The inductance of an interdigitated power/ground network can be efficiently modeled by treating the inductance as a local phenomena.

Substrate extraction is discussed in Section 3.2. Different models that estimate noise propagation in a lightly doped substrate and an epi-type substrate with a grounded back side metal are reviewed. The primary advantage of these compact models is efficiently estimating substrate currents as compared to more computationally expensive numerical techniques such as FDM or BEM. The fitting coefficients within compact models are typically adjusted for every technology. Models with fewer fitting parameters that can accurately predict the current flow within the substrate are therefore quite attractive.

Chapter 4

Compact Substrate Model

Substrate noise coupling continues to be a primary concern in highly heterogeneous mixed-signal circuits such as transceivers where the digital and analog/RF functions are placed on the same monolithic substrate [96]. The demand for higher and diverse integration exacerbates this issue due to the reduced physical distance between the digital aggressor and the sensitive analog/RF blocks.

A variety of noise reduction and isolation techniques exist to alleviate substrate noise coupling. The evaluation of these techniques and quantification of the substrate noise at the boundary of the sensitive circuit require a computationally efficient analysis methodology which simultaneously considers the circuit activity, power/ground network, and substrate network.

Existing substrate network extraction techniques fail when analyzing large scale circuits due to the increasing computational complexity of large scale circuits, making the efficient estimation of the substrate noise prohibitive. Current approaches to

model the substrate can typically be divided into two classes. The first class includes those techniques that discretize the substrate into a 3-D $R(C)$ mesh to determine the impedances such as the finite difference method (FDM) [101], [102] and the boundary element method (BEM) [103], [104]. Although highly accurate, the primary limitation of these approaches is the significant increase in computational complexity with circuit size, prohibiting the efficient analysis of large scale mixed-signal circuits [105].

The second class of substrate modeling methods is the use of macromodels to represent the impedance between two ports within a substrate [99], [100], [106]. The primary advantage of this approach is fast estimation of the substrate impedance with reasonable accuracy, supporting the efficient evaluation of different isolation structures without extracting the entire substrate. The difficulty in using these models is the requirement to fit several process dependent parameters.

Compact models are presented in this chapter that require *only one* fitting parameter as opposed to multiple parameters proposed in existing work [99], [100], [106].. Furthermore, these proposed models are applicable to lightly doped substrates which are more challenging to model [107], but are commonly used in mixed-signal and analog circuits due to enhanced isolation [100], [108]. Note that the majority of existing models is valid only for heavily doped substrates [99], [106], [98], [109], where the bulk can be represented as a single equipotential node [107].

The rest of the chapter is organized as follows. The proposed closed-form expressions for the substrate resistances are described in Section 4.1. These compact models are also used to evaluate the isolation efficiency of several structures such as a guard ring and a triple well. The models are applied in Section 4.2 to determine the optimum number of guard rings for a large aggressor block. An industrial circuit with a lightly doped substrate is used to compare different common isolation structures, as described in Section 4.3. The chapter is summarized in Section 4.4.

4.1 Substrate Models

Several models to efficiently estimate the substrate resistance are described in this section. The proposed half-ellipse model to characterize the substrate resistance between two ports is described in Subsection 4.1.1. The isolation between the bulk substrate and a p-well block is compared in Subsection 4.1.2. Models are also described for several commonly used signal isolation structures. Specifically, a circuit model for a guard ring and a triple-well with a guard ring structure is described, respectively, in Subsections 4.1.3 and 4.1.4.

4.1.1 Substrate Resistance Between Two Ports

A lightly doped (bulk type) substrate is assumed since a model of a bulk type substrate is significantly more complicated than an epi type substrate and the bulk

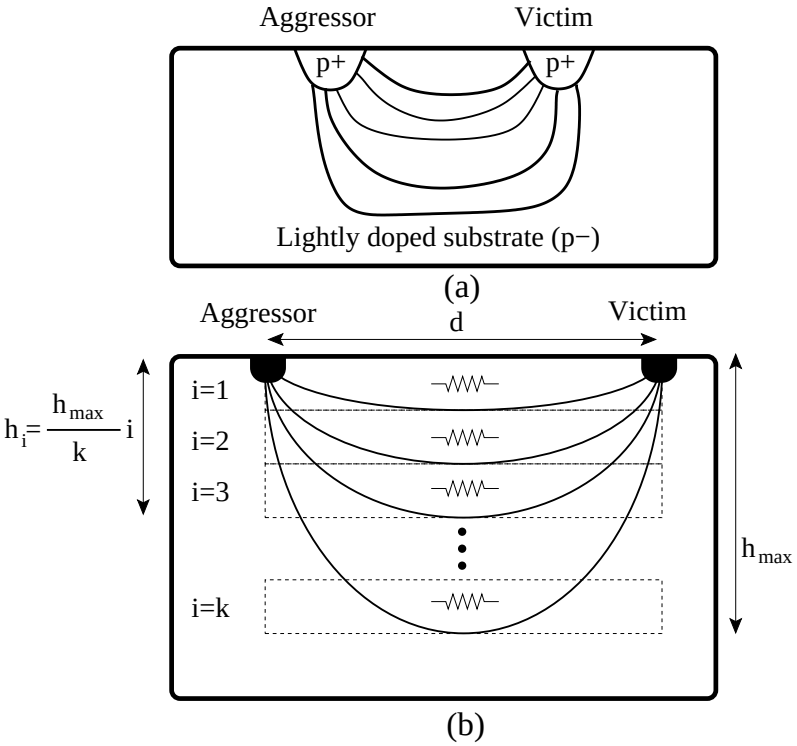


Figure 4.1: Current propagation between two ports within a lightly doped substrate: (a) based on experimental and device simulations, and (b) approximation based on the half-ellipse model.

cannot be represented as a single equipotential node. A bulk type substrate is more appropriate for mixed-signal and analog circuits where substrate coupling is a primary concern. The current propagation path between two ports, *e.g.*, an aggressor and victim, within a lightly doped substrate is depicted in Fig. 4.1(a) [110]. A large portion of the current flows near the surface, but a smaller portion of the current flows deeper within the substrate. The resistance of these deeper paths is therefore higher than those paths near the surface. This current flow is approximated with a *half-ellipse*, as depicted in Fig. 4.1(b). The perimeter of a half-ellipse is used to

estimate the resistance of the path. Since the perimeter of the half ellipse is greater within the deeper portion of the substrate, as shown in Fig. 4.1(b), the substrate resistance is also greater in the proposed model.

Each half-ellipse shown in Fig. 4.1(b) represents a current path, and therefore, is modeled as a substrate resistance. The value of each resistance is based on the perimeter P of the half-ellipse [111],

$$P(d, h_i) = \frac{\pi}{2}\left[3\left(0.5d + h_i\right) - \sqrt{\left(1.5d + h_i\right)\left(0.5d + 3h_i\right)}\right], \tag{4.1}$$

where d is the distance between two ports and $h_i = \frac{h_{max}}{k}\, i$. h_{max} is a technology dependent fitting parameter that determines the depth of the current within the substrate. i is an index from 1 to k, where k determines the number of vertical substrate resistances within the model. A higher k produces a more accurate result at the expense of a linear increase in computation. The results presented in this chapter assume $k = 100$. Dimensions d and h_i are also illustrated in Fig. 4.1(b). Note that these dimensions are proportional to the two radii of the half-ellipse. d is constant for each path, but h_i varies based on the depth of the current path within the substrate. The resistance R_i of each path is approximated using (4.1) as

$$R_i = \rho \frac{P(d, h_i)}{(h_{max}/k)w}, \tag{4.2}$$

where ρ and w are, respectively, the resistivity of the substrate and width of the port. The effective resistance R_{eff} between two ports is therefore the sum of the parallel resistors,

$$R_{eff} = \frac{1}{\left(\sum_{i=1}^{k} \frac{1}{R_i}\right)}. \tag{4.3}$$

The effective resistance can also be expressed in the integral form, as k approaches infinity,

$$R_{eff} = \frac{1}{\int_0^{h_{max}} \frac{w}{P(d,h)}\,\mathrm{d}h}. \tag{4.4}$$

The proposed half-ellipse based model is compared with SubstrateStorm, a commercial BEM based substrate extraction tool, also referred to as SNA [112]. This comparison is illustrated in Fig. 4.2 to evaluate the accuracy of the model. The normalized substrate resistance between two ports is shown in this figure as a function of the distance between these ports. Note that the model accurately captures the nonlinear dependence of the substrate resistance over a wide range of distance, where the RMS error is 14%. The computational complexity of the proposed model is negligible as compared to the commercial SNA extraction tool.

4.1.2 P-Well Block Isolation

A p-well block is an isolation structure typically used in mixed-signal and analog circuits to increase the resistivity of the substrate, thereby reducing the coupling

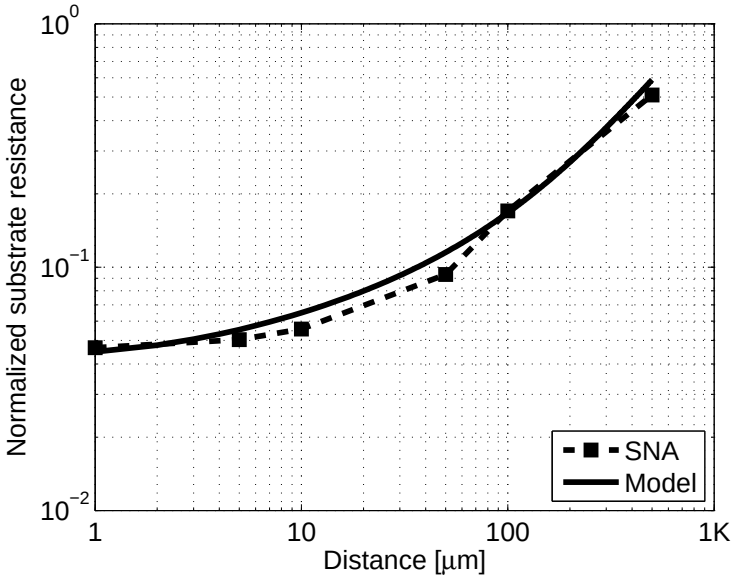


Figure 4.2: Comparison of the model with a commercial substrate extraction tool (SNA). The proposed model accurately captures the nonlinear dependence of the substrate resistance with distance.

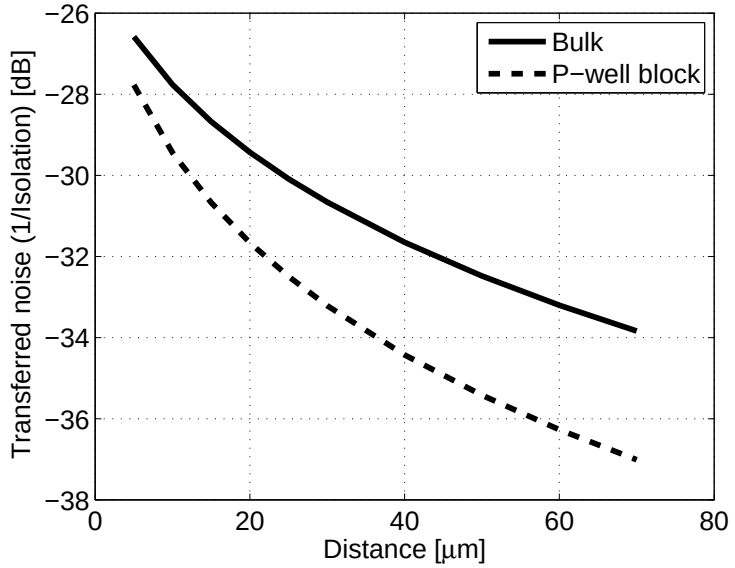


Figure 4.3: Comparison of the isolation efficiency of a bulk substrate and p-well block.

noise between an aggressor and a sensitive node. The model proposed in the previous section is also valid for p-well block isolation. The parameter h_{max}, however, needs to be fitted for this structure since the current propagation characteristics of a p-well block are different than a bulk substrate.

The magnitude of the transferred noise is shown in Fig. 4.3 for two cases: (1) a bulk substrate and (2) a p-well block. If the distance between the aggressor and victim port is small, the isolation efficiency of a p-well block is comparable to a bulk substrate. For example, at 5 μm, the p-well block only provides an additional 1 dB isolation. The advantage of a p-well block becomes more apparent as the distance between the ports increases, assuming that the size of the p-well block is also larger.

4.1.3 Guard Ring Isolation

A guard ring refers to the p+ substrate contacts (or n+ taps for the n-well) placed around an aggressor or victim and connected to a ground pad (or power pad for the n-well), as illustrated in Figs. 4.4(a) and (b). The guard ring eliminates coupling noise by providing a low impedance path for the injected noise current within the substrate (or n-well), thereby improving the noise characteristics of the victim. Note that a guard ring can be placed around only the aggressor, victim, or both the aggressor and victim.

The proposed circuit model of a guard ring is depicted in Fig. 4.4(c). Resistors

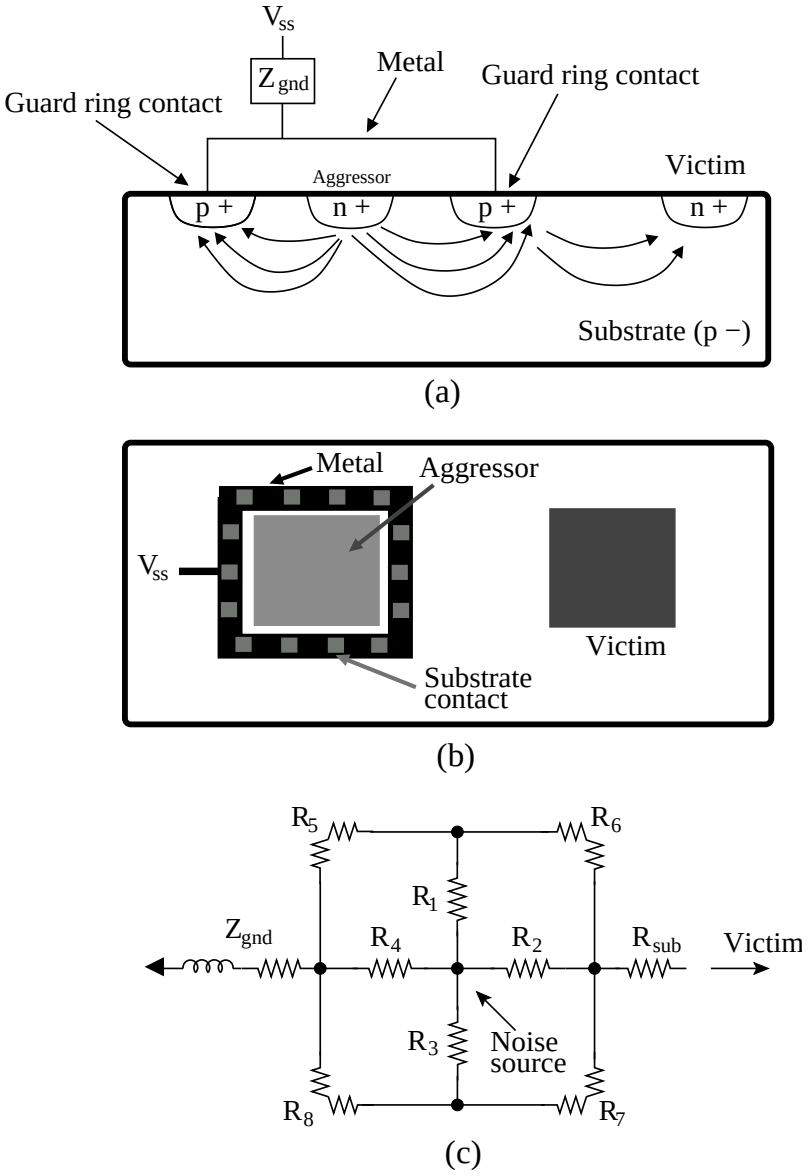


Figure 4.4: Illustration of a guard ring: (a) cross-sectional view, (b) top view, and (c) equivalent circuit.

R_1 to R_4 represent the substrate resistance from the noise source to the ring. These resistances are described by (4.3), as mentioned in Section 4.1.1. R_5 to R_8 represent the resistance of the metal surrounding the ring. These resistances can be determined directly from the sheet resistance since the width and length of the metal are known. The impedance between the ring and ground pad is modeled by Z_{gnd}. Finally, the substrate resistance between the ring and victim is represented by R_{sub} which is also described by (4.3).

Several parameters such as the width and connectivity of the ring, as well as the pad location, significantly affect the overall efficiency of the ring. For example, the isolation achieved by a ring is illustrated in Fig. 4.5 as a function of the distance between the ring and ground pad. As shown in this figure, the proposed model exhibits reasonable accuracy as compared to SNA. Note that as the distance between the ring and ground pad increases, the isolation degrades due to the higher impedance Z_{gnd} of the ground network. The impedance of the ground network connected to the ring should therefore be lower to increase the efficiency of the ring.

4.1.4 Triple Well with Guard Ring

Another technique to further increase the isolation efficiency of a guard ring is to use a deep n-well, also referred to as a triple well or isolated p-well, as illustrated in Fig. 4.6(a). The efficiency is increased by the junction capacitances C_{j1} and C_{j2}, as

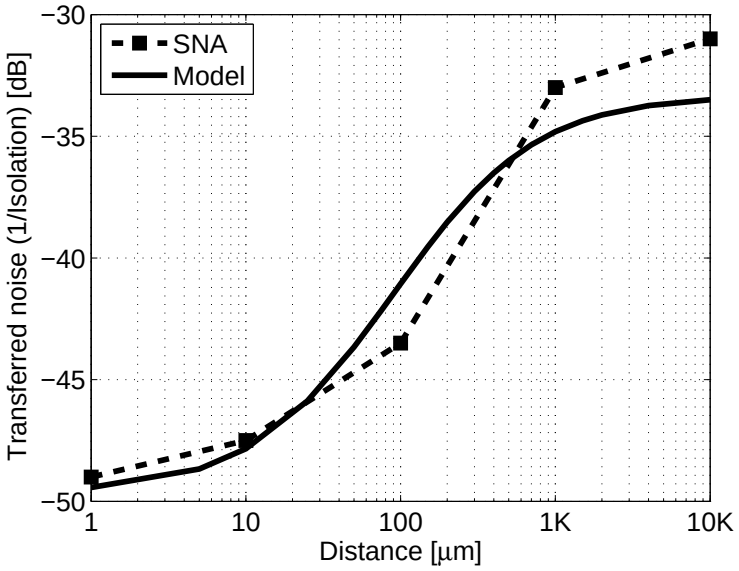


Figure 4.5: Magnitude of transferred noise as a function of the distance between the ring and ground pad.

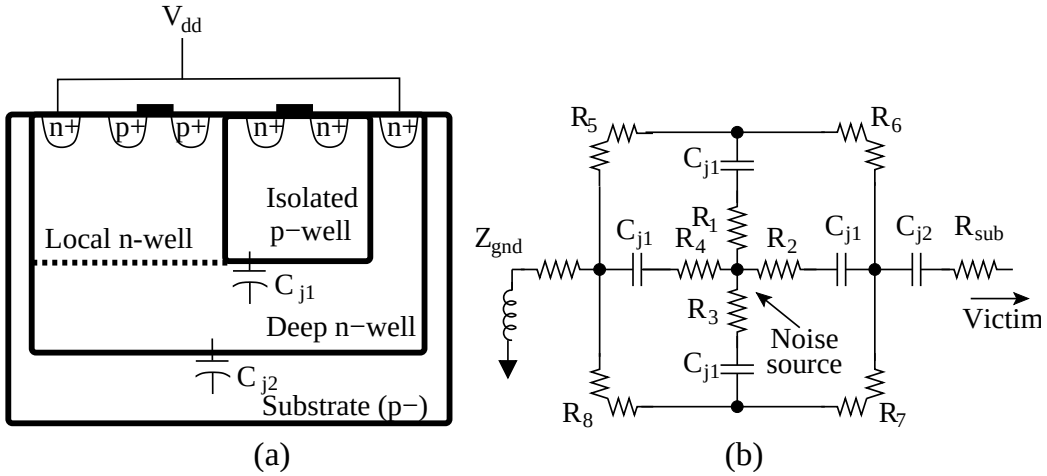


Figure 4.6: Triple well structure to further increase the efficiency of a guard ring: (a) cross-sectional view and (b) equivalent circuit.

shown in this figure. The proposed circuit model of a triple well with a guard ring is depicted in Fig. 4.6(b). The model is similar to the guard ring model with the addition of the junction capacitances. The junction between the isolated p-well and deep n-well is represented with four capacitances (C_{j1}) and the junction between the deep n-well and substrate is represented by the capacitance C_{j2}. These capacitances are dependant on the dimensions of the deep n-well and other technology parameters. Note that the substrate resistances, R_1 to R_4 and R_{sub}, are determined as described in Section 4.1.1.

The isolation obtained with a triple well with a guard ring is compared with SNA in Fig. 4.7 for several circuit sizes and two frequencies. The proposed model accurately captures the effect of circuit size on the noise isolation characteristics, as illustrated in this figure. Note that the isolation efficiency is significantly lower at higher frequencies since the isolating effect of the capacitances diminishes with increasing frequency. Also note that the isolation efficiency of the triple well structure degrades as the size of the aggressor circuit grows although the injected noise is the same. This behavior is due to the larger junction capacitances within a large scale circuit, also demonstrated in [113]. It is therefore desirable to divide a large deep n-well into smaller sections to improve the isolation efficiency.

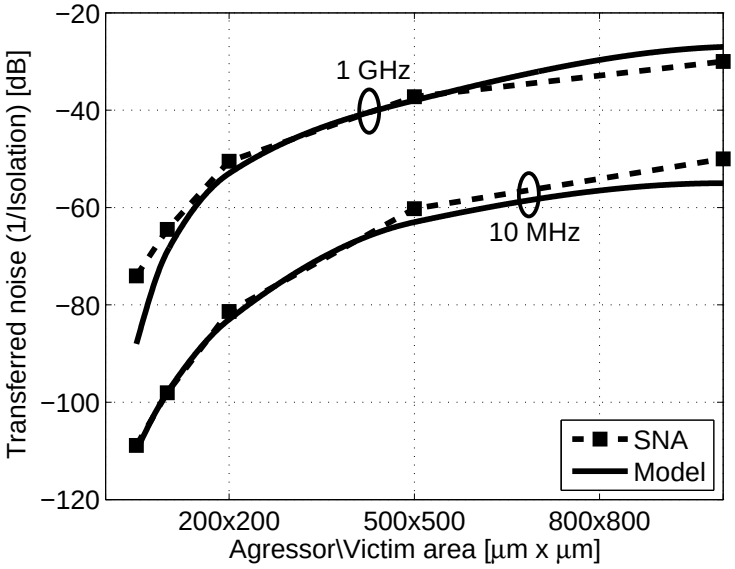


Figure 4.7: Magnitude of transferred noise as a function of circuit size for a triple well with a guard ring structure.

4.2 Optimum Number of Guard Rings

The guard ring is less efficient when using the isolation for a large aggressor block since part of the current injected into the substrate bypasses the guard ring and reaches the victim, as depicted in Fig. 4.4(a). A common practice to improve the isolation efficiency is to divide the aggressor block into smaller sections and use a separate, smaller ring for each subblock, thereby reducing the overall impedance of the guard ring [114]. The disadvantage of using multiple rings is a substantial increase in overall area.

To better understand this tradeoff and determine the optimum number of guard

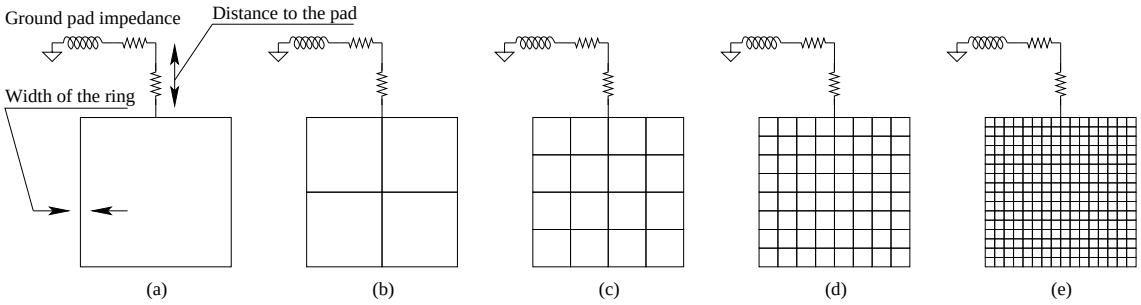


Figure 4.8: Five different guard ring isolation configurations: (a) one ring, (b) four rings, (c) 16 rings, (d) 64 rings, and (e) 256 rings.

rings, the proposed compact models are utilized to analyze noise isolation for a different number of guard rings. A circuit block consisting of 256 standard cells is assumed as the aggressor. As shown in Fig. 4.8, the noise isolation efficiency is determined for five different cases: one, four, 16, 64, and 256 rings.

The overall area is maintained constant to provide a fair comparison of the noise isolation efficiency. Under this constraint, an additional ring is placed at the expense of a reduced guard ring width. The overall fixed area (based on a single ring) is

$$Area = \left(2 \cdot width_{gr} + \sqrt{n} \cdot width_{scell} \cdot\right)^2, \tag{4.5}$$

where $width_{gr}$ and $width_{scell}$ are, respectively, the width of a guard ring and standard cell. n is the number of standard cells within the large circuit block and is equal to 256 for these examples. Since the width of a standard cell is constant, placing multiple guard rings requires a reduction in $width_{gr}$ to maintain the same overall area.

Table 4.1: Five guard ring isolation configurations are evaluated under a constant area constraint. The guard rings are located in close proximity to the ground pad.

Isolation configuration	One ring	Four rings	16 rings	64 rings	256 rings
$width_{gr}$	1	2/3	2/5	2/9	2/17
TN [dB]	-33.0	-35.9	-31.4	-27.1	-23.7
$width_{gr}$	2	4/3	4/5	4/9	4/17
TN [dB]	-44.9	-47.4	-42.2	-36.7	-32.1
$width_{gr}$	4	8/3	8/5	8/9	8/17
TN [dB]	-56.3	-58.1	-52.5	-46.3	-40.5

*TN - transferred noise

Results of the analysis obtained using the proposed models are listed in Table 4.1. For each case, the guard ring is located in close proximity to the ground pad. Thus, the ring width plays a dominant role in the overall impedance of the ring. To maintain a constant area, the ring width is reduced as the number of rings increases. Based on these results, utilizing four rings is the most efficient configuration since a further increase in the number of rings reduces the overall isolation efficiency due to the higher ring impedance. Note that the width of the guard rings is normalized to 2.5 μm.

In the second case, the same five guard ring configurations are analyzed where the distance between the guard ring and ground pad d_{gp} is varied. The results are listed in Table 4.2. A greater distance to the ground pad degrades the isolation efficiency. When $d_{gp} = 150$ μm, the number and width of the guard rings have significantly less effect on the isolation efficiency since the impedance of the ground network rather than the ring impedance is dominant.

Table 4.2: Five guard ring isolation configurations are evaluated under a constant area constraint. The guard rings are located at different distances to the ground pad.

	Isolation configuration	One ring	Four rings	16 rings	64 rings	256 rings
	$width_{gr}$	2	4/3	4/5	4/9	4/17
$d_{gp} = 1$ μm	TN [dB]	-44.9	-47.4	-42.2	-36.7	-32.1
$d_{gp} = 15$ μm	TN [dB]	-30.7	-30.8	-27.9	-25.1	-22.9
$d_{gp} = 150$ μm	TN [dB]	-10.2	-9.8	-8.8	-8.1	-7.5

*TN - transferred noise

Thus, dividing an aggressor block into smaller subblocks to improve the noise isolation produces the expected results when the ring impedance dominates over the impedance of the ground network. In this case, an optimum number of rings exists under a constant area constraint. This optimum number is a weak function of the guard ring width.

4.3 Comparison Among Multiple Isolation Schemes

The comparison is performed using a large scale block within an industrial transceiver circuit designed in a 90 nm CMOS technology with a lightly doped substrate. The isolation obtained from several different configurations are evaluated in this section using the proposed models. Specifically, five cases are compared for the same industrial mixed-signal circuit: (1) no isolation, (2) both aggressor and victim are surrounded with guard rings with dedicated ground pads, (3) the guard ring of the

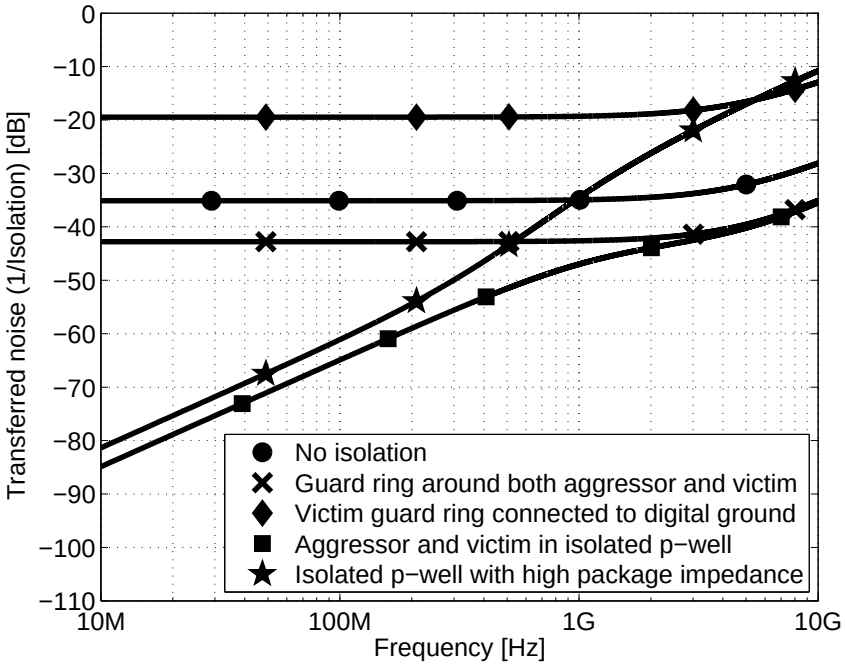


Figure 4.9: Comparison of the isolation efficiency of several different configurations obtained using the proposed models.

victim is connected to the digital ground pad, (4) the aggressor and victim are both placed in an isolated p-well with a low package impedance, and (5) the aggressor and victim are both placed in an isolated p-well with a high package impedance. Note that the physical area of each case is maintained equal to provide a fair comparison. The results are illustrated in Fig. 4.9. An important observation is that the isolation obtained in the third case is worse than no isolation since the switching noise of the digital circuit directly couples to the victim. A guard ring should therefore have a dedicated ground pad to be effective. Another interesting observation is the effect of the package impedance on the efficiency of the triple well. Specifically, if the package

impedance is sufficiently high, the first case (no isolation) achieves a better result at higher frequencies than a triple well. This result is due to the greater junction capacitance since the size of the local n-well is larger when a deep n-well is used, as shown in Fig. 4.6. Also note that a triple well is significantly more effective than a guard ring at lower frequencies due to capacitive isolation, but the difference in efficiency diminishes with higher frequency.

4.4 Summary

Closed-form compact expressions have been proposed to efficiently estimate the substrate resistance within a lightly doped substrate by approximating the current flow with a half-ellipse. Only one fitting parameter is required as opposed to existing models that use multiple fitting parameters. The proposed models are used to (1) determine the optimum number of guard rings for a large aggressor block under a constant area constraint and (2) develop a circuit description of common signal isolation structures such as a guard ring and triple well. The efficiency of these isolation structures is accurately characterized using the proposed models and is evaluated on an industrial mixed-signal circuit. These results can be used as guidelines to improve the overall signal integrity of mixed-signal circuits.

Chapter 5

Simultaneous Shield and Repeater Insertion

Further increases in integrated circuit (IC) scaling require more efficient devices, circuits, and systems in terms of power, delay, noise, and area. Efficient optimization processes are therefore required. To achieve this capability, many different design techniques are used. In many cases, only one technique is implemented; however, two or more techniques applied simultaneously may provide higher performance. A methodology that considers multiple design objectives while satisfying system requirements typically utilizes lower resources. Optimization processes and related design techniques applied to high performance ICs are the topic of this chapter.

A standard optimization process is based on a *cost* function (also known as an *objective* function). There are two steps involved in this process, building a function and determining the optimal value of the function. In this chapter, a different, resource based, optimization process is presented.

Any design constraint may be characterized as a resource. Some constraints, such as power and area, are more commonly treated as a resource. Other design objectives, such as delay or noise, are less commonly referred to as a resource. A practical application is composed of a combination of optimization processes and multiple design techniques. A methodology that considers these issues in an integrated fashion is the focus of this chapter. Two different techniques that provide immunity to coupled noise, shield and repeater insertion, have been combined based on resource optimization to exemplify this process. Each of the techniques exhibits different power, delay, noise, and area resource characteristics.

This chapter is organized as follows. Limitations to the standard optimization process that motivates resource based optimization processes are described in Section 5.1. This process is simultaneously applied to shield and repeater insertion in Section 5.2. Each resource model is also reviewed in this section. A practical case study is presented in Section 5.3. In Section 5.4, simultaneous shield and repeater insertion techniques are compared with only shielding and only repeater insertion. Finally, this chapter is concluded in Section 5.5.

5.1 Resource Based Optimization Process

Limitations in standard optimization processes are described in subsection 5.1.1. The theory and limitations of resource based optimization processes are presented in

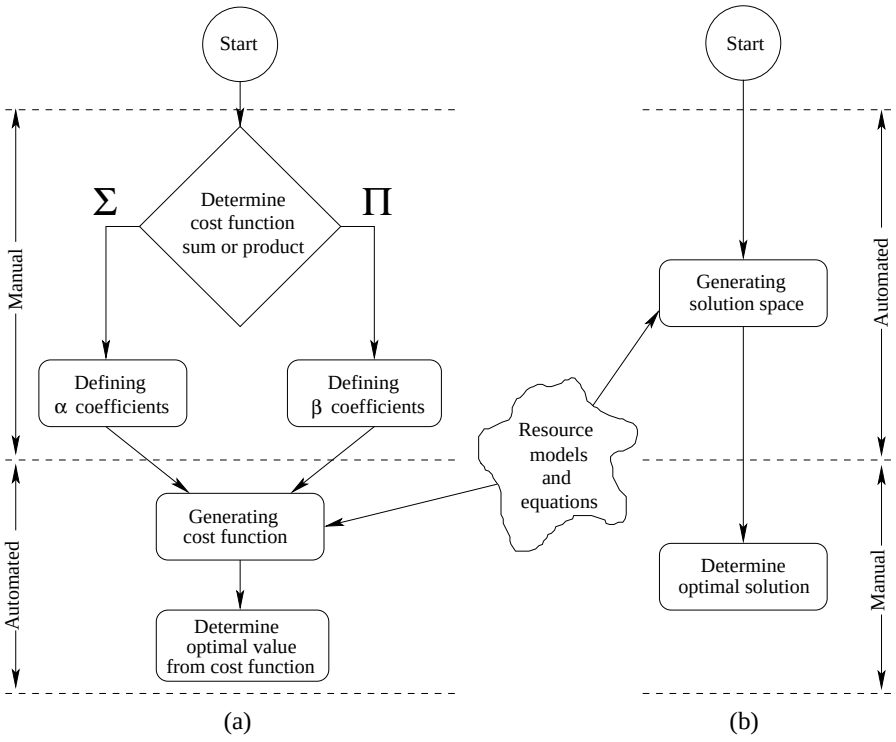


Figure 5.1: Optimization flow diagram. (a) Standard and (b) resource based optimization process.

subsection 5.1.2 and 5.1.3, respectively. Different design techniques are introduced in subsection 5.1.4.

5.1.1 Limitations in Standard Optimization Processes

A general flow for a standard optimization process is shown in Fig. 5.1(a). The primary disadvantage of this flow is the need for user involvement before the optimization process is initiated. The *cost* function and coefficients must be allocated for each resource. For the same system, two users may choose different coefficients

and thereby produce different results. Additionally, some resources have changing importance. These aspects constrain the standard optimization process.

5.1.2 Resource Based Optimization Processes

To overcome these limitations, a different resource based optimization process is proposed. The user involvement occurs at the end of this process. In Fig. 5.1(b), a flow diagram of this resource based optimization process is presented.

In order to provide insight into the resource based optimization flow, consider a system where

$$area = f_1(width), \tag{5.1}$$

$$noise = f_2(width). \tag{5.2}$$

A fundamental assumption in (5.1) and (5.2) is that the width determines the area and noise. Conversely, the area or noise may determine the width. By inverting (5.1), the same system is described by

$$width = f_1^{-1}(area), \tag{5.3}$$

$$noise = f_2(width). \tag{5.4}$$

Substituting (5.3) into (5.4), the same system can be characterized by

$$noise = f_2\left[f_1^{-1}(area)\right]. \tag{5.5}$$

This system representation describes the relationship between the two resources and can be presented as a tradeoff line.

Power, area, noise, and delay are four primary design criteria. The number of variables, *e.g.*, line width, shield width, number of repeaters, and power supply, is typically much greater. Any system can be represented by n variables and $n+1$ resources,

$$\left.\begin{aligned} res_1 &= f_1(a_1, a_2, a_3, \ldots, a_n) \\ res_2 &= f_2(a_1, a_2, a_3, \ldots, a_n) \\ &\vdots \\ res_n &= f_n(a_1, a_2, a_3, \ldots, a_n) \\ res_{n+1} &= f_{n+1}(a_1, a_2, a_3, \ldots, a_n) \end{aligned}\right\}, \tag{5.6}$$

where $res_1, res_2, \ldots, res_{n+1}$ are the resources, such as power, delay, noise, and area, and $a_1, a_2, \ldots, a_n$ are variables, such as the line width, shield width, and length.

Inverting the first n equations in (5.6),

$$\left.\begin{aligned} a_1 &= g_1(res_1, res_2, \ldots, res_n) \\ a_2 &= g_2(res_1, res_2, \ldots, res_n) \\ &\vdots \\ a_n &= g_n(res_1, res_2, \ldots, res_n) \\ res_{n+1} &= f_{n+1}(a_1, a_2, a_3, \ldots, a_n) \end{aligned}\right\}. \tag{5.7}$$

To exemplify this process, if n equations in (5.6) are invertible, (5.7) describes the same system. The first n equations in (5.7) are substituted into the last equation in (5.7), resulting in

$$\begin{aligned} res_{n+1} = f_1[\, & g_1(res_1, res_2, \ldots, res_n), \\ & g_2(res_1, res_2, \ldots, res_n), \\ & \ldots, \\ & g_n(res_1, res_2, \ldots, res_n)]. \end{aligned} \tag{5.8}$$

Representing the system by (5.8), the interaction is among the resources and not among the design variables. The function described in (5.8) represents a solution space. The behavior of each resource among the other resources is referred to here as a tradeoff surface.

5.1.3 Limitations in Resource Based Optimization Processes

Resource based optimization also exhibits limitations. These limitations can be categorized as

- Model inaccuracies
- Function invertability

In a standard optimization process, inaccuracy in the models produces quantization error. In resource based optimization, however, this error is cumulative. Due to these additive errors, the models used in this optimization process must be sufficiently accurate. Otherwise, only the fidelity of the final function may be useful.

Function invertability is a different limitation in resource based optimization processes. For $y = f(x)$ where x can not be explicitly extracted, certain techniques are required to provide invertability. Some of these techniques are truncation, Taylor expansion, and approximation, which can lead to greater model inaccuracy.

In Section 5.2, a case study where these resource based process limitations are demonstrated is presented. The limitations are described and strategies for overcoming these constraints are provided.

5.1.4 Local Optimization Techniques

Several techniques have been proposed in the literature to overcome interconnect noise, such as shielding, repeater insertion, differential signaling, active regeneration, intentional skewing, and bus swizzling. Each of these techniques protect the interconnect from coupled noise in a different way and require different resources. The following section focuses on two commonly used techniques, shield and repeater insertion.

5.2 Shield And Repeater Insertion

Placing a shield beside and inserting repeaters along a victim line are chosen to exemplify the resource based optimization process. The width of the shield line, and the number and size of the repeaters are chosen to express noise on the victim line as a function of power, area, and delay resources. Repeater insertion, shielding, and basic resource expressions are summarized in the following section.

As compared to [115] where a *cost function* is used, this work is based on resource optimization. In [115], two different design techniques, shield and repeater insertion, are used to enhance noise characteristics without incurring a large penalty in area. The problem has been divided into a number of sections, as illustrated in Fig. 5.2, where only a single design technique is used for each section. The primary disad-

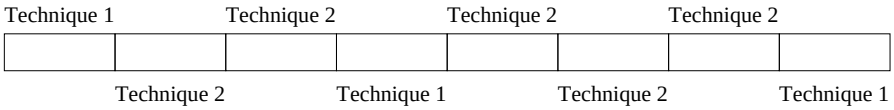


Figure 5.2: Problem is divided into a number of smaller sections, where only a single design technique is applied for each section.

vantage of this approach is the result will likely be non-optimal, depending upon the order of the design techniques being used. A different order of implementation may yield different results. In [115], the noise is modeled based on the Devgan metric [116], while in this work the shield noise model is based on [117].

5.2.1 Repeater Insertion

Repeater insertion is a well known design technique to reduce the delay required to propagate a signal along a line [29]. The objective is to divide the interconnect into smaller sections, reducing the quadratic delay dependency on length to a linear dependency, thereby reducing the overall delay [57]. If the number of repeaters is too small, the delay due to the interconnect will be dominant. If the number of repeaters is too large, the repeater delay dominates. The optimal number of repeaters that minimizes the overall delay has been presented in [29, 9], and [57].

An additional advantage of repeater insertion is reducing the coupled noise from adjacent interconnects. It is impractical, however, to insert excessive repeaters due to delay, power and area constraints.

5.2.2 Shielding

Shielding inserts an additional line between a victim line and an aggressor line. This technique can be divided into two major categories: passive and active shielding [61]. The focus of this chapter is on passive shielding. A passive shield line is connected to the power/ground network, filtering the noise from the aggressor away from the victim line. The technique is highly effective, although significant area is required.

5.2.3 Resources

Four primary resources for simultaneous shield and repeater insertion are considered: power, delay, noise, and area. In this chapter, the resource models are based on a 0.18 μm CMOS technology.

a. Power

Two primary power dissipation sources are considered. The first source, dynamic power, is used to charge and discharge the interconnect and transistor capacitances. The second source, short-circuit power, also occurs when the transistors switch. During the switching time, the current from the power to ground network passes through the NMOS and PMOS transistors. This power component is typically in the range of 5% to 10% of the overall transient power. The total transient power is the summation

of the dynamic and short-circuit power,

$$power = power_{dyn} + power_{sc}. \tag{5.9}$$

The dynamic power is

$$power_{dyn} = \alpha C_{eff} V_{dd}^2 f, \tag{5.10}$$

where V_{dd} and f are, respectively, the power supply voltage and operating frequency. α is a switching coefficient characterizing the switching behavior and C_{eff} is the effective capacitance,

$$C_{eff} = k\left(\frac{C_{line}}{k} + C_{transistor}\right) = C_{line} + c_o hk. \tag{5.11}$$

C_{line}, c_o, h, and k are the line capacitance, minimum gate capacitance, ratio between the final and minimal transistor widths, and the number of inserted repeaters along the victim line, respectively. The short-circuit power for one transistor is [57]

$$power_{sc} = \left|\ln\left(\frac{v_{tn}}{V_{dd} + v_{tp}}\right)\right| \frac{C + \vartheta_{do} RC}{\vartheta_{do}} I_{peak} f V_{dd}, \tag{5.12}$$

where v_{tn} and v_{tp} are, respectively, the threshold voltage of the NMOS and PMOS transistors. R and C are the lumped load resistance and capacitance, respectively. ϑ_{do} is the saturation velocity, also defined in [57], and I_{peak} is the maximum saturation

current of the switching transistor and is expressed as

$$I_{peak} = \frac{\mu_n c_{ox}}{2} \frac{w}{l} \left(\frac{V_{dd}}{2} - v_{tn} \right)^2, \tag{5.13}$$

where μ_n, c_{ox}, w, and l are the N-type mobility, oxide capacitance, width, and length of the transistor, respectively. Expressing (5.12) in h and k, the following terms are substituted.

$$C = c_o h + \tfrac{c_{int}}{k}, \tag{5.14}$$

$$R = \tfrac{r_{int}}{k}, \tag{5.15}$$

$$\vartheta_{do} = \vartheta_{do_o} h = \tfrac{h}{r_o}, \tag{5.16}$$

$$w = w_o h, \tag{5.17}$$

where r_o, w_o, and ϑ_{do_o} represent, respectively, the minimum resistance, width, and saturation velocity of the transistor. r_{int} and c_{int} are the total resistance and capacitance of the victim line, respectively. The NMOS and PMOS threshold voltages are assumed to be equal, permitting the total short-circuit power to be expressed as

$$power_{sc} = k \left| \ln \left(\frac{v_t}{V_{dd} + v_t} \right) \right| \frac{(c_o h + \frac{c_{int}}{k})(1 + \frac{h}{r_o} \frac{r_{int}}{k})}{\frac{h}{r_o}} \cdot \cdot \frac{\mu_n c_{ox}}{2} \frac{h w_o}{l} \left(\frac{V_{dd}}{2} - v_t \right)^2 f V_{dd}. \tag{5.18}$$

b. Delay

Minimizing the overall interconnect delay in a repeater system has been investigated in [29]. In [57], a more accurate delay expression is presented based on the saturation velocity characteristic,

$$delay = k\alpha_1 \frac{C + \vartheta_{do} RC}{\vartheta_{do}}, \tag{5.19}$$

where α_1 is relative to the propagation delay and equal to 0.693 for 50% of the voltage waveform (or 2.3 for 90%). Substituting (5.14) to (5.16) into (5.19), the signal propagation delay is

$$delay = k\alpha_1 \frac{\left(c_o h + \frac{c_{int}}{k}\right)\left(1 + \frac{h}{r_o}\frac{rint}{k}\right)}{\frac{h}{r_o}}. \tag{5.20}$$

Two resources, *power* and *delay*, only affect the repeater insertion process. The two resources, *noise* and *area*, are defined simultaneously for both shield and repeater insertion.

c. Noise

Noise modeling in shielded interconnect has been investigated in [30, 117]. From the shield model used in [117] and illustrated in Fig. 5.3, the noise as a function of the shield line width is approximated by

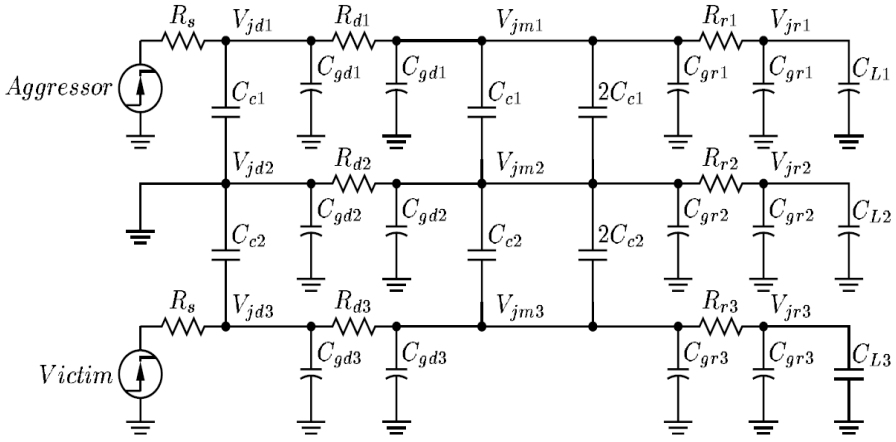


Figure 5.3: Model of shielding effect with coupling noise [30].

$$noise_{sh} = C_1 e^{-C_2 w_{sh}}, \tag{5.21}$$

where w_{sh} is the width of the shield line, and C_1 and C_2 are constants extracted from the model. The noise voltage is normalized to V_{dd}, beginning from C_1 with no shield line present ($w_{sh} = 0$), and exponentially decreasing with wider shield lines. The exponential term emphasizes the effectiveness of this technique. Repeater insertion divides the overall length of the line into smaller sections. Assuming a uniform distribution of the noise along the victim line, the total noise of the line is

$$noise_{rep} = \frac{1}{k}, \tag{5.22}$$

dividing the noise by the number of inserted repeaters. The total effect of inserting a shield line and repeaters is expressed as a product,

$$noise = noise_{sh} \cdot noise_{rep} = C_1 e^{-C_2 w_{sh}} \frac{1}{k}. \tag{5.23}$$

d. Area

A schematic layout of a shielded line with repeaters is presented in Fig. 5.4. The width ratio between the PMOS and NMOS transistors is three. The PMOS transistor is designed in a stack structure to reduce the overall width. Half of the NMOS and PMOS transistors are under the signal line, resulting in a total repeater width of hw_o. Note that the power, ground, and aggressor lines are not shown and are not considered in the area expression. The area of the structure illustrated in Fig. 5.4 is

$$area = length(w_{line} + w_{sh} + s + hw_o), \tag{5.24}$$

where $length$, w_{line}, and s represent the total length, signal line width, and spacing between the signal line and shield line, respectively. Two terms in this equations, h and w_{sh}, are the design variables.

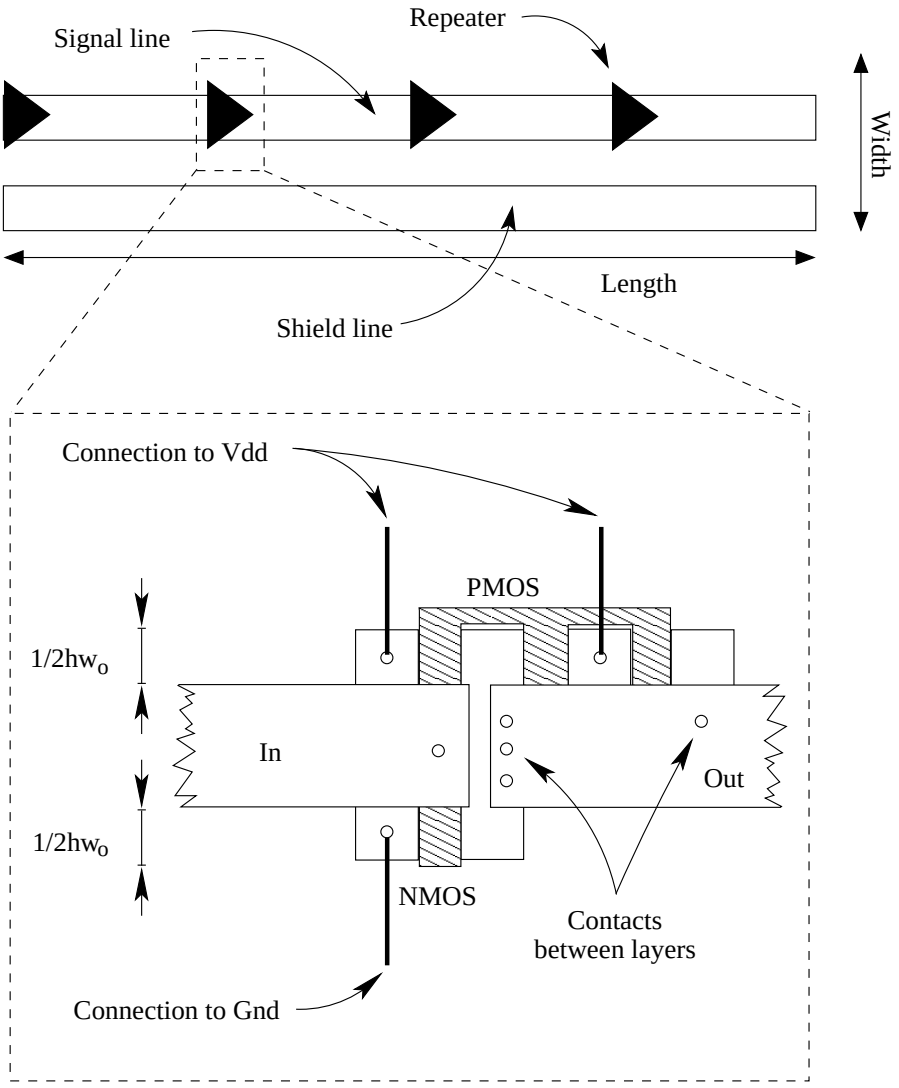


Figure 5.4: Schematic layout of a signal line with shield line and repeaters to reduce coupling noise.

5.2.4 Coupling Noise with Resource Based Optimization

The resource models are summarized in (5.25) to (5.28) and expressed in terms of the resources and variables,

$$power = f_1(h, k), \tag{5.25}$$

$$delay = f_2(h, k), \tag{5.26}$$

$$noise = f_3(w_{sh}, k), \tag{5.27}$$

$$area = f_4(w_{sh}, h). \tag{5.28}$$

Due to the two common variables, a resource based optimization procedure is initiated with (5.25) and (5.26). The overall power equations are noninvertible, demonstrating the limitation of this procedure. The truncation method is therefore used, where the short-circuit power term is dropped, resulting in a successful inversion,

$$h = g_1(power_{dyn}, delay), \tag{5.29}$$

$$k = g_2(power_{dyn}, delay). \tag{5.30}$$

The *power* becomes $power_{dyn}$ to emphasize that only dynamic power is considered. The short-circuit power is added later in the procedure. Equations (5.29) and (5.30)

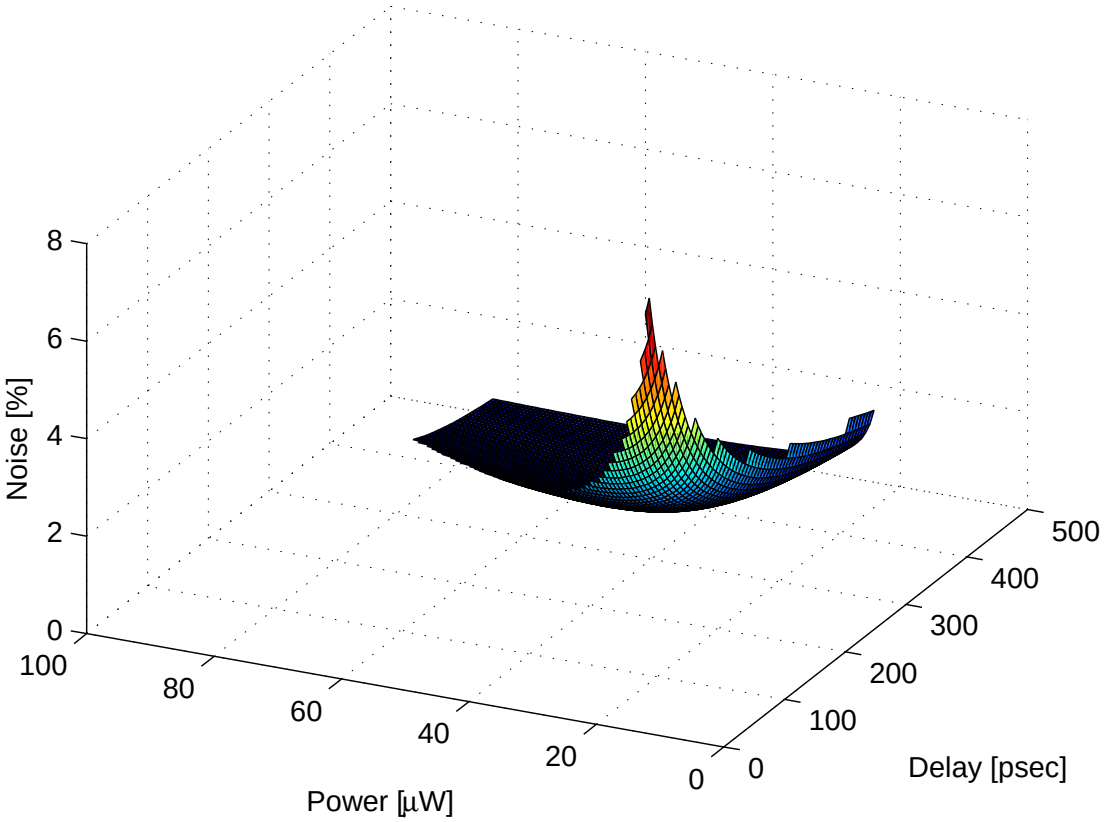


Figure 5.5: Noise as a function of power and delay in a system with shields and repeaters.

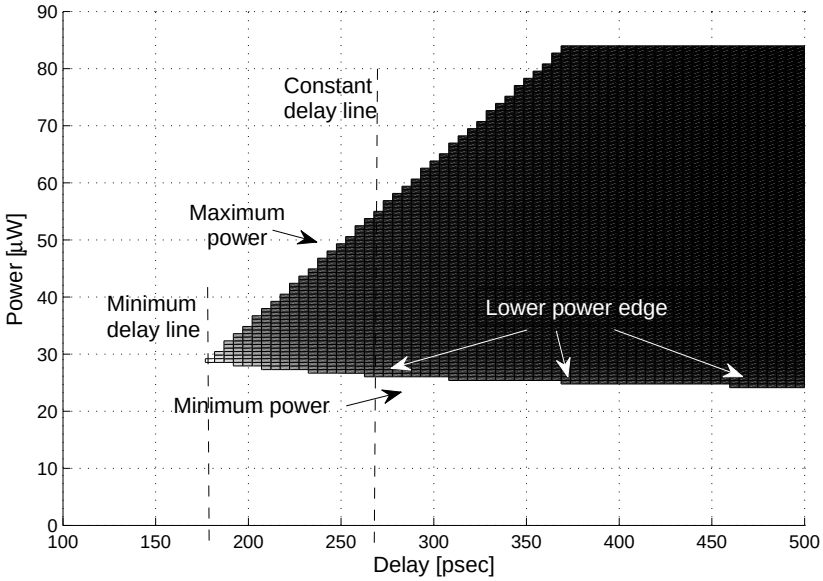


Figure 5.6: Top view of Fig. 5.5. The lighter color represents a larger amount of noise.

are substituted into (5.28),

$$area = f_4(w_{sh}, g_1(power_{dyn}, delay)). \tag{5.31}$$

Inverting (5.31), the width of the shield line is

$$w_{sh} = g_4(area, power_{dyn}, delay). \tag{5.32}$$

Substituting (5.29), (5.30), and (5.32) into (5.27), the noise function is

$$noise = F(area, power_{dyn}, delay). \tag{5.33}$$

Note that the noise is not a function of the number or size of the repeaters or width of the shield line.

5.3 Simulation Results

A case study with inserted repeaters and a shielded victim line is considered. The area, power, delay, and noise are evaluated for this system. Several physical parameters are chosen to reflect practical design characteristics. Specifically, s = 0.5 μm, $lenght$ = 1 mm, V_{dd} = 1.8 volts, v_t = 0.5 volts, l = 0.18 μm, c_{int} = 250 fF, r_{int} = 11 Ω, w_{line} = 2 μm, w_o = 0.5 μm, C_1 = 7.25%, and $C_2 = 1.33 \cdot 10^6$ m^{-1}. By increasing the area, the noise is reduced since wider shield lines and additional repeaters are possible. The noise monotonically decreases as a function of area; therefore, the area is set to a value of $4.15e3$ μm^2, a practical design value.

Each solution of (5.33) represents a specific h, k, and w_{sh}, which determines the short-circuit power from (5.18). The short-circuit power is added to the dynamic power, permitting the overall power dissipation to be estimated.

A graph presenting *noise* as a function of *power* and *delay* is illustrated in Fig. 5.5. Note the relationship among power, delay, and noise, generating a tradeoff surface, permitting different tradeoffs to be made. The top view of the graph illustrated in Fig. 5.5 is shown in Fig. 5.6, where the lighter region indicates a higher noise. For this design case, a 180 psec delay is the minimum delay, as depicted in Fig. 5.6. This delay

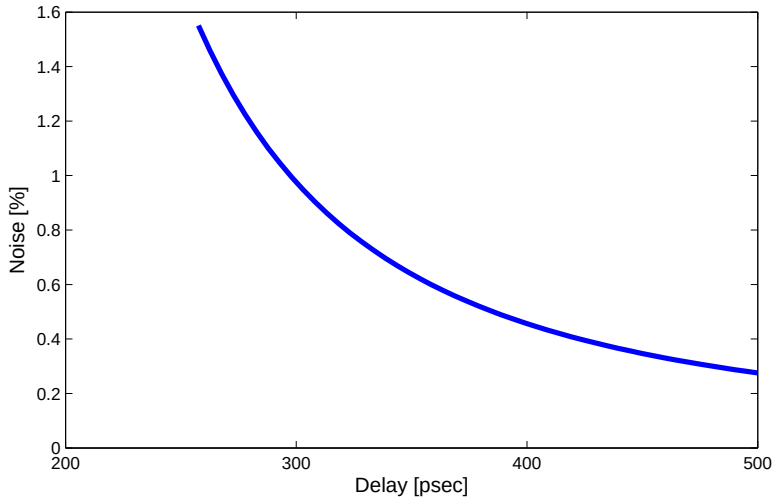


Figure 5.7: Noise as a function of delay at a constant power (50 μW) and maximum allowed area (4.15e3 μm^2).

is not the same as determined in [29, 9] and [57], since power, noise, and area are also considered. The lower edge of the power curve, illustrated in Fig. 5.6, saturates to a minimum power value. This curve does not reach zero due to the minimum power required to charge and discharge the line capacitance.

In Fig. 5.7, noise is presented as a function of delay at a constant power and maximum allowed area. An increase in delay will reduce the coupling noise, since more repeaters or wider shield lines are available. The exponentially increasing curve (with decreasing delay), illustrated in Fig. 5.7, indicates the noise penalty from choosing a value close to the minimum delay. Note that by relaxing the delay constraint, the coupling noise is significantly smaller.

Noise as a function of power at the maximum allowed delay and area is illustrated

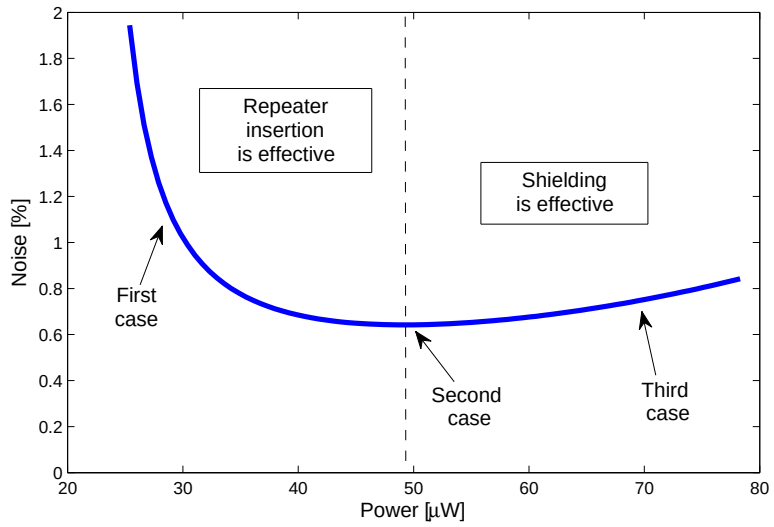


Figure 5.8: Noise as a function of power at the maximum allowed delay (350 psec) and area (4.15e3 μm^2).

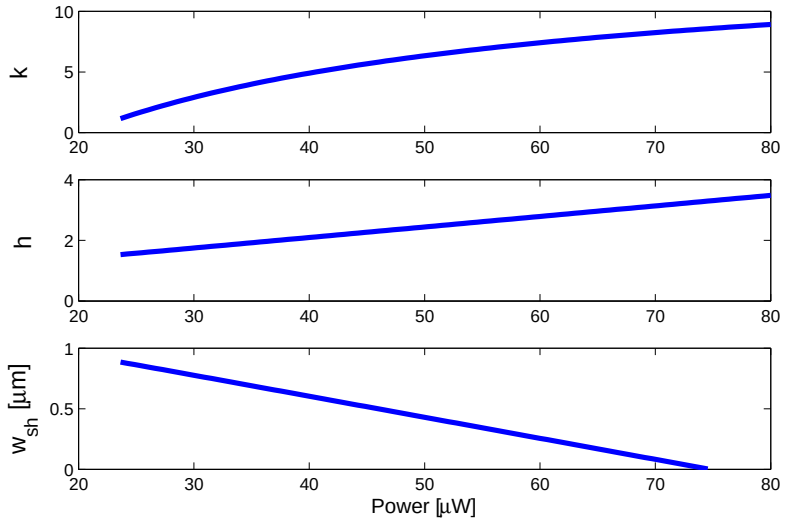


Figure 5.9: k, h, and w_{sh} as a function of power at the maximum delay (350 psec) and area (4.15e3 μm^2).

in Fig. 5.8. The graph consists of two different regions: the noise is reduced and the noise is increased by increasing the power. This parabolic noise behavior can be exploited to determine the minimum noise for this circuit. To motivate these results, three cases, depicted in Fig. 5.8, have been evaluated. The first case, at a power of 29 μW, produces a 1.1% noise (normalized to V_{dd}). The noise voltage in this case is 21 mV. The noise for the second case located at a power of 49 μW is 0.65% (or 11.5 mV). The final case at a power of 70 μW produces 0.8% (or 14 mV) noise. The 10 mV noise difference between the first and second case exemplifies the tradeoff. The noise difference between the second and third case is smaller, but significant.

The effects of k (number of repeaters), h (width of the repeater), and w_{sh} (width of shielding line) as a function of power are illustrated in Fig. 5.9. The area and delay are maintained at maximum values. With an increase in the power, the number and width of the repeaters increase at a different rate, maintaining a constant delay. Simultaneously, the width of the shield lines decreases, providing more space for larger repeaters while maintaining the area constant. The larger number of repeaters reduces the noise; however, the reduction in the shield width increases the noise. Adding repeaters at lower power levels reduces the noise more than adding repeaters at higher power levels. Hence, at lower power levels, the most efficient noise reduction technique is repeaters, while at higher power levels, the most efficient noise reduction technique is shield lines, as illustrated in Fig. 5.8. Both of these techniques reduce the

Table 5.1: Three design cases shown in Fig. 5.8 evaluated in SPICE

	k (number of repeaters)	h·0.5 (width of the repeaters)	w_{sh} (width of the shield line)
First case	2	0.8 μm	0.8 μm
Second case	6	1.2 μm	0.5 μm
Third case	8	1.5 μm	0.1 μm

noise, exhibiting a parabolic noise behavior, allowing the minimum noise design to be determined. In this case, the minimum noise is at 49 μW total power and contributes only 0.65% (or 11.5 mV) noise.

This concept is evaluated on a system composed of a victim interconnect with several repeaters and a shield line. Three design cases, listed in Table 5.1, are considered.

The power, delay, and noise are determined from SPICE simulations. The analytic model and SPICE results are compared in Fig. 5.10 and Table 5.2 for three cases, listed in Table 5.1 and depicted in Fig. 5.8. In Table 5.2, the change in delay, power, and noise is determined relative to the minimum noise design case (second case). In the analytic model, the delay is maintained constant; however, small changes in the delay are noted from SPICE. The power resulting from the analytic model and SPICE is similar. The noise evaluated from SPICE also exhibits good agreement with the analytic model. The SPICE results demonstrate the same parabolic noise behavior when simultaneously applying shield and repeater insertion. The noise is lower in

Table 5.2: Analytic and SPICE results for three design cases from Table 5.1 and Fig. 5.8

	k	h	w_{sh}	Delay [psec]	Change in Delay [%]	Power [μW]	Change in Power [%]	Noise [mV]	Change in Noise [%]
Analytic	2.04	1.63	0.83	350	0.0	28.9	41.1	21.1	82.5
	5.91	2.33	0.48	350		49.0		11.6	
	8.04	3.04	0.13	350	0.0	69.6	42.1	13.7	18.2
SPICE	2	1.63	0.83	520	6.6	44.9	22.0	15.7	73.0
	6	2.33	0.48	557		57.6		9.1	
	8	3.04	0.13	563	1.1	76.6	33.1	14.0	54.5

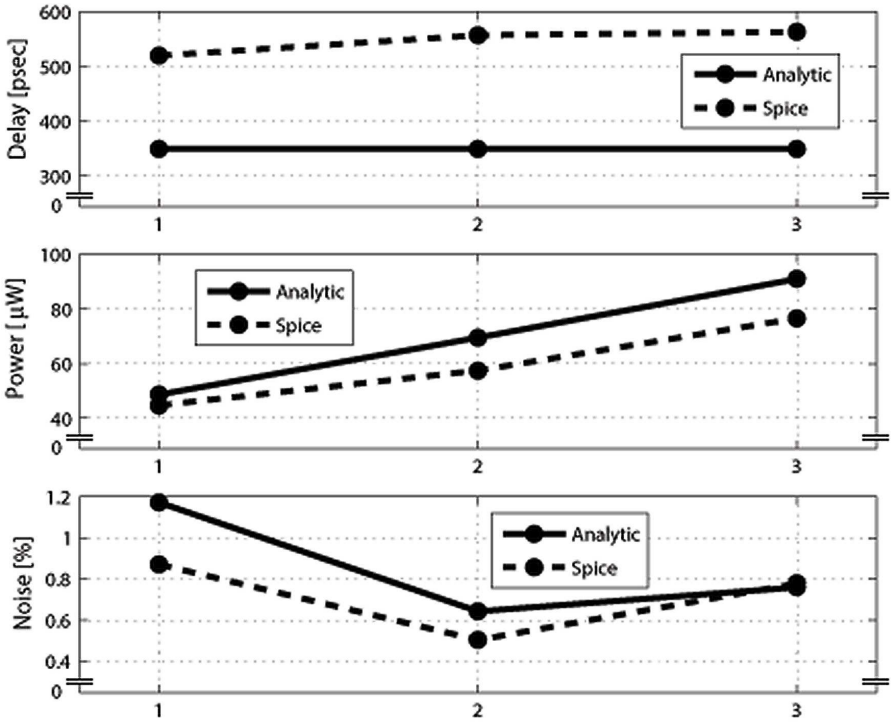


Figure 5.10: Delay, power, and noise for three different design cases. Analytic and SPICE results are compared.

the second design case than the first and third design cases, confirming the parabolic noise behavior. The minimum noise is achieved with simultaneous shield and repeater insertion, while satisfying power, area, and delay constraints.

5.4 Comparison of Shield and Repeater Techniques

A comparison of simultaneous shield and repeater insertion with only shielding (without repeater insertion) and only repeater insertion (without shielding) is discussed in this section. The same resources are compared: power, delay, area, and noise. A constant area of 4.15e3 μm^2 is assumed.

In only shielding, all of the area except for the victim line and spacing is dedicated to the shield line. A 1.65 μm shield line is inserted between the aggressor and victim lines. The reduction in coupled noise is only due to the shield line and, according to (5.23), when $k = 1$ (a single driver repeater) the coupled noise is 0.81% (or 14.5 mV). The power dissipation is minimal, only 22.2 μW, since dynamic power is only dissipated for the line and driver repeater, and a small amount of short-circuit power to switch the driver repeater. The delay, however, increases to 515 psec.

In the repeater insertion case (without shielding), emphasis is placed on achieving a target delay of 350 psec, as in the simultaneous shield and repeater insertion case. Consequentially, minimum noise is targeted. To minimize the noise, the largest number of repeaters is required. To satisfy the target delay and area constraints, the highest number of repeaters is determined to be ten. In this case, all of the area is occupied by the repeaters. The noise is reduced from $C_1 = 7.25\%$ to $7.25\% \cdot 1/10 = 0.725\%$ or

Table 5.3: Comparison among shielding, repeater insertion, and shield and repeater insertion techniques. For each technique, the area is maintained equal to provide a fair comparison. For the only repeater and simultaneous shield and repeater insertion techniques, 350 psec is the target delay.

	Noise	Area	Power	Delay
Only Shielding	14.5 mV	4.15e3 μm^2	22.2 μW	515 psec
Only Repeaters	13.0 mV	4.15e3 μm^2	86.7 μW	350 psec
Simultaneous Shield and Repeater Insertion	11.5 mV	4.15e3 μm^2	49.0 μW	350 psec

13 mV. The power consumption for this system is comparably high, 86.7 μW. The results are compared in Table 5.3.

Note in Table 5.3 that the noise is similar among all of the cases. A noise advantage of two to three millivolts is determined for the simultaneous shield and repeater insertion case. If the delay is not constrained, the more appropriate technique is shielding only, since minimal power is dissipated in this case. In the cases where delay is also considered, the only repeater insertion technique achieves the target delay with comparable noise performance. The power dissipation, however, is almost twice that of the simultaneous shield and repeater case.

5.5 Summary

Resource based optimization is described and compared to standard optimization processes in this chapter. The resource based optimization process is evaluated for a system that simultaneously considers shield and repeater insertion. The methodology is used to investigate area, power, delay, and noise tradeoffs. The coupled noise as a function of power with maximum allowed delay and area is evaluated, demonstrating a parabolic behavior. This approach permits the minimum noise design to be determined. The analytic model exhibits good agreement with SPICE. Over 50% reduction in coupled noise is demonstrated as compared to three design cases by applying this resource based optimization process. To motivate simultaneous shield and repeater insertion, three cases are evaluated and compared: shielding only, repeater insertion only, and simultaneous shield and repeater insertion. The noise performance is comparable among all of these techniques. In only shielding, however, the delay is higher, while in only repeater insertion, the power is higher. In practical cases where the delay, power, and area are constrained, simultaneous shield and repeater insertion exhibits the best performances.

Chapter 6

Interdigitated Power/Ground Network - A Single Metal Layer

With high operating frequencies and scaled geometries, the power and ground (P/G) distribution network requires enhanced optimization to provide significant current flow. The higher currents increase voltage losses within the P/G network, while decreasing power supply voltages provide lower noise margins.

With flip-chip packaging, the package inductance is reduced [118], making the on-chip inductance more significant. Since the voltage variations within a P/G network are due to IR [119, 120] and $L\frac{di}{dt}$ [121] voltage drops, the effective resistance and inductance are the primary foci of the optimization process. At higher frequencies, the inductive impedance is dominant, requiring accurate estimation of the effective inductance.

An interdigitated P/G distribution network structure where a few wide lines are replaced by a large number of narrow lines is often used to reduce the inductive

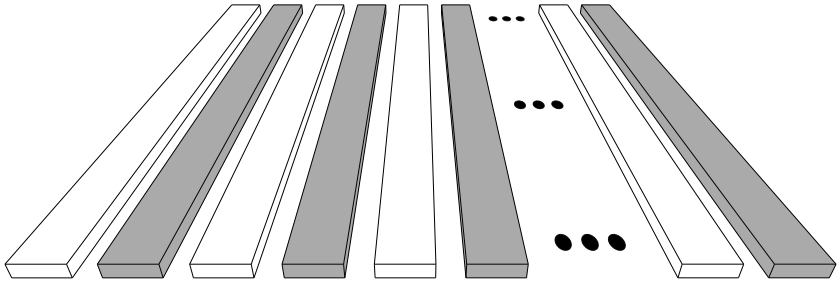


Figure 6.1: A single metal layer of an interdigitated P/G distribution structure. The darker and lighter lines represent the power and ground lines, respectively.

effect [122, 44]. Different P/G structures have been compared in [72, 91], where the interdigitated structure is shown to achieve the greatest reduction in inductance.

Each layer of a multi-layer interdigitated structure consists of individual interdigitated power and ground wires. A single metal layer is depicted in Fig. 6.1.

The need for efficient P/G networks has been recognized, and several techniques to optimize the P/G distribution network have been developed [123, 124, 125], focusing only on the resistive component. More advanced techniques for designing power networks have been described [126, 127]; however, only the package inductance is considered, neglecting the on-chip inductance. With more advanced packaging techniques (such as flip-chip), the on-chip inductive noise ($L\frac{di}{dt}$) is also important [118, 128]. To consider on-chip inductance in power supply networks, a technique to simplify the mesh model of an RLC power network is proposed [90], assuming the loads are treated as identical current sources. The significance of the on-chip inductance within paired

and interdigitated power/ground network structures is described in [129], where the inductance is treated as a local effect. In [130], the inductance model considers the mutual inductance between close and distant power/ground wires in interdigitated structures. Based on this model, a closed-form expression characterizing a single layer of an interdigitated P/G network structure is utilized, permitting the optimal width of a power/ground network that minimizes the network impedance to be determined.

This chapter is organized as follows. The problem is formulated in Section 6.1. The resistance and inductance of an interdigitated power and ground structure is discussed in Section 6.2. In Section 6.3, a closed-form expression for the line width that produces the minimum impedance for a single metal layer is presented. The accuracy of the optimal line width and related issues are discussed in Section 6.4. This chapter is concluded in Section 6.5.

6.1 Problem Formulation

An interdigitated P/G distribution network is typically allocated over an entire upper metal layer, where the network is designed for minimum impedance. The overall impedance of a power/ground network is

$$Z_{eff} = R_{eff} + j2\pi f L_{eff}, \tag{6.1}$$

where R_{eff} and L_{eff} are, respectively, the effective resistance and inductance of the network. f is the target frequency.

The objective is to minimize the overall impedance of a network at a specific frequency, reducing the voltage drop on the power/ground network. For constant area, two design variables, the line-to-line spacing s and width of the lines w, characterize a single metal layer of an interdigitated power/ground network. The absolute value of the effective impedance as a function of line-to-line spacing and line width is

$$|Z_{eff}(s,w)| = \sqrt{R_{eff}^2(s,w) + 4\pi^2 f^2 L_{eff}^2(s,w)}. \tag{6.2}$$

For constant area, different line widths, shown in Fig. 6.2, produce a different network impedance. A higher line width reduces the resistance, however, the inductance increases due to the fewer number of interdigitated power/ground pairs. The primary objective is to determine the width that produces the lowest impedance for a single metal layer within an interdigitated power/ground distribution network.

6.2 Impedance of an Interdigitated P/G Distribution Network

The overall impedance of a power/ground network is a combination of the effective resistance and inductance. In the following subsections, a model for the effective

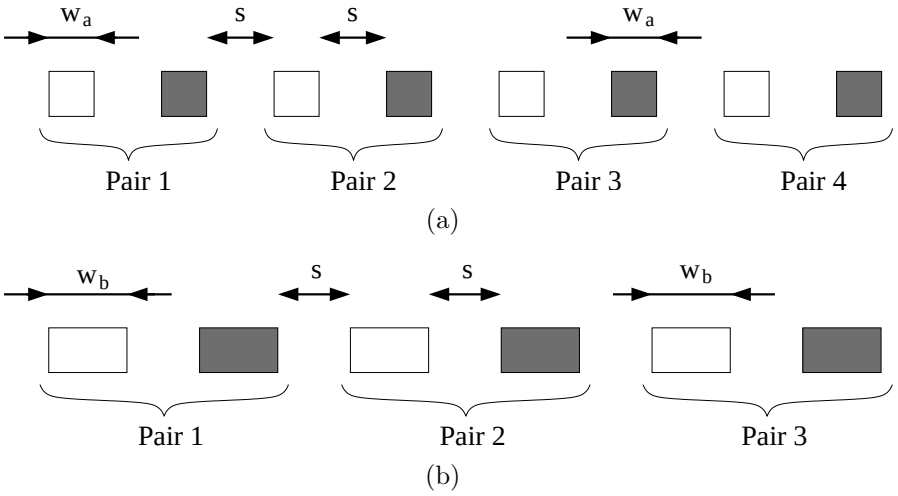


Figure 6.2: For the same physical area, two different interdigitated power/ground networks are presented. The spacing s is maintained constant. The width w is different for each network, where the width w_a of network (a) is thinner than the width w_b of network (b). Increasing the width requires fewer interdigitated power/ground line pairs since the area is maintained constant.

resistance and inductance of an interdigitated network assuming a constant area is presented.

6.2.1 Effective Resistance

The overall area allocated for a power/ground network is

$$A = l\left[N\left(w + s + w + s\right)\right] = 2lN\left(w + s\right), \tag{6.3}$$

where l and N are, respectively, the length of a single power or ground wire and the number of interdigitated pairs. Since N pairs of an interdigitated power and ground wires are in parallel, the effective resistance of the network R_{eff} is

$$R_{eff} = \frac{1}{N}\rho\frac{2l}{tw}, \tag{6.4}$$

where t is the thickness of the metal. Substituting (6.3) into (6.4), the effective resistance for a constant area is

$$R_{eff} = \frac{4l(w+s)}{A}\rho\frac{l}{tw}, \tag{6.5}$$

where ρ is the metal resistivity.

Due to the skin effect, the resistance increases with higher signal frequencies. In

this chapter, however, the skin effect is neglected since the highest target frequency is 10 GHz, resulting in a skin depth of 0.65 μm for copper wires. Half the thickness of the highest metal layer for a 65 nm CMOS technology [64] is less than the skin depth, justifying the assumption of a uniform current distribution. In addition, with advancements in technology, the thickness of the metal layers is decreasing, making the skin effect less significant.

6.2.2 Effective Inductance

The effective inductance of a single metal layer is

$$\frac{1}{L_{eff}} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} + ... + \frac{1}{L_N}, \tag{6.6}$$

where L_1, L_2, L_3, and L_N are, respectively, the effective inductance of the first, second, third, and N^{th} pair of an interdigitated P/G distribution network (refer to Fig. 6.2). Assuming the current flows in opposite directions in power and ground wires, the effective inductance of every pair can be determined [92].

In Fig. 6.3, the effective inductance normalized to the lowest effective inductance is depicted for each pair in a 100-pair interdigitated P/G distribution network. The difference in inductance is small among all of the pairs in an interdigitated power/ground network, excluding those pairs closest to the boundary. The effect of the boundary is

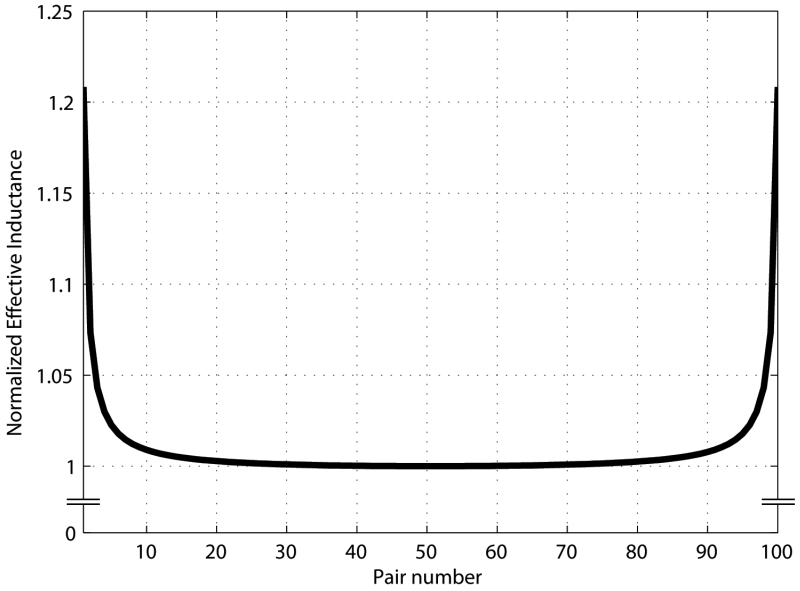


Figure 6.3: Normalized effective inductance for each pair in a 100 pair interdigitated P/G distribution network.

neglected, assuming all of the individual inductances are equal, permitting the effective inductance of a single layer within an interdigitated P/G distribution network to be determined (see Appendix A for derivation). A similar assumption is considered in [91], neglecting the mutual inductance between distant pairs, effectively treating the inductance as a local phenomena. By not neglecting distant mutual effects, the effective inductance can be estimated with higher precision. A derivation of the proposed effective inductance expression is presented in Appendix A, based on the Wallis

formula [131], resulting in

$$L_{eff} = \frac{1}{N}\frac{\mu_0 l}{\pi}\left[ln\left(\frac{w+s}{w+t}\right) + \frac{3}{2} + ln\left(\frac{2}{\pi}\right)\right], \tag{6.7}$$

where N and μ_0 are the number of power and ground pairs and the permeability of a vacuum, respectively. The upper boundary of the proposed inductance model is determined in Appendix B, resulting in

$$ErrorBound_{N\geq 1} = \frac{ln\left(\frac{\pi}{2}\right)}{ln\left(\frac{d}{w+t}\right) + \frac{3}{2}}. \tag{6.8}$$

The *proposed* model, represented by (6.7), is compared among FastHenry [93], a multipole 3-D inductance extraction program, the *Grover* [92], and *Mezhiba* [91] models in Fig. 6.4. The *Grover* model describes the inductance of each pair which is calculated individually and includes every mutual inductance component [92]. While the individual inductance of each pair is determined, the effective inductance of a single layer within an interdigitated P/G network structure is estimated assuming the individual inductive lines are in parallel. Hence, the *Grover* model includes every mutual term among all of the wires in a system. The effective inductance utilizes an approximation from [91], where the inductance is treated as a local effect, and the mutual inductance between other pairs is neglected. This model

is called the *Mezhiba* model [91]. The *proposed* model, represented by (6.7), determines the effective inductance assuming the number of P/G pairs is infinite. Since the magnitude of the mutual terms quickly declines to zero as a function of distance, this assumption is highly accurate in P/G networks with a large number of interdigitated power and ground pairs.

The error of the *Grover* model is lowest; however, the computational complexity exponentially increases with the number of wires in a P/G distribution network. The error and complexity, presented in Figs. 6.5 and 6.6, respectively, are evaluated among the *Grover*, *Mezhiba*, and *proposed* models. The complexity of the *proposed* and *Mezhiba* models is independent of the number of P/G pairs, while the *proposed* model converges to the values extracted by FastHenry, providing enhanced accuracy as compared to the *Mezhiba* model.

Substituting (6.3) into (6.7), the effective inductance of an interdigitated network is

$$L_{eff} = \frac{2l(w+s)}{A}\frac{\mu_0 l}{\pi}\left[ln\left(\frac{w+s}{w+t}\right) + \frac{3}{2} + ln\left(\frac{2}{\pi}\right)\right]. \tag{6.9}$$

6.3 Optimal Width for Minimum Impedance

The effective resistance (6.5) and inductance (6.9) as a function of wire width are compared with FastHenry [93] and illustrated in Fig. 6.7. A 1 x 1 mm^2 area of the top metal layer for a 65 nm CMOS [64] technology is chosen. The target frequency is

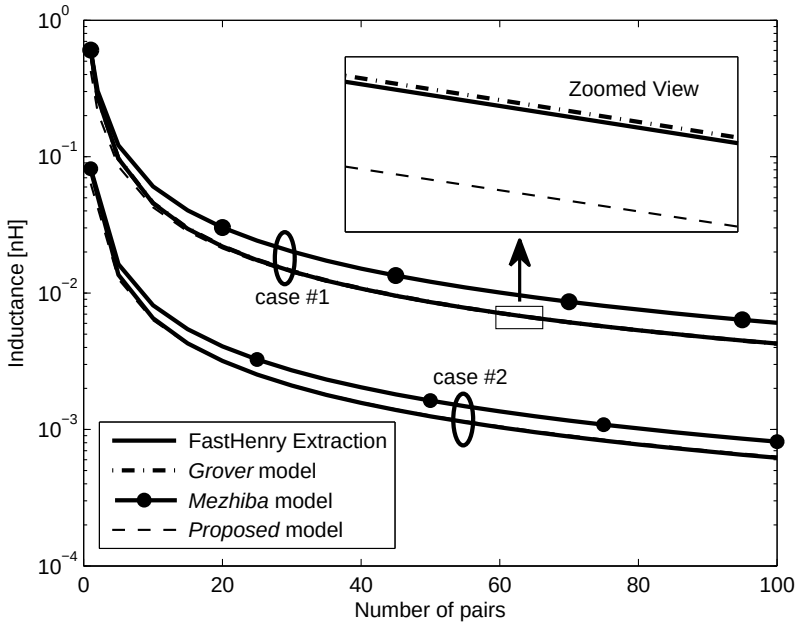


Figure 6.4: Comparison of FastHenry, *Grover*, *Mezhiba*, and *proposed* models for two different design cases.

5 GHz. For constant area, according to (6.5), wider power and ground wires reduce the effective resistance. With multiple thin lines, a large area is consumed by the line-to-line spacing, increasing the effective resistance of the network. The inductance under a constant area constraint has the opposite effect since the mutual component of the inductance is dominant in an interdigitated P/G distribution structure. Additional lines increase the mutual inductance, reducing the effective inductance according to (6.9). Since the effective resistive and inductive impedance in interdigitated structures behaves inversely with increasing wire width, the objective is to minimize the overall impedance of the network at a target frequency.

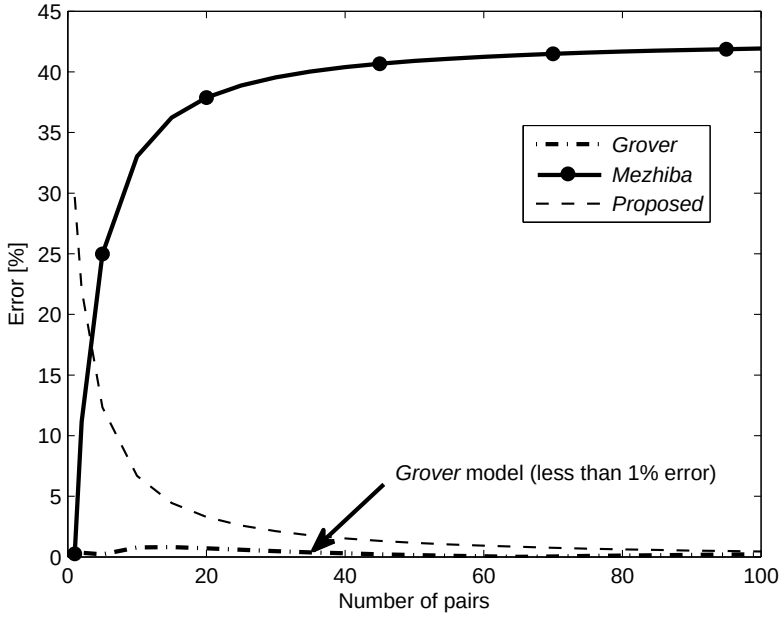


Figure 6.5: Error comparison for the *Grover*, *Mezhiba*, and *proposed* models. All of the models are compared to FastHenry.

In the network, the line-to-line spacing is based on the interconnect characteristics. The network impedance from (6.2) is therefore

$$|Z_{eff}(w)| = \sqrt{R_{eff}^2(w) + 4\pi^2 f^2 L_{eff}^2(w)}, \tag{6.10}$$

a function of line width. In Fig. 6.8, (6.10) is compared to FastHenry for several line widths.

Since the effective inductance in (6.9) is a transcendental function of width, a closed-form analytic solution can not be determined for the wire width that minimizes

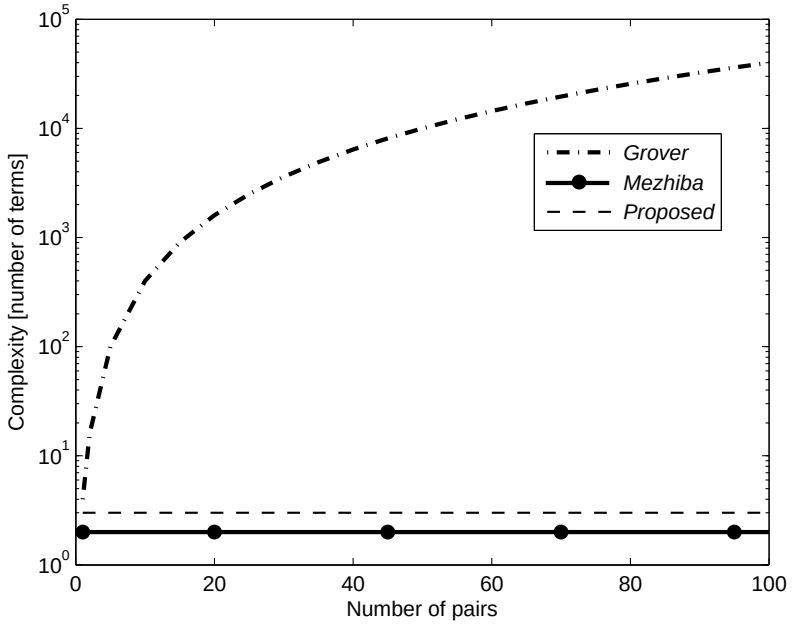


Figure 6.6: Comparison of computational complexity among the *Grover*, *Mezhiba*, and *proposed* models.

the impedance. A closed-form expression can, however, be determined for the special case where the line-to-line spacing is equal to the thickness of the metal, resulting in

$$w_{opt}^{(0)} \approx \sqrt[3]{0.91 \frac{s\rho^2}{\mu_o^2 t^2 f^2}}. \tag{6.11}$$

A detailed derivation of (6.11) is presented in Appendix C. A numerical solution based on n iterations of the Newton–Raphson method is used to determine the optimal width for all other spacings,

$$w_{opt}^{(n)} = w_{opt}^{(n-1)} - \frac{F\prime(w_{opt}^{(n-1)})}{F\prime\prime(w_{opt}^{(n-1)})}, \tag{6.12}$$

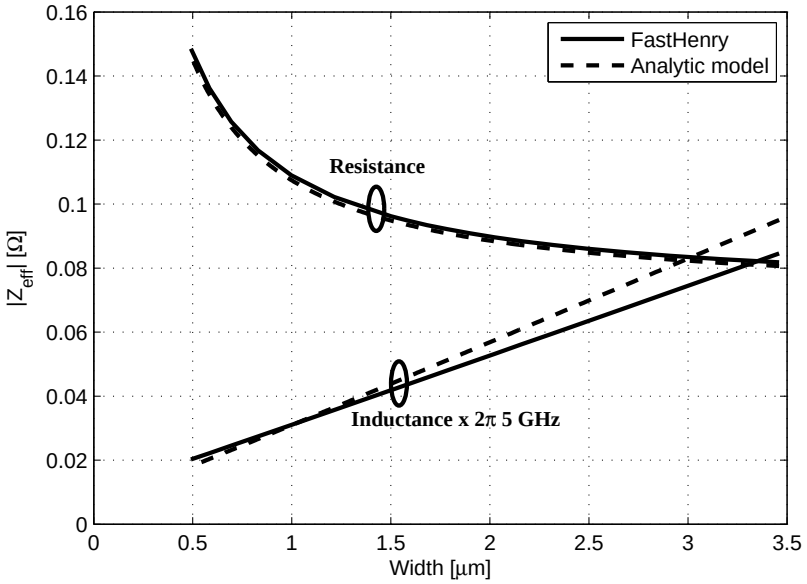


Figure 6.7: Effective resistance and inductance at 5 GHz as a function of wire width for a single layer within an interdigitated P/G distribution network. The overall area is maintained constant.

where $F \equiv |Z_{eff}(w)|$ and $w_{opt}^{(n-1)}$ is the $(n-1)^{st}$ estimate of the optimal wire width. The initial estimate is based on (6.11). The number of iterations can typically be increased to enhance the accuracy of the optimal width.

Considering resistance and inductance (both self and mutual) of a network, (6.11) combined with (6.12) determines the optimal line width of an interdigitated power and ground network. The optimal line width produces the minimum impedance network at a target frequency.

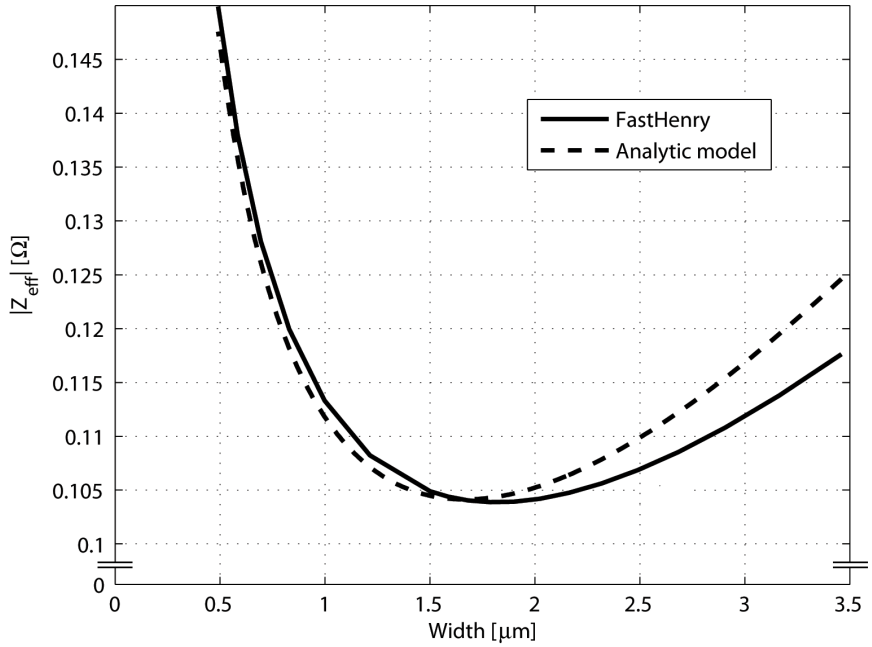


Figure 6.8: Magnitude of the impedance of (6.10) and FastHenry.

6.4 Optimal Width Characteristics

The optimal width for different thicknesses and spacings is evaluated in this section. The required number of iterations of the Newton–Raphson method is determined. The range of frequency for the optimum width is also reviewed. Finally, the effect of the decoupling capacitance on the optimal width is discussed.

6.4.1 Effect of Thickness and Spacing

A comparison among FastHenry, (6.11), and $w_{opt}^{(1)}$ based on the first iteration of the Newton–Raphson method is shown in Fig. 6.9 for several different metal thicknesses.

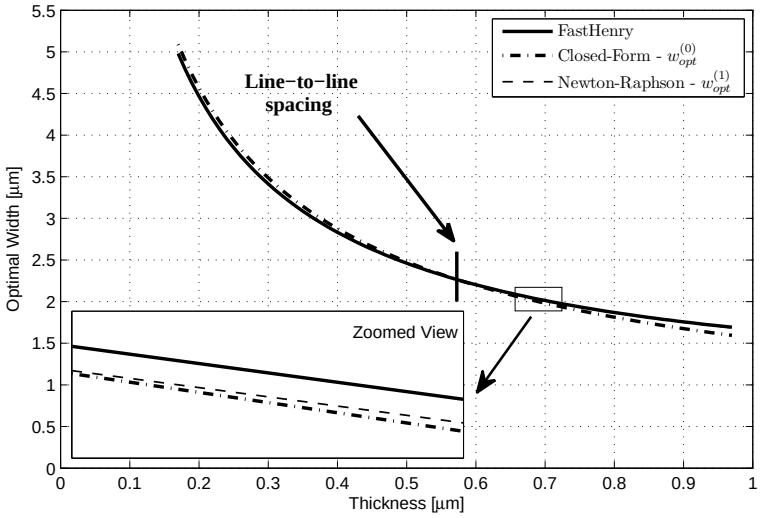


Figure 6.9: Closed-form $w_{opt}^{(0)}$ and $w_{opt}^{(1)}$ based on the first iteration of the Newton-Raphson method as compared with FastHenry for different thicknesses.

The spacing is chosen as the mid point between the thinnest and thickest metal layers for a 65 nm CMOS technology. A 5 GHz frequency is assumed. The error between FastHenry and (6.11) reaches 6%, while the error between FastHenry and $w_{opt}^{(1)}$ is below 1%. For those cases where the target accuracy is below the error of the initial estimate, the closed-form expression of $w_{opt}^{(0)}$ is computationally efficient in determining the P/G line width. If higher accuracy is required, the interdigitated P/G wire width can be determined according to (6.12).

A comparison among FastHenry, (6.11), and $w_{opt}^{(1)}$ based on the first iteration of the Newton–Raphson method is shown in Fig. 6.10 for different spacings. The spacing ranges from 0.54 μm (the lowest permitted spacing) to 15 μm. When the spacing

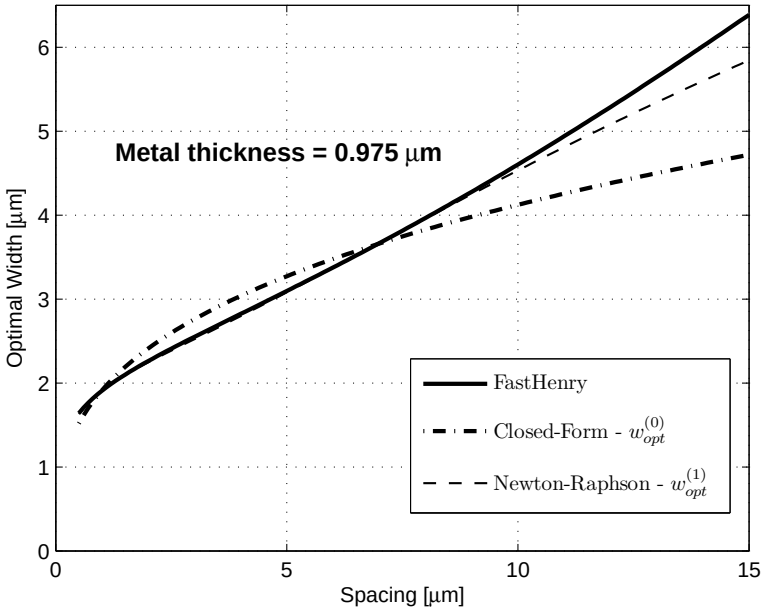


Figure 6.10: Closed-form $w_{opt}^{(0)}$ and $w_{opt}^{(1)}$ based on the first iteration of the Newton-Raphson method as compared with FastHenry for different spacings.

is equal to the thickness, $w_{opt}^{(0)}$ and $w_{opt}^{(1)}$ are equal since the logarithmic term of (6.9) is zero. At spacings below 7 μm, (6.11) exhibits an error below 9% as compared to FastHenry. For spacings up to 15 μm, the error between FastHenry and (6.11) reaches 26%, while the error with only one iteration of the Newton–Raphson method is below 9%. For spacings greater than 15 μm, additional iterations of the Newton–Raphson method are necessary.

The different number of iterations to determine w_{opt} is evaluated in Fig. 6.11. The error is relative to FastHenry. Note that (6.11) assumes the spacing and thickness are equal. The closed-form expression is therefore only accurate for small spacings.

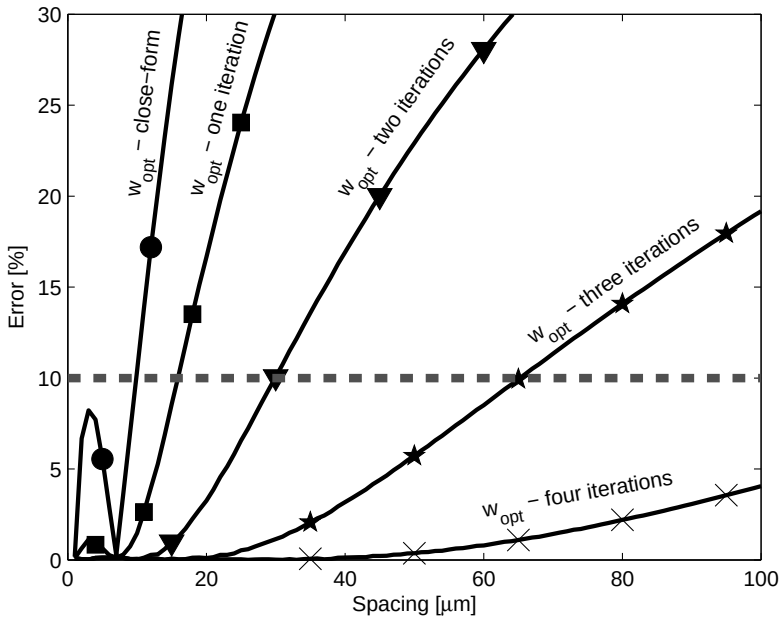


Figure 6.11: Error of w_{opt} is evaluated for several spacings using closed-form and one to four iterations of the Newton–Raphson method. The error is relative to FastHenry.

For wider spacings, the Newton–Raphson method is preferred to accurately determine the optimal width. A larger number of iterations is needed for wider spacings since the error decreases significantly for large number of iterations. Spacings up to 100 μm are evaluated, suggesting that four iterations are sufficient to determine the optimal width within 10% accuracy as compared to FastHenry.

6.4.2 Frequency Range

Since the width is optimized for a target frequency, the effect on the frequency range of interest (from DC to the target frequency) is important. In the following

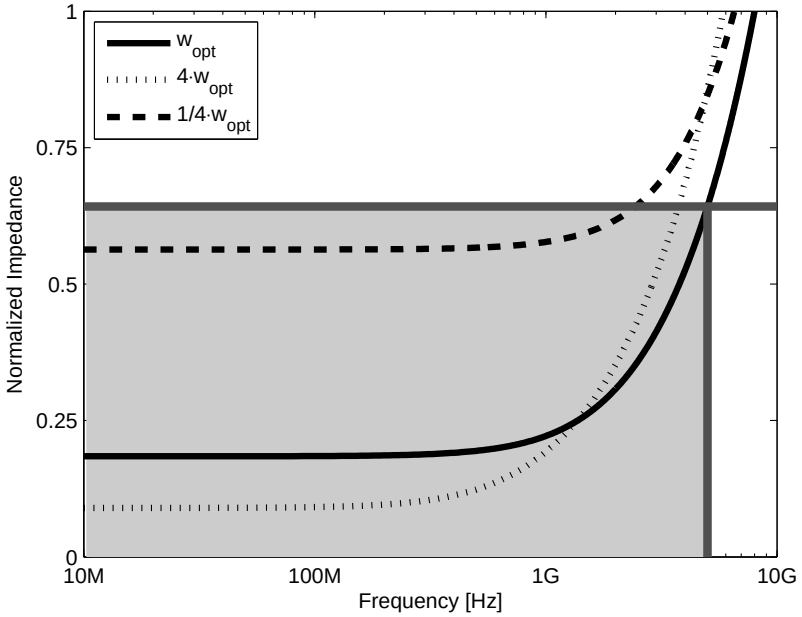


Figure 6.12: Impedance over the frequency range of interest. Three different P/G network line widths are depicted. The impedance is minimum at the target frequency with the optimum width.

discussion, a target frequency of 5 GHz is assumed. In Fig. 6.12, the impedance of the network as a function of frequency is depicted. Three values of the width are evaluated - the optimal width, a width four times greater than the optimal width, and a width four times smaller than the optimal width. Note that the area is maintained constant. An increasing width corresponds to a fewer number of interdigitated pairs within the P/G network (a thinner line corresponds to a higher number of interdigitated pairs).

As illustrated in Fig. 6.12, the minimum impedance at 5 GHz is achieved using the optimal width, while a lower and higher line width increases, respectively, the

resistive and inductive component of the overall impedance. At low frequencies, the P/G network with wider lines produces a lower impedance, although more than the required impedance at the target frequency. In the example shown in Fig. 6.12, the network impedance with wider lines is below the target impedance only below 3.7 GHz. As depicted in Fig. 6.12, the P/G network with a smaller width satisfies the impedance requirements only up to 2.5 GHz.

6.4.3 Capacitive Component

The three capacitive components within a P/G distribution network are intentionally placed decoupling capacitance, non-switching circuitry behaving as a capacitance between the power and ground wires, and the line capacitance between the power and ground wires. Since the non-switching circuitry and decoupling capacitance are significantly larger than the line capacitance, the line capacitance can be neglected. The capacitance is therefore not significantly affected by the change in the line width of the P/G network.

Note that the overall impedance is different when considering the decoupling capacitance. The optimal width that minimizes the impedance of a network is similarly determined as (6.12), since the decoupling capacitance is independent of line width.

6.5 Summary

The characteristics of a single metal layer of an interdigitated power and ground network are investigated in this chapter. A closed-from expression characterizing the inductance of a single layer within a P/G distribution network is presented. The solution is compared to previous work and FastHenry, exhibiting good accuracy. The error for the *proposed* model is highest for a few P/G pairs; however, due to the small number of lines, the *Grover* model can be used in these cases. With an increasing number of P/G network pairs, the error of the *proposed* model decreases rapidly, permitting the effective inductance of a P/G distribution network to be accurately and efficiently estimated. The magnitude of the effective inductance is bounded by the values determined from the *proposed* and *Mezhiba* models. The bound dramatically decreases with increasing number of pairs. The effective inductance of a single layer within an interdigitated P/G distribution network structure can therefore be accurately determined for any number of power and ground line pairs.

The effect of resistive and inductive impedances under a constant area constraint is observed for different wire widths (or number of interdigitated pairs). Under a constant area constraint, conflicting behavior of the effective resistive and inductive impedances is observed with increasing wire width. This conflicting behavior is explored to achieve the minimum impedance of a single interdigitated metal layer. A closed-form expression for the line width is determined, achieving high accuracy for

small spacings as compared to FastHenry. The accuracy of the optimal width is enhanced over a wide range of spacings utilizing iterative approaches (up to four iterations). The effect of different spacings, thicknesses, and frequencies on the optimal width is reviewed. A design methodology that considers the network resistance and inductance (both self and mutual) of a single interdigitated metal layer network that produces the minimum impedance is presented in this chapter.

Chapter 7

Interdigitated Power/Ground Network - Multiple Metal Layers

An interdigitated P/G distribution structure is typically located on several metal layers. Each layer consists of interdigitated power and ground wires, where the direction of the wires is perpendicular to the direction of the wires in the previous layer, as depicted in Fig. 7.1. Routing flexibility and reduced inductance are two primary advantages of an interdigitated P/G distribution structure. With advancements in technology, additional metal layers are provided [7], permitting the dedication of several metal layers to the P/G network. Due to electromigration, the maximum current is limited; therefore, a larger number of metal layers passes higher current to the system under the same electromigration constraint.

The approach for designing a multi-layer power and ground network is different that a single metal layer, since the interaction among the layers should be considered. The optimal width of the top metal layer in a multi-layer system is determined based

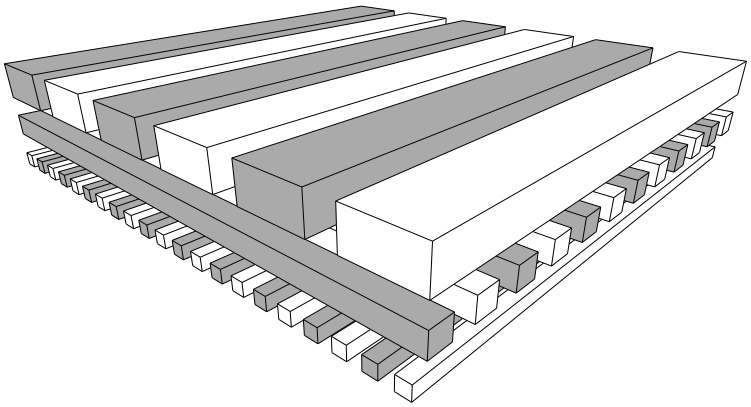


Figure 7.1: Global interdigitated P/G distribution structure. The darker and lighter lines represent, respectively, the power and ground lines.

on (6.12), as described in Chapter 6. The current density also needs to be determined for a multi-layer system.

This chapter is organized as follows. The electromigration and current density is introduced in Section 7.1. In Section 7.2, a method to lower the current density across multiple metal layers is discussed. A method to minimize the network impedance across multiple metal layers is provided in Section 7.3. The tradeoff between the impedance of a P/G network and the current density is presented in Section 7.4. The chapter is summarized in Section 7.5.

7.1 Current Density

Multi-layer systems can be approximated by the network shown in Fig. 7.2, where the resistance and inductance is, respectively, the effective resistance and inductance

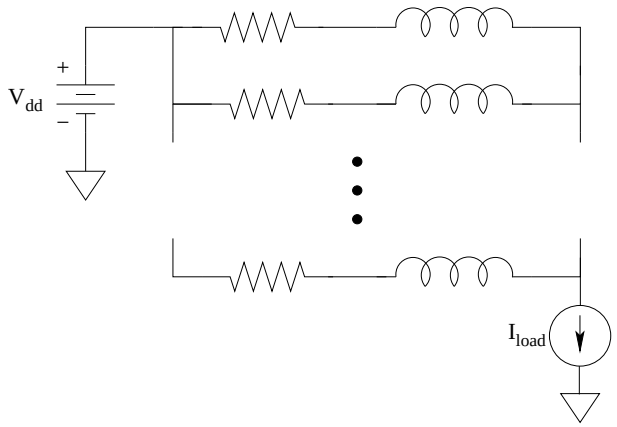


Figure 7.2: Multi-layer P/G distribution network model. Each resistance and inductance represent, respectively, the effective resistance and inductance of a single layer within a P/G network.

of a single layer within a P/G distribution network [91]. This model treats the system as worse case since all of the current is assumed to flow through the entire layer. Electromigration should also be considered when optimizing a multi-layer system.

The current density CD of an arbitrary layer m is

$$CD_m = \frac{|i_m|}{CroSec_m}, \tag{7.1}$$

where i_m and $CroSec_m$ are the current and cross section of layer m, respectively. The skin effect is considered in determining the cross-section of the layer,

$$CroSec_m = \begin{cases} N_m w_m t_m & 2\delta > w \quad (7.2) \\ N_m w_m t_m & 2\delta > t \quad (7.3) \\ 2\delta \left[N_m (w_m + t_m) - 2\delta \right] & \text{otherwise,} \quad (7.4) \end{cases}$$

where δ is the skin depth. The skin depth is defined as

$$\delta \equiv \sqrt{\frac{1}{\pi f \mu_0 \sigma}}, \quad (7.5)$$

where σ is the conductivity of the material.

Allocating additional metal layers for the power and ground distribution network distributes the overall current among a larger number of metal layers. The current density of a particular metal layer is therefore lower.

Two different approaches are considered for optimizing a multi-layer P/G network. In the first approach, the current density per layer is maintained equal for all of the layers, while providing a low P/G network impedance. The second approach minimizes the impedance, while considering electromigration. A tradeoff exists between the current density and the impedance of a P/G distribution network. A lower impedance reduces the voltage drop, providing a higher noise margin.

7.2 First Approach - Equal Current Density

The first optimization approach for an interdigitated P/G distribution network structure is discussed in this subsection. The limiting current density is the highest current density among the layers. In this approach, the current density among the layers is maintained equal, minimizing the limiting current density of a P/G network. A lower limiting current density enhances the reliability of the multi-layer system. The current flowing through an arbitrary layer m is

$$|i_m| = \frac{V_{drop}}{|Z_m|}, \tag{7.6}$$

where V_{drop} and Z_m are the voltage across the entire P/G distribution network and the impedance of the m^{th} layer, respectively. Substituting (7.6) into (7.1), the current density of the m^{th} metal layer is

$$CD_m = \frac{V_{drop}}{|Z_m|} \frac{1}{CroSec_m}. \tag{7.7}$$

Two layers, m and n, provide the same current density when

$$|Z_m|\, CroSec_m = |Z_n|\, CroSec_n. \tag{7.8}$$

While the width of a single metal layer is optimized for minimum impedance, the width of the other metal layers is determined to maintain equal current density, as described by (7.8). Pseudo-code for determining the width of the individual metal layers within a multi-layer system based on maintaining equal current density among the metal layers is provided in Appendix D.

Based on the proposed methodology, an eight layer P/G distribution network is described for a 65 nm CMOS technology. To evaluate the proposed methodology, all of the metal layers are available for the P/G distribution, although in practical cases some metal layers are used for the signals, clock network, and shield lines. In Fig. 7.3, the physical parameters for a single metal layer of an interdigitated P/G distribution network are illustrated. The width of each metal layer is determined by the proposed algorithm, as described in Appendix D.

Based on a 65 nm CMOS technology, a width of 1.66 μm is initially determined from (6.12) for the top (eighth) metal layer to minimize the impedance of a single metal layer. The width of the additional metal layers is based on maintaining equal current density according to (7.8). The current density per multiple metal layers is depicted in Fig. 7.4. Increasing the width of the lower metal layer affects the current density in the lower layer as well as the upper metal layer. Increasing the width of the lower metal layer changes the impedance of the lower metal layer (decreasing the resistance and increasing the inductance). The current is distributed among the

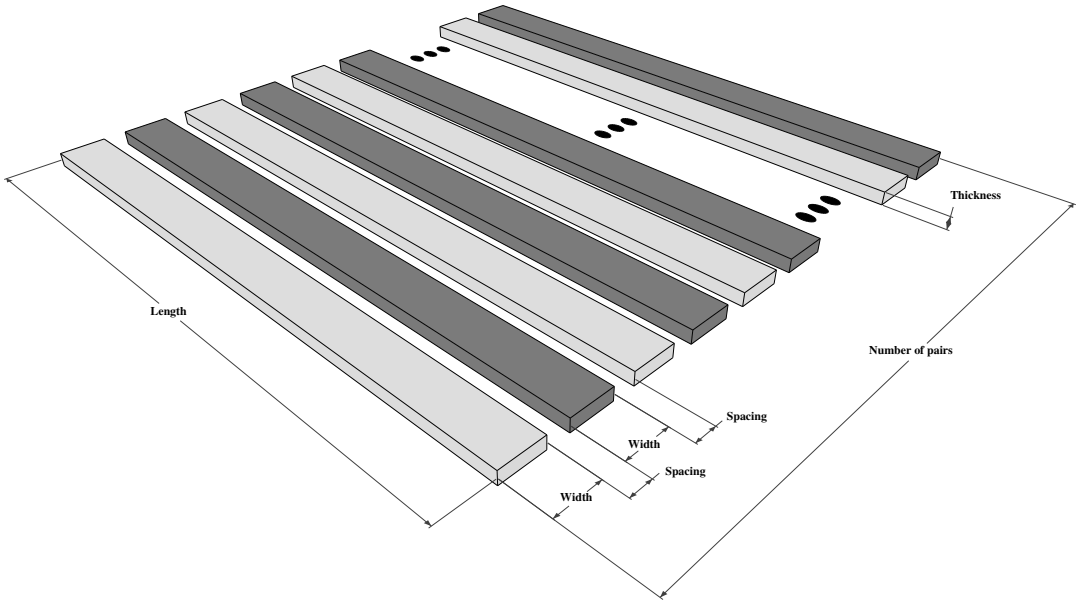


Figure 7.3: Physical characteristics of a single metal layer. The same notation is applied for each metal layer in a multi-layer P/G distribution network.

different metal layers based on the relative impedance of the metal layers, resulting in larger current in the metal layer with lower impedance while changing the current density in all of the metal layers.

The width is determined at the intersection of the current density of multiple metal layers. The intersection is the width where equal current density among the multiple metal layers is maintained, lowering the limiting current density. As inferred in Fig. 7.4, this intersection occurs at a greater width for the lower metal layers, since the metal layers are thinner. This structure is therefore called the *pyramid* structure. The spacing, thickness, width, and number of interdigitated pairs for each metal layer in an eight layer P/G system is summarized in Table 7.1.

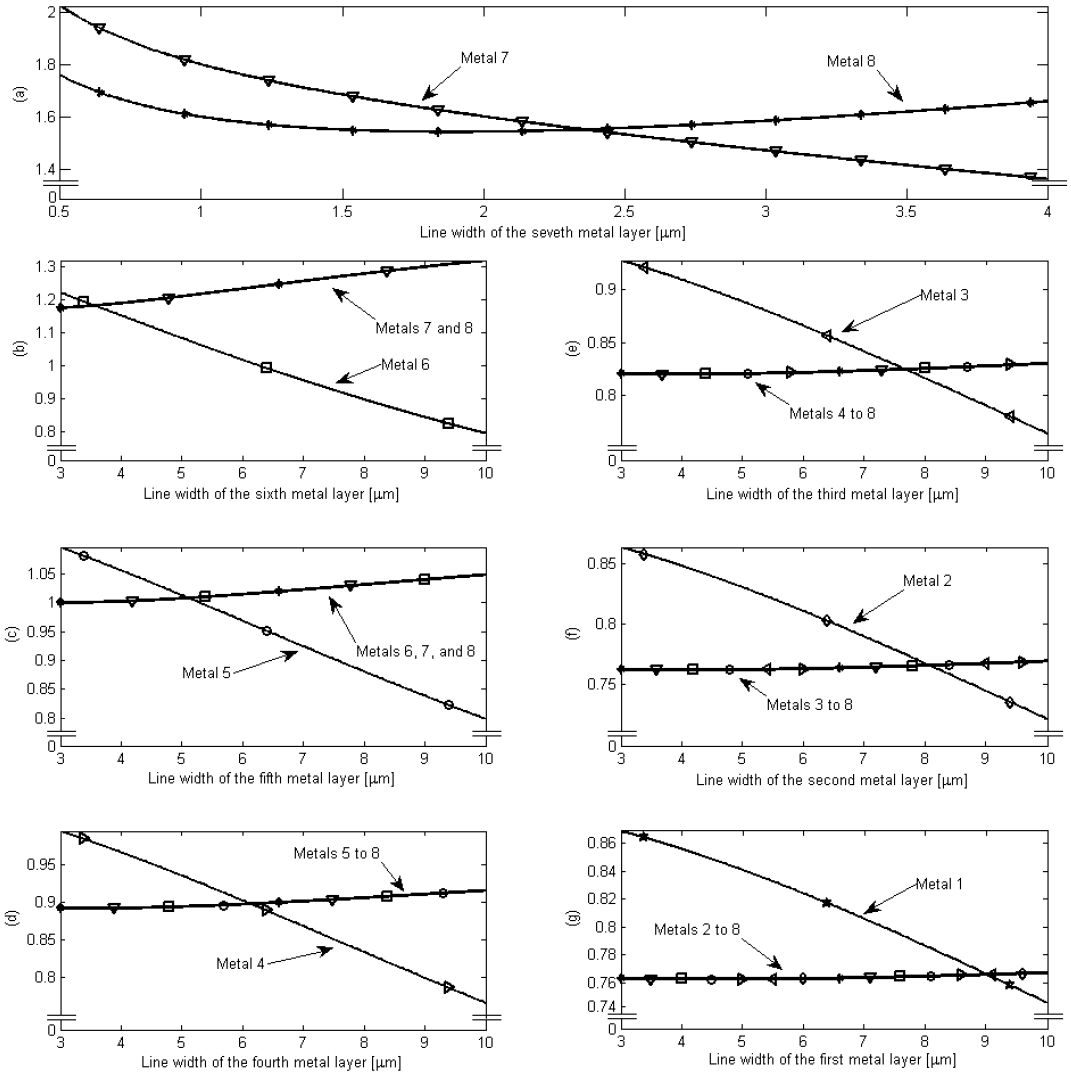


Figure 7.4: Current density for multiple metal layers. The current density of the (a) seventh and eighth, (b) sixth, seventh, and eighth, (c) fifth, sixth, seventh, and eighth, (d) fourth to eighth, (e) third to eighth, (f) second to eighth, and (g) first to eighth metal layers. The width is determined at the intersection of the current density of the multiple metal layers. The $y - axis$ for each figure is the current density in units of mA/μm^2, while the total current is assumed to be 1 A.

Table 7.1: Spacing, thickness, width, and number of interdigitated pairs for an eight metal layer system. The eighth metal layer is the top metal layer. Since the lines are wider, the number of interdigitated pairs is lower for a constant area.

Metal Layer	Thickness [μm]	Spacing [μm]	Width [μm]	Number of pairs
8	0.975	0.540	1.66	227
7	0.650	0.360	2.36	183
6	0.430	0.240	3.56	131
5	0.300	0.165	5.11	94
4	0.250	0.140	6.13	79
3	0.200	0.110	7.67	64
2	0.190	0.105	8.07	61
1	0.170	0.105	9.02	54

The normalized current density is shown in Fig. 7.5 for an eight metal layer P/G network based on the equal current density methodology. While the maximum current density for a specific technology, physical area, and current is known, the required number of metal layers for an interdigitated P/G distribution network can be determined, as illustrated in Fig. 7.5.

Two additional P/G network structures are compared with the proposed *pyramid* structure. These three structures are illustrated in Fig. 7.6. The width of the individual metal layers for the *pyramid* structure is listed in Table 7.1. This structure is shown in Fig. 7.6(a). Note in the *pyramid* structure, the power and ground lines in the lower metal layers are wider. In conventional metal systems, the power and ground lines are wider at the higher metal layers, as illustrated in Fig. 7.6(b). For this structure, the width of the metal layers is the opposite of the *pyramid* structure,

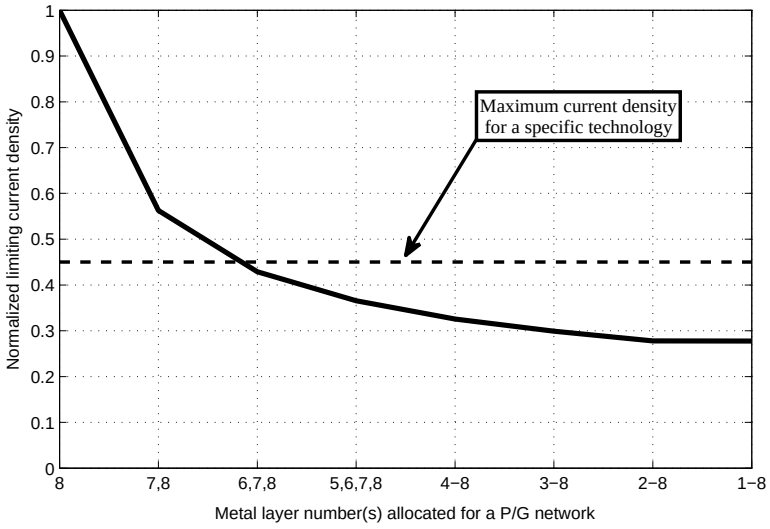
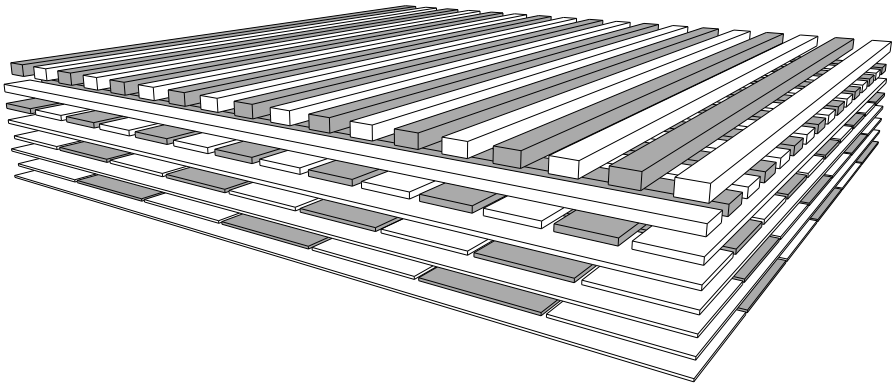


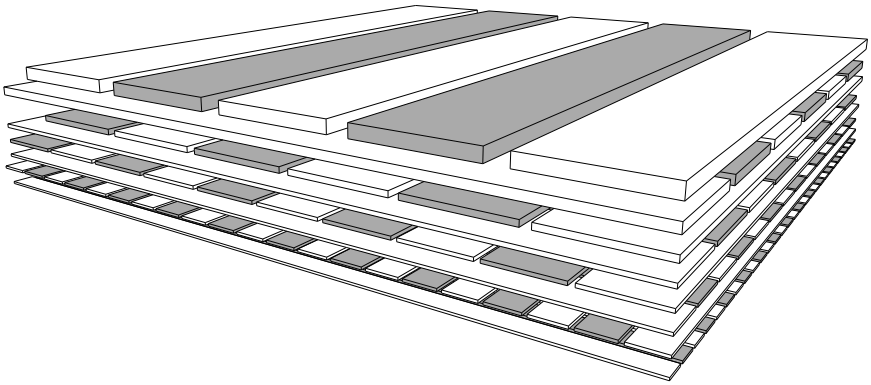
Figure 7.5: Normalized limiting current density for different metal layers. The $x-axis$ represents the metal layer number(s) allocated for a P/G network. The current density is highest when allocating only a single metal layer (layer number eight) for a P/G network. The current density is reduced as additional metal layers are added.

and is therefore called the *inverted pyramid* (standard) structure. In Fig. 7.6(c), the width of each metal layer is maintained constant at 5.5 μm; therefore, this structure is referred to as the *equal width* structure. The width, number of interdigitated pairs, effective impedance, and limiting current density for these three structures are listed in Table 7.2. For the current density evaluation, the metal layers are extracted individually using FastHenry.

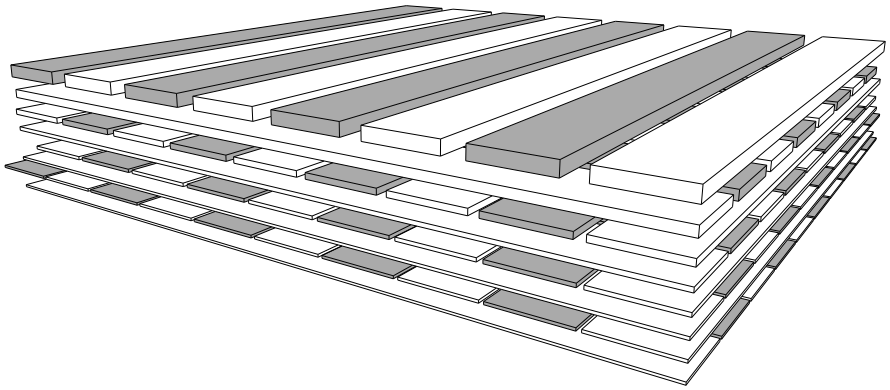
In the *pyramid* structure, the current density is maintained equal among the layers, lowering the limiting current density. Since the thickness decreases with lower metal layers, the lines are wider, maintaining a constant current density. In the



(a)



(b)



(c)

Figure 7.6: Three P/G structures; (a) *pyramid* (proposed) structure - the width decreases with higher metal layers, (b) *inverted pyramid* (standard) structure - the width increases with higher metal layers, and (c) *equal width* structure - the width is maintained equal among all of the metal layers.

Table 7.2: Three structures are compared for equal current density. The thickness, spacing, width, and number of interdigitated pairs per metal layer for each structure are listed.

Metal Layer	Thickness [μm]	Spacing [μm]	*Pyramid* structure		*Inverted pyramid* structure		*Equal width* structure	
			Width [μm]	Number of pairs	Width [μm]	Number of pairs	Width [μm]	Number of pairs
8	0.975	0.540	1.7	227	9.0	52	5.5	82
7	0.650	0.360	2.4	183	8.1	59	5.5	85
6	0.430	0.240	3.6	131	7.7	63	5.5	87
5	0.300	0.165	5.1	94	6.1	79	5.5	88
4	0.250	0.140	6.1	79	5.1	95	5.5	88
3	0.200	0.110	7.7	64	3.6	136	5.5	89
2	0.190	0.105	8.1	61	2.4	202	5.5	89
1	0.170	0.105	9.0	54	1.7	280	5.5	89
Effective Impedance [mΩ]			30.6		46.0		38.2	
Limiting Current Density [mA/μm^2]			0.766		1.400		1.044	

inverted pyramid structure, the higher metal layers are wider, permitting greater routing flexibility; the reliability of the metal, however, decreases since the limiting current density is 82% higher as compared to the *pyramid* structure. In the *inverted pyramid* structure, most of the current flows in the higher metal layers increasing the effective impedance of the overall system. The impedance is 50% higher than in the *pyramid* structure. The *equal width* structure exhibits a higher effective impedance

and current density of 24% and 36%, respectively, as compared to the *pyramid* structure. This trend is consistent with the change in importance of the inductance as compared to the resistance at higher frequencies.

7.3 Second Approach - Minimum Impedance

The focus of the second optimization approach is to minimize the impedance of the overall P/G distribution network. Assuming the metal layers are in parallel while optimizing each layer for minimum impedance, the lowest impedance of the overall system is achieved. The number of required metal layers is based on the current density constraint. Pseudo-code for this optimization algorithm is presented in Appendix E. The impedance of each of the eight metal layers is illustrated in Fig. 7.7.

Three different interdigitated P/G distribution structures, illustrated in Fig. 7.6, are compared. The structures are referred to by the same names as in the previous subsection, however, the widths are determined based on the minimum impedance algorithm rather than the equal current density algorithm. The width of the power and ground lines for the *pyramid* (proposed) structure is based on the algorithm presented in Appendix E. In the *inverted pyramid* (standard) structure, the width increases with higher metal layers. The width of the metal layers is the opposite of the *pyramid* structure. The width of all eight metal layers is maintained constant at

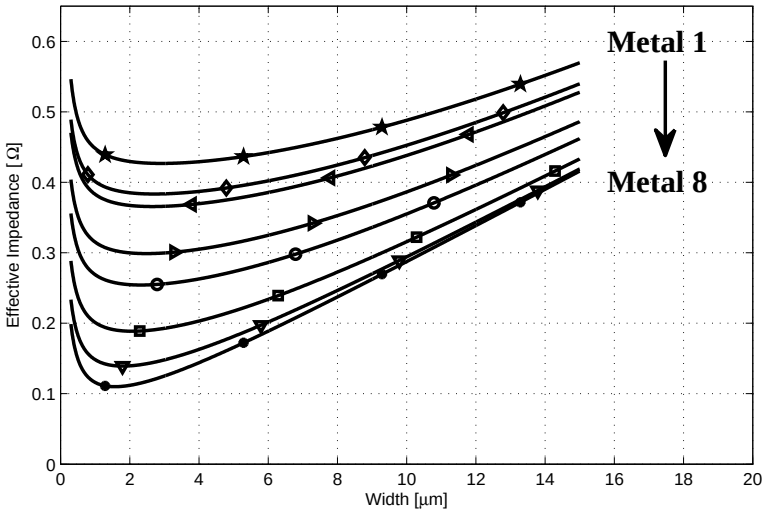


Figure 7.7: Effective impedance as a function of width for each of eight metal layers within an interdigitated P/G distribution network. The overall area of each metal layer is maintained constant.

2.4 μm for the *equal width* structure. In Table 7.3, the thickness, spacing, width, and number of interdigitated pairs are listed for each structure. The effective impedance and limiting current density for each structure are also summarized in Table 7.3.

The lowest effective impedance is achieved in the *pyramid* structure. The effective impedance is 6% and 3% higher for the *inverted pyramid* and *equal width* structures, respectively, as compared to the *pyramid* structure. The limiting current density in the *pyramid* structure is enhanced, respectively, by 8% and 4% as compared to the *inverted pyramid* and *equal width* structures. Hence, the effective impedance achieved by the proposed *pyramid* structure is lower. This improvement is due to the relative importance of the inductance as compared to the resistance in high frequency systems.

Table 7.3: Three structures are compared for minimum impedance. The thickness, spacing, width, and number of interdigitated pairs per metal layer for each structure are listed.

Metal Layer	Thickness [μm]	Spacing [μm]	*Pyramid* structure		*Inverted pyramid* structure		*Equal width* structure	
			Width [μm]	Number of pairs	Width [μm]	Number of pairs	Width [μm]	Number of pairs
8	0.975	0.540	1.7	227	2.9	144	2.4	170
7	0.650	0.360	1.9	225	2.7	162	2.4	181
6	0.430	0.240	2.1	212	2.7	172	2.4	189
5	0.300	0.165	2.3	199	2.5	187	2.4	194
4	0.250	0.140	2.5	189	2.3	201	2.4	196
3	0.200	0.110	2.7	180	2.1	225	2.4	199
2	0.190	0.105	2.7	177	1.9	254	2.4	199
1	0.170	0.105	2.9	165	1.7	283	2.4	199
Effective Impedance [mΩ]			29.5		31.5		30.5	
Limiting Current Density [mA/μm^2]			0.843		0.909		0.875	

7.4 Discussion

The following section is divided into four subsections: a comparison between the two aforementioned design approaches, a discussion of routability and the grid area ratio, an estimate of the optimal power/ground line width for different frequencies and technologies, and an investigation of the critical frequency for the design of multi-layer power/ground networks.

Table 7.4: Comparison between two optimization approaches for a one, two, three, and eight metal layer system. CD is the current density.

Number of metal layers	First Approach		Second Approach	
	Z_{eff} $[m\Omega]$	Limiting CD [mA/μm^2]	Z_{eff} $[m\Omega]$	Limiting CD [mA/μm^2]
1	105.1	2.71	105.1	2.71
2	59.5	1.54	59.4	1.60
3	45.6	1.18	45.2	1.25
8	30.6	0.77	29.5	0.84

7.4.1 Comparison

Evaluating both approaches, a tradeoff is observed between the impedance (or voltage drop) and the limiting current density of a P/G distribution network. When focusing only on the current density, the optimal solution suggests the P/G network should be as wide as possible; however, at high frequencies, the effective impedance increases significantly due to the higher inductance. Both approaches consider the effective impedance and current density, while the current density is the focus of the first approach, and the effective impedance is the focus of the second approach. A comparison between both approaches for a one, two, three, and eight metal layer system is listed in Table 7.4.

As observed from Table 7.4, the effective impedance is lowest using the second approach, while the limiting current density is lowest if the first approach is used. A tradeoff between the limiting current density and effective impedance is noted. A difference of less than 10% between the two approaches for the impedance and current

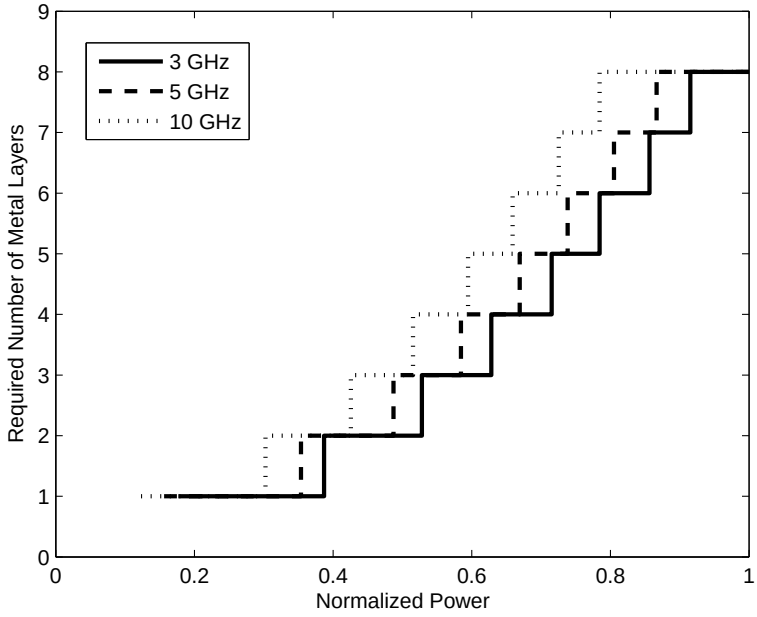


Figure 7.8: Required number of metal layers for a P/G network as a function of normalized power evaluated at three different frequencies.

density is demonstrated. However, when additional constraints (such as routability) are considered, the optimal width may not be available for that particular layer of metal within the power/ground distribution network. In this situation, the optimization process is focused on minimizing the impedance or current density, resulting in a greater difference between the two approaches. These two approaches are presented here to satisfy both optimization flows.

7.4.2 Routability

To develop the proposed methodology, these examples assume all of the metal layers and physical area can be used for the power/ground network. For a practical on-chip power and ground distribution network, routability, cost, and other issues should also be considered. Routability is an important issue primarily affecting the lower metal layers. Global power/ground networks tend to utilize the higher metal layers. To consider routability, a metric, the grid area ratio, is defined as the ratio of the metal resources occupied by the power/ground network to the total metal area [72, 129],

$$A \equiv \frac{w + s_0}{p}, \tag{7.9}$$

where w, s_o, and p are, respectively, the line width, minimal spacing between power and ground lines, and the line pitch. The line pitch is the width and spacing between the power and ground lines. If the spacing between the power and ground lines is minimum, the grid area ratio is one. The grid area ratio is depicted in Fig. 7.9 for different spacings. As anticipated, increasing the distance between the lines reduces the grid area ratio. As illustrated in Fig. 7.9, the grid area ratio is higher for the lower metal layers, resulting in reduced routability for the lower metal layers as compared to the higher metal layers where routability is better.

In Fig. 7.10, several P/G networks with different line widths between the top and

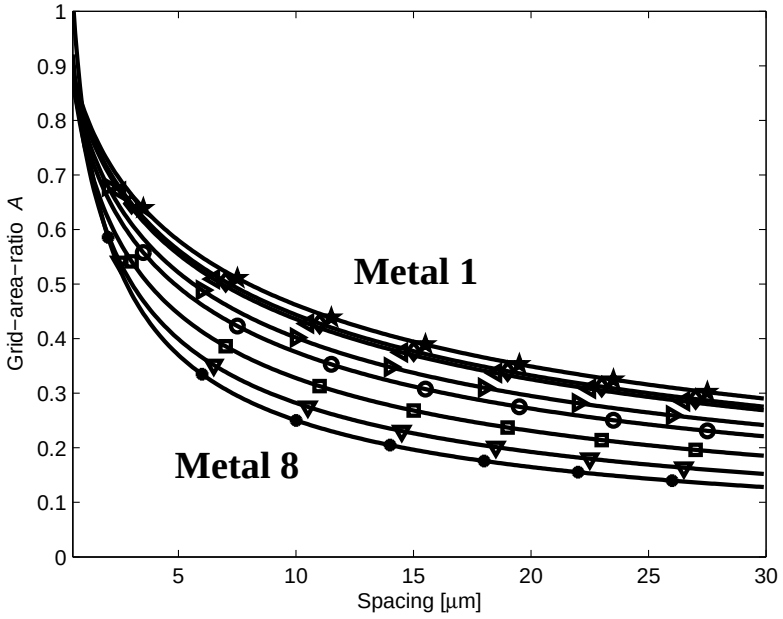


Figure 7.9: Grid area ratio as a function of spacing between the lines for different metal layers. The line width is based on (6.12) to minimize the network impedance.

bottom metal layers are evaluated. A 5 GHz target frequency and 10 μm line-to-line spacing between the power and ground lines is chosen. Four interdigitated metal layers are allocated for the power network. In Fig. 7.10, the x-axis is w_{top}/w_{bottom}, permitting a comparison between the impedance and grid-area-ratio for several *pyramid*, *equal width*, and *inverted pyramid* structures. The vertical line at $w_{top}/w_{bottom} = 1$ represent the *equal width* structure. The region to the left of the *equal width* structure represents *pyramid* structures with increasing width at the bottom metal layers and decreasing width at the top layers. The region to the right represents *inverted pyramid* structures with decreasing width at the bottom metal layers and increasing

width at the top layers. The lowest impedance among these structures is the *pyramid* structure with a line width based on (6.12). The grid-area-ratio however is lower in the *inverted pyramid* structure, indicating a tradeoff between the impedance and routability. The primary disadvantage of the proposed *pyramid* structure is therefore a higher grid-area-ratio (lower routability) as compared to the conventional *inverted pyramid* structure. A power network to the right of the minimum impedance *pyramid* structure may therefore be a reasonable compromise to provide effective routability with a reasonable loss in network impedance.

7.4.3 Fidelity

The required number of metal layers for the specified power levels is depicted in Fig. 7.8. The technology parameters are chosen based on a 65 nm CMOS technology with an area of 1 mm x 1 mm. The results are evaluated at three different frequencies, indicating that an additional metal layer is required at higher frequencies.

The optimal width as a function of the number of metal layers at 3 GHz and 10 GHz is illustrated in Fig. 7.11 for a 65 nm, 45 nm [132], and 32 nm [133] CMOS technology. The optimal width is determined based on the Newton-Raphson method as described by (6.12). At higher frequencies, the optimal width is thinner since the inductive impedance is greater. The optimal width increases with thinner, less inductive metal layers to satisfy the minimum impedance constraint. With technology

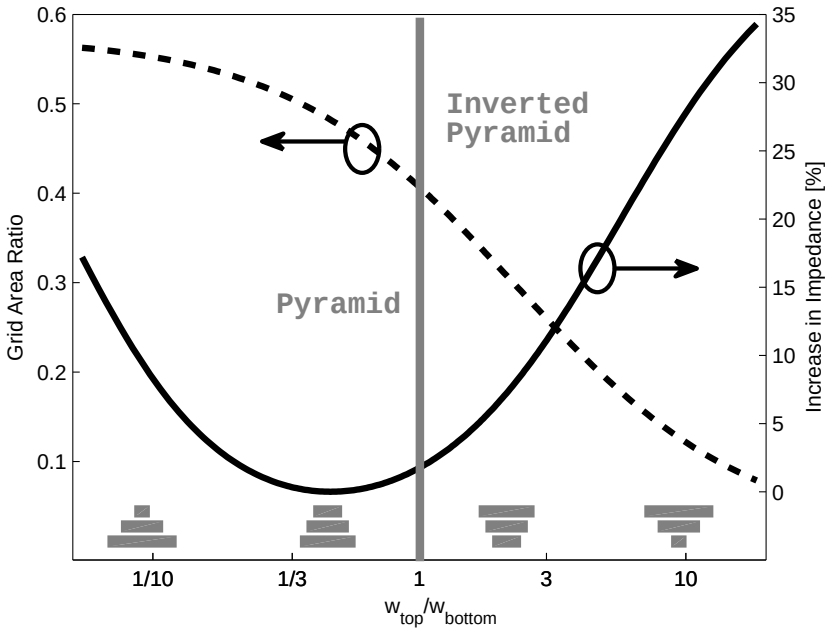


Figure 7.10: Grid-area-ratio and increase in impedance for several interdigitated P/G structures. Four metal layers are allocated for the power network. The vertical line represents the *equal width* structure. The left region is for *pyramid* structures, while the right region is for *inverted pyramid* structures. The minimum impedance is achieved by the *pyramid* structure, where the line width is based on (6.12).

scaling, metal thicknesses typically decrease, requiring wider lines to compensate for the increase in resistivity.

The effect of frequency on the design of an interdigitated power/ground distribution network is significant. At lower frequencies, where the resistive impedance is dominant, wide wires are used to reduce the impedance, while maintaining a constant cross-section to satisfy equal current density. At higher frequencies, the inductive impedance is dominant, suggesting that the power/ground lines should be less wide.

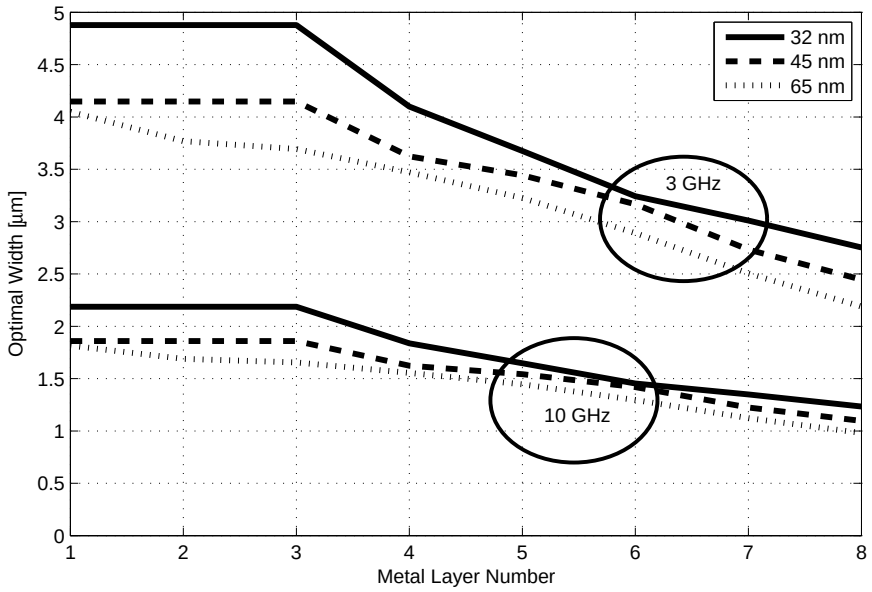


Figure 7.11: Optimal width to minimize the effective impedance of each metal layer based on a 65 nm, 45 nm, and 32 nm CMOS technology for two different frequencies.

7.4.4 Critical Frequency

The relationship between the two highest metal layers is investigated, permitting the concept of a critical frequency to be defined. The critical frequency is defined here as the frequency at which the impedance of two metal layers is equal. Assuming the width of both metal layers is the same, the critical frequency is determined for a variety of spacings, as depicted in Fig. 7.12. The arrows indicate the direction of decreasing grid area ratio A (increasing routability). The critical frequency is evaluated for three different line widths.

The critical frequency for different grid area ratios is illustrated in Fig. 7.13. A

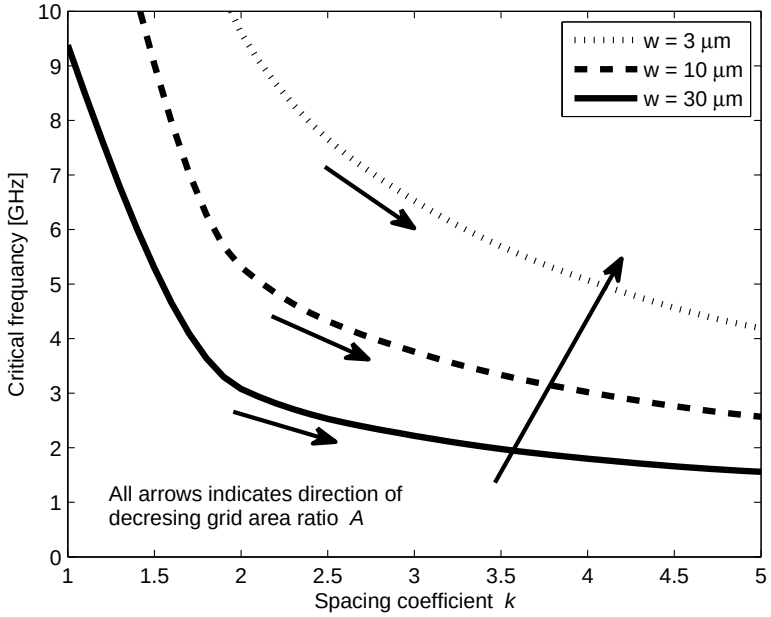


Figure 7.12: Critical frequency at which the impedance of the higher metal layer is equal to the impedance of the lower metal layer. The width of the power/ground lines is maintained equal for both metal layers. The distance between the lines is the minimum spacing of each layer multiplied by a spacing coefficient k.

line width of 10 μm is assumed. The area above the target frequency is the region where the higher metal layer is dominant (the impedance of the higher metal layer is greater), while the area below the target frequency is the region where the lower metal layer is dominant (the impedance of the lower layer is greater). In Fig. 7.14, the effect of different line widths on the critical frequency is shown.

From Fig. 7.13, the higher metal layer is the dominant metal layer in terms of the impedance for signal frequencies above 3 GHz (for high routability) and 16 GHz (for low routability), assuming a 10 μm line width for both metal layers.

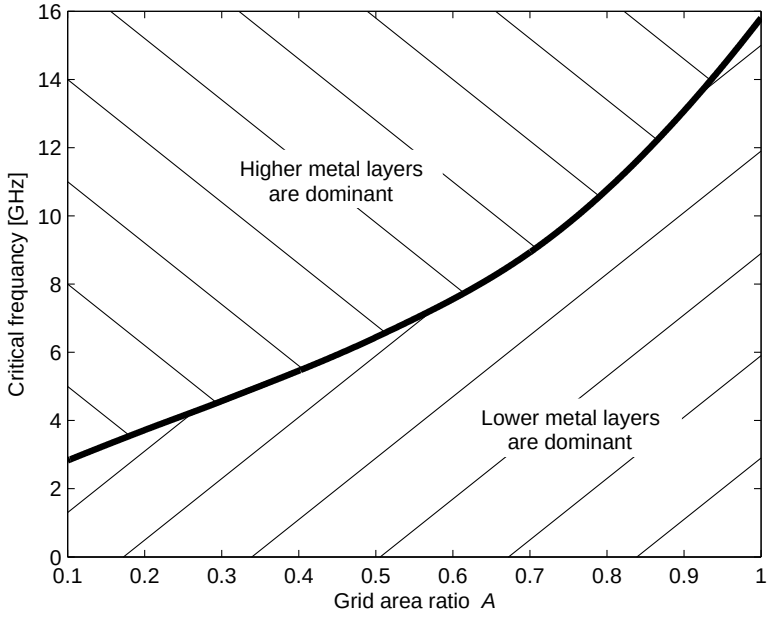


Figure 7.13: Critical frequency as a function of grid area ratio. A line width of 10 μm is assumed for all power/ground lines. The area above the line indicates the region where the higher metal layer is dominant, while the region below the line indicates the region where the lower metal layer is dominant.

7.5 Summary

The characteristics of impedance and current density in interdigitated multi-layer power/ground distribution network is investigated in this chapter. Considering the current density and network impedance, two approaches for designing a multi-layer interdigitated P/G distribution network are evaluated. For each approach, the *pyramid* (proposed), *inverted pyramid* (standard), and *equal width* P/G structure are considered. The proposed methodology for equal current density improves the effective impedance and limiting current density by 50% and 82%, respectively, as compared

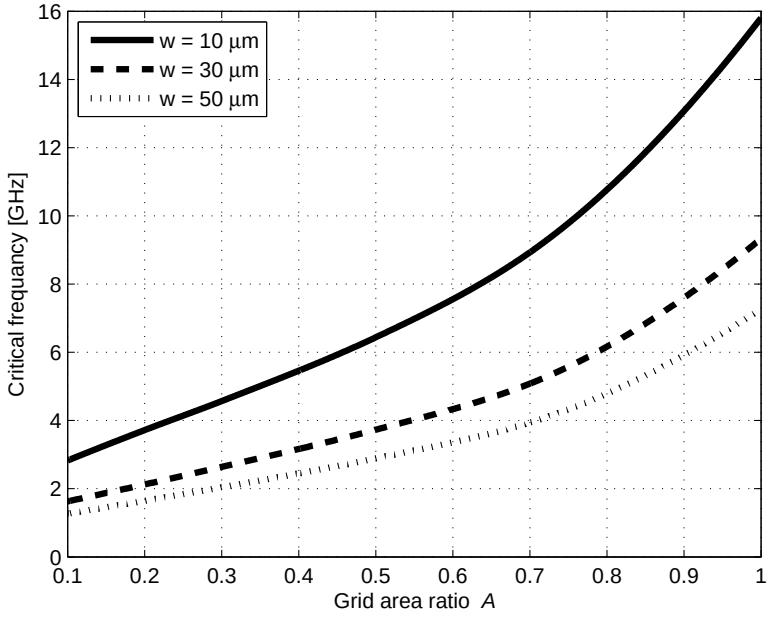


Figure 7.14: Critical frequency as a function of grid area ratio for three line widths.

with the conventional *inverted pyramid* structure. The methodology for the minimum impedance achieves a 6% and 8% enhancement, respectively, as compared to the *inverted pyramid* structure for the effective impedance and current density. This behavior is due to the relative change in importance of the inductance as compared to the resistance in high frequency systems.

Based on the proposed analytic model, the optimal width of each metal layer for minimum effective impedance is determined for 65 nm, 45 nm, and 32 nm CMOS technologies. The grid area ratio is introduced, demonstrating enhanced routability in the higher metal layers as compared with the lower metal layers. Under different

routability constraints, the frequency at which the higher metal layers are more dominant as compared to the lower metal layers is determined. Several *pyramid*, *equal width*, and *inverted pyramid* structures are compared in terms of the impedance and grid-area-ratio, indicating a tradeoff between the impedance and routability.

Chapter 8

Power Network Optimization Based on Link Breaking Methodology

The power distribution network is conventionally designed to achieve a target impedance over a wide range of frequencies [118]. This target impedance is based on minimizing the maximum voltage drop within an on-chip power network, where this drop is typically less than 10% [134]. The overall on-chip power distribution network is therefore designed to satisfy a worst case scenario at a specific location within the grid.

A change in voltage at the power node of a gate can significantly increase the delay of a logic gate [119, 135, 136], degrading the overall performance of a system [137]. Since different circuits are affected differently by a drop in the power supply voltage, the power distribution network should be designed to satisfy multiple constraints. The voltage level for those gates along the critical path can tolerate the least voltage

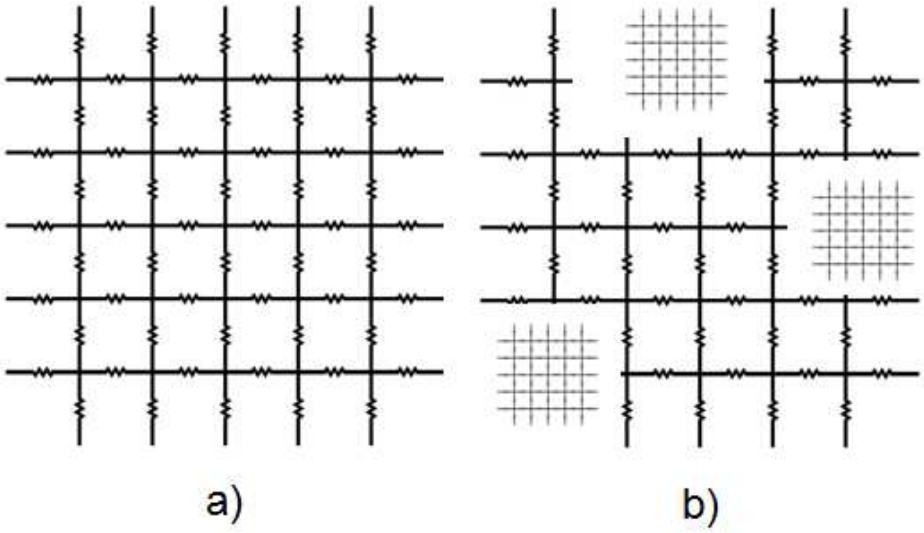


Figure 8.1: Mesh structured power distribution network. (a) single power distribution network focused on reducing the network impedance. (b) multiple power distribution networks lower the noise at the expense of increasing the network impedance.

degradation, while the gates along a non-critical path may satisfy speed constraints despite a higher voltage drop [39]. Circuits such as PLLs (phase lock loops) and VCOs (voltage controlled oscillators) are highly sensitive to changes in the power supply [138], while digital logic circuits can tolerate much higher variations in the power supply voltage. The voltage level of a power distribution network across an entire IC is typically maintained within 10% degradation, while for a PLL, the voltage level should satisfy a maximum 2% voltage degradation. To satisfy these constraints, the current supplied to the PLL is filtered by a DC-to-DC converter or a large on-chip decoupling capacitance placed near the PLL [139]. The decoupling capacitors and DC-to-DC converters however consume large area and can dissipate significant power [140].

Separate power networks can be designed to independently supply current to different parts of a circuit; thereby shielding different parts of an IC from each other. Separate power networks are widely used in mixed-signal circuits, where the current is supplied by different power networks to the analog and digital circuits [141]. For systems requiring the same voltage level, this approach however may inefficiently utilize metal resources due to the additional area and routing constraints [39]. The number of I/O pads is also a limiting resource, preventing the use of an excessive number of separate power networks [142]. In Fig. 8.1, a single and multiple separate power networks are illustrated. With a single network, as shown in Fig. 8.1(a), the sensitive circuit (for example, a PLL) and aggressor circuit (exemplified by a large digital logic circuit) share the same power network, thereby lowering the power network impedance. A sensitive circuit can however be highly affected by the noise generated from an aggressor circuit. With multiple power networks, as shown in Fig. 8.1(b), one network can be dedicated to an aggressor circuit while another network can be dedicated to the sensitive circuits, minimizing noise coupling between the aggressor and sensitive circuits. This approach however results in an increase in the power network impedance. The methodology proposed in this chapter utilizes a single power network to provide a low network impedance, while disconnecting (or breaking) links within the on-chip power network between an aggressor and sensitive circuit, thereby reducing noise coupling to the sensitive circuits.

This chapter is organized as follows. The primary design objective for reducing voltage variations is formulated in Section 8.1. An example where links within the on-chip mesh structured power distribution network are disconnected is described in Section 8.2. The sensitivity of the victim circuits to variations in the voltage within the power network is characterized in Section 8.3. In Section 8.4, the proposed link breaking methodology is described. An algorithm for breaking links for a large number of aggressor and victim circuits connected to a common on-chip power distribution network is also described in this section. In Section 8.5, several design cases are evaluated. The degradation in the supply voltage and propagation delay before and after applying the proposed link breaking methodology is summarized. Additional discussion related to enhancing the voltage levels within an on-chip power distribution network and the computational runtime of the algorithm is presented in Section 8.6. Finally, the conclusions are summarized in Section 8.7.

8.1 Reduction in Voltage Variations

The voltage drop ΔV_x at node x within a mesh structured power network, illustrated in Fig. 8.2, is a superposition of the voltage drop independently produced by each current source. Disconnecting a link on a mesh structured network increases the voltage drop at node x produced by current I_x. The voltage drop at node x produced by other currents $I_{j,j\neq x}$ is however reduced. If only a single node x is considered, the

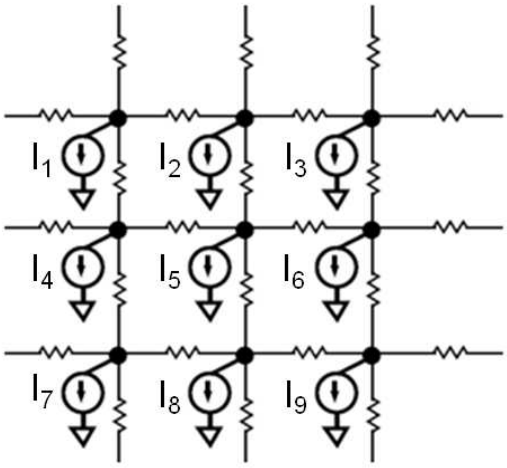


Figure 8.2: Mesh structured power network with current sources.

objective is to minimize the overall voltage drop ΔV_x.

Consider the case where circuits A and B are connected to a simple power distribution network, as illustrated in Fig. 8.3(a). The current sunk by circuits A and B is, respectively, I_A and I_B. The impedance from the power supply to the circuit is, respectively, Z_A and Z_B. The impedance of the power network between circuits A and B is denoted as Z_{AB}.

The voltage drop on the power distribution network at node A (the location where circuit A is connected to the power network) due to the current sunk by circuit A is

$$\Delta V_A = I_A \cdot [Z_A || (Z_{AB} + Z_B)] . \tag{8.1}$$

The voltage drop on the power distribution network at node A due to the current

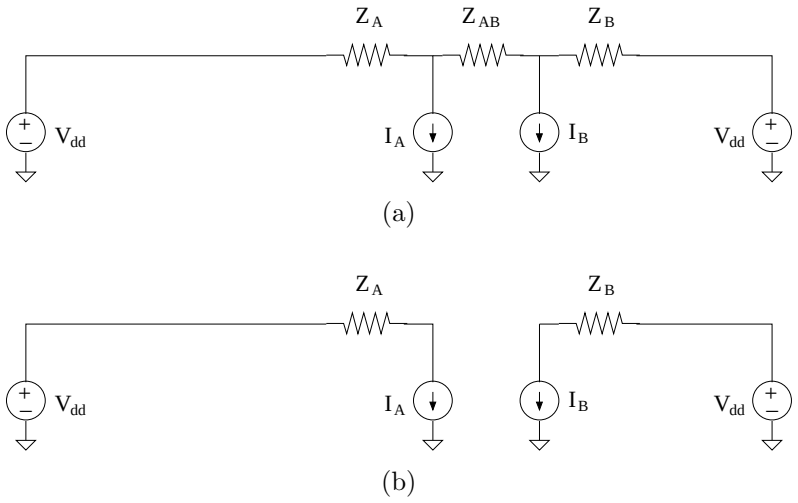


Figure 8.3: Two circuits connected to a simple power distribution network, (a) common power network for both circuits and (b) separate power networks for each circuit.

sunk by circuit B is treated as noise injected by circuit B at circuit A. This voltage drop is

$$\Delta V_A = I_B \cdot [(Z_A + Z_{AB}) \,||Z_B] \cdot \frac{Z_A}{Z_A + Z_{AB}}. \tag{8.2}$$

The overall voltage drop at node A is the superposition of (8.1) and (8.2),

$$\begin{aligned} \Delta V_A =& I_A \cdot [Z_A|| \,(Z_{AB} + Z_B)] + \\ & I_B \cdot [(Z_A + Z_{AB}) \,||Z_B] \cdot \frac{Z_A}{Z_A + Z_{AB}}. \end{aligned} \tag{8.3}$$

Similarly, the voltage drop at node B is

$$\Delta V_B = I_B \cdot [Z_A || (Z_{AB} + Z_B)] + I_A \cdot [(Z_A + Z_{AB}) || Z_B] \cdot \frac{Z_B}{Z_B + Z_{AB}}. \quad (8.4)$$

Assuming circuit B is an aggressor ($I_B \gg I_A$) and circuits A and B are located in close physical proximity ($Z_{AB} \ll Z_A \; and \; Z_B$), the voltage drop at nodes A and B is dominated by the voltage drop ΔV_B. To protect circuit A from circuit B, link Z_{AB} should be disconnected, as illustrated in Fig. 8.3(b), resulting in a voltage drop at nodes A and B,

$$\Delta V_A = I_A Z_A \quad (8.5)$$

$$\Delta V_B = I_B Z_B. \quad (8.6)$$

In this example, the objective is to determine if the link Z_{AB} needs to be broken. If the voltage drop at node A is lower when link Z_{AB} is disconnected as compared with the configuration where Z_{AB} is connected, link Z_{AB} should be broken. Note that by disconnecting link Z_{AB}, the voltage drop at node B also changes, requiring the voltage drop at node B to be evaluated and maintained below some limit.

Since every circuit is an aggressor and a victim, the problem formulation and objective needs to be generalized. Two parameters are therefore assigned to each circuit, one characterizing the aggressiveness and the second the sensitivity of a circuit. The

aggressor parameter is directly related to the current sunk by a circuit. Simultaneously, every circuit exhibits a different sensitivity to variations in the power network voltage. For example, a PLL is highly sensitive to voltage variations as compared to digital logic. Two circuits with a different critical path may also exhibit a different sensitivity to voltage variations: a slower critical path requires a smallest power drop, while a fast critical path can better tolerate a large voltage drop on the power network. A sensitivity factor is therefore assigned to each circuit connected to the power network. A more detailed discussion of the sensitivity factor is presented in Section 8.3.

In a system with multiple aggressors and victims, the objective is to minimize the effect of the voltage drop over the entire system. To improve the performance of an IC, the voltage drop is reduced in those circuits with high sensitivity at the expense of increasing the voltage drop in the less sensitive circuits.

Breaking a link between two circuits in a mesh structured power distribution network does not however completely isolate these circuits, rather resulting in an increase in the impedance between the two circuits. The larger impedance between the circuits lowers the noise coupling between the two nodes. Three specific nodes, the victim, aggressor, and power supply, within a mesh structured power distribution network, are illustrated in Fig. 8.4. The objective is to increase the network impedance between the victim and the aggressor nodes (Z_{AB}), reduce the influence of the aggressor

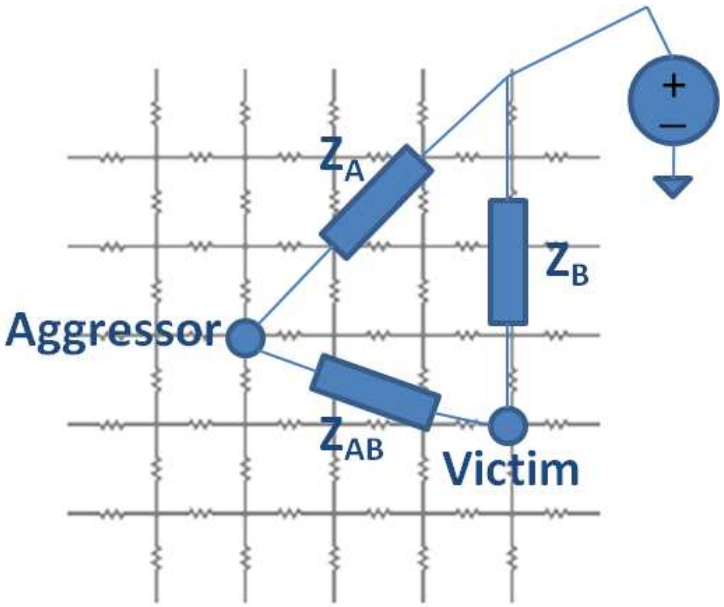


Figure 8.4: Aggressor and victim circuits sharing a mesh structured power distribution network. The objective is to increase Z_{AB}, while insignificantly increasing Z_A, resulting in shielding the victim from an aggressor.

on the victim node, while only minimally increasing the effective impedance between the aggressor and the power supply (Z_A).

The normalized effective resistance between the left and right nodes as a function of a specific disconnected link at a particular location (along the x-axis) is depicted in Fig. 8.5. A 20×20 node mesh structured network is illustrated in Fig 8.6. The x-axis describes the location (or link number depicted in Fig. 8.6) of the disconnected link between two nodes. The largest increase in the effective resistance is achieved when breaking the link closest to either node. An 11% increase in the resistance is caused by breaking a single link. This change confirms that breaking links within a mesh structured power distribution networks may result in a large change in the effective

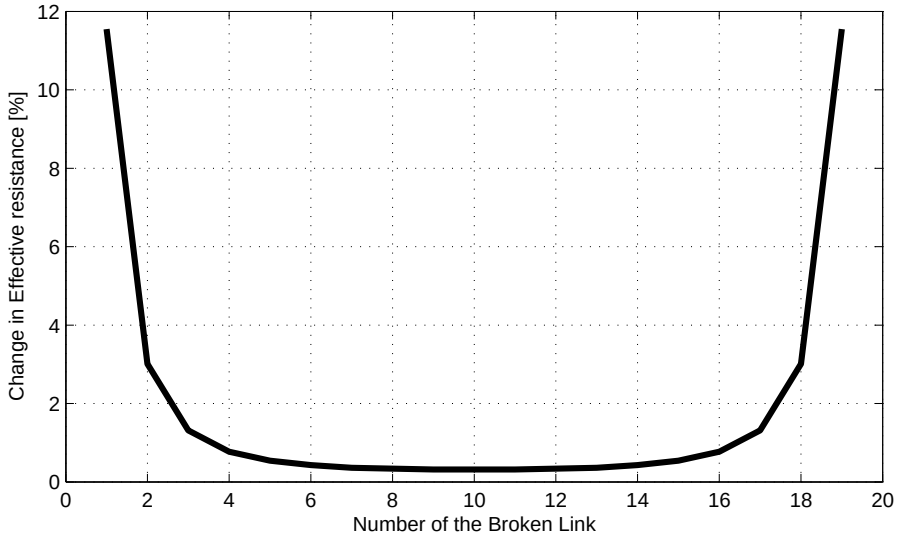


Figure 8.5: A change in the effective resistance between the left and right nodes within a 20×20 mesh structured power distribution network (see Fig. 8.6) as a function of a specific location of a disconnected link between two nodes.

impedance; effectively shielding the victim from the aggressor.

8.2 Single Aggressor and Victim Example

A single aggressor and single victim example is provided in this section, intuitively illustrating the problem and solution. A 25×25 node mesh structured power network is illustrated in Fig. 8.7(a). Two power pads are located at the top/right and bottom/left nodes. The aggressor, a large current sink, is connected in the left/center region of the mesh network, while the victim circuit is connected in the right/center region. The current sunk from the power network by an aggressor is two orders

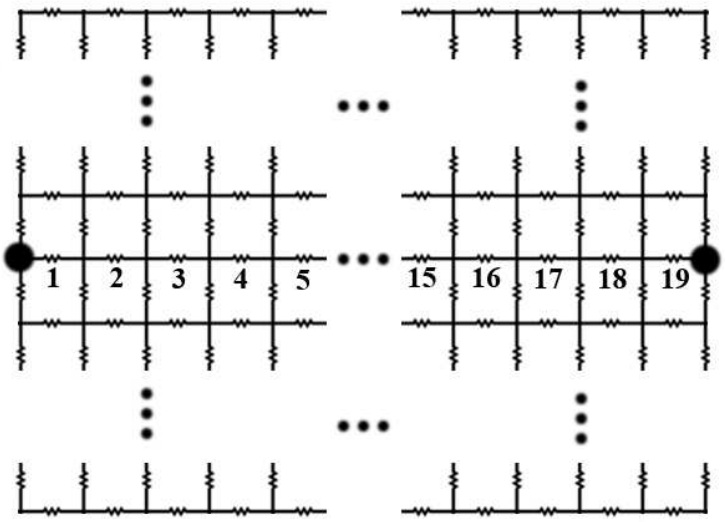
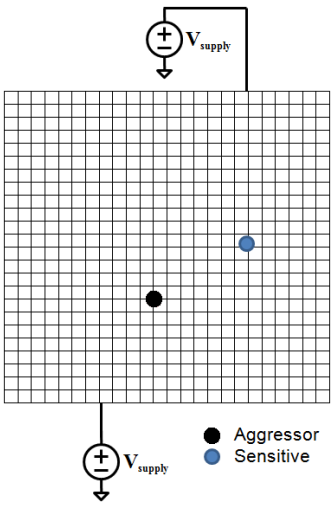


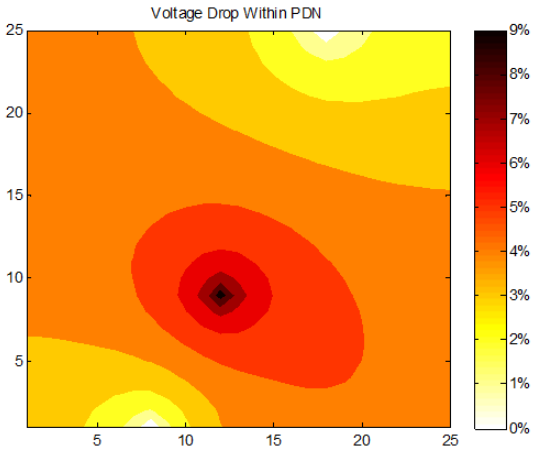
Figure 8.6: 20×20 node mesh structured network. The effective resistance is between the two bold nodes. The links are numbered based on the location along the x-axis.

greater than the current sunk by the victim circuit. A one volt power supply voltage is assumed. The voltage drop is shown in Fig. 8.7(b) as a shade of color, where the darker shade represents a higher voltage drop. The aggressor and victim nodes exhibit, respectively, a 96 and 47 millivolt voltage drop.

As discussed previously, the design objective is to reduce the voltage drop at the victim node, while insignificantly increasing the voltage drop at the aggressor node. The disconnected links are therefore required to be far from the aggressor and close to the victim. The links are removed around the victim node, isolating the victim from the rest of the network while maintaining a single connection to the power supply (the lowest voltage drop). A large number of additional connections may also be provided

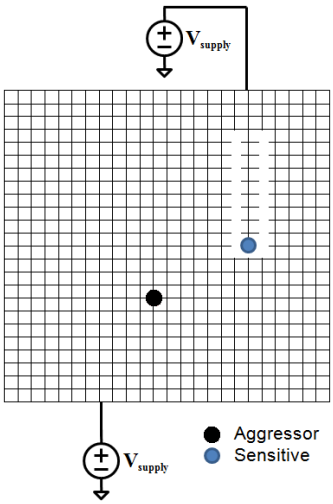


(a)



(b)

Figure 8.7: Two circuits, an aggressor and a victim, are connected to a 25×25 node mesh structured power distribution network. (a) schematic of the power network and (b) voltage drop within the power distribution network before disconnecting any links.



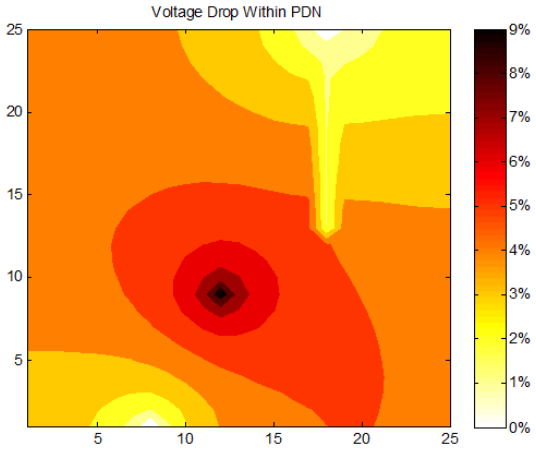


Figure 8.8: Two circuits, an aggressor and a victim circuits, are connected to a 25×25 node mesh structured power distribution network. (a) schematic of the power network and (b) voltage drop within the power distribution network after disconnecting nine pairs of links. Note that the voltage drop for the victim circuit is significantly lower after disconnecting the links.

based on reliability and current density constraints. The procedure is repeated until the target low voltage drop at the victim node is achieved. Note that the voltage drop at the aggressor node is simultaneously monitored while disconnecting links. The procedure is discontinued once the voltage drop at the aggressor node exceeds the target limit or the desired voltage drop at the victim node is achieved.

The voltage drop for the revised 25×25 node mesh structured network is illustrated in Fig. 8.8(a). Nine sets of links have been disconnected, producing a 20 millivolt voltage drop at the victim node, while the voltage drop at the aggressor node increased from 96 to 98 millivolt, as depicted in Fig. 8.8(b). The improvement in the variation of the power voltage at the victim node is 135%, while a voltage degradation of only 2.1% is observed at the aggressor node. Since in practical applications each node within the network can be simultaneously both an aggressor and a victim, the methodology has been developed to address this issue based on maximizing the overall performance of the circuit.

8.3 Sensitivity Factor

The sensitivity factor describes the relative importance of a change in voltage on the performance of a circuit. A method to describe the sensitivity factor is to investigate the sensitivity of the supplied voltage on the performance (for example,

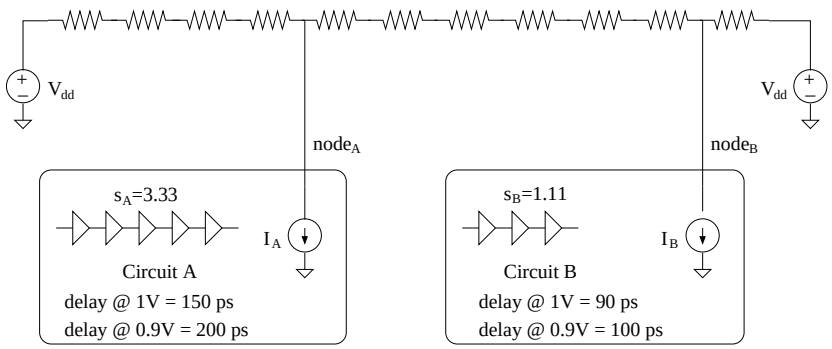


Figure 8.9: Example to determine the sensitivity factor, where two circuits have different propagation delays.

the propagation delay) of a particular circuit. The sensitivity factor is [143]

$$s = \left.\frac{\frac{\Delta delay}{delay(x)}}{\frac{\Delta V}{V(x)}}\right|_{x=V_{dd}} = \frac{\Delta delay}{\Delta V} \cdot \frac{V_{dd}}{delay_{min}}, \tag{8.7}$$

where $\Delta delay$ and $delay_{min}$ are, respectively, the change in the delay and the minimum delay of a circuit. The minimum delay is achieved assuming a full V_{dd} at the power rail of the circuit. ΔV is the change in the supply voltage at the node supplied to the circuit. The sensitivity factor is dependant on the type of circuit.

Consider an example where two circuits are connected to the power network, as depicted in Fig. 8.9. With a power supply of one volt (full V_{dd}) applied to $node_A$ and $node_B$, the propagation delay of the critical path within circuit A is 150 ps, while the propagation delay of the critical path within circuit B is 90 ps. Reducing the power level by 10%, the delay of circuits A and B is, respectively, degraded to 200 ps and

100 ps. The resulting sensitivity factor for circuits A and B is therefore, respectively,

$$s_A = \frac{200ps - 150ps}{1 - 0.9} \cdot \frac{1}{150ps} = 3.33, \tag{8.8}$$

$$s_B = \frac{100ps - 90ps}{1 - 0.9} \cdot \frac{1}{90ps} = 1.11. \tag{8.9}$$

8.4 Link Breaking Methodology

An algorithm for determining which links should be removed, thereby shielding the sensitive circuits is described in this section. The described algorithm is only one of many algorithms and is not necessarily the optimal algorithm for the link breaking methodology. Since each circuit within a network can be characterized as both an aggressor and a victim, each node of interest is associated with a matrix composed of two parameters $[i, s]$. Parameter i is an aggressor related parameter, and is equal to the current sunk from the network. Parameter s is related to the victim parameter, expressing the sensitivity of the circuit connected to the node. The objective is to enhance overall performance, such as minimize the average (8.10) or worst case (8.11) delay.

$$delay_{average} = \frac{1}{k} \sum_{j=1}^{k} \left(delay_j \right), \tag{8.10}$$

$$delay_{worst} = max \left(delay_1, delay_2, ..., delay_k \right), \tag{8.11}$$

LINK-BREAKING
1. Determine voltage drops at all k nodes
2. Calculate initial $delay_{ini}$ function based on (8.11)
3. Generate x randomly perturbed systems
4. Determine voltage drops at k nodes for x systems
5. Calculate *delay* function based on (8.11) for x systems
6. For every x systems
7. Generate six different networks,
   where a link is broken at every direction
8. Determine new *delay* values, maintaining network with lowest *delay*
9. Goto 7, if enchantment is achieved
10. Select system with lowest *delay*
11. If $delay_{ini} > delay$ then $delay_{ini} \leftarrow delay$ and goto 3

Figure 8.10: Pseudocode of the link breaking algorithm.

where

$$delay_j = delay_{min-j} \cdot \left[\frac{s_j}{V_{dd}} \cdot \Delta V_j + 1 \right]. \tag{8.12}$$

ΔV_j is a change in the voltage at node j due to currents $i_1, i_2, ..., i_k$ and the impedance of the mesh structured power distribution network. $delay_{min-j}$ is the minimum propagation delay of circuit j achieved by applying the maximum supply voltage V_{dd}. s_j is the sensitivity factor of circuit j.

Pseudocode of the LINK-BREAKING algorithm for the proposed methodology is provided in Fig. 8.10, with the objective of minimizing the worst case propagation delay. Other algorithms can be used which may yield enhanced computational efficiency or a global solution to the link breaking methodology. In line 1, the voltage drop at k nodes (all aggressor/sensitive nodes) is determined. Based on the voltage

and sensitivity of the circuits, the initial value of the delay function $delay_{ini}$ is determined, as listed in line 2. The revised number of power networks x is generated, where each network is perturbed by removing a random link. In lines 4 and 5, the voltage drop and delay function are determined for each of the perturbed networks. A search for a local minimum is evaluated for each perturbed system in lines 6 to 9. The network with the lowest *delay* value is selected in line 10. The process is repeated until the value of the delay function cannot be further reduced.

Since k nodes of interest are typically lower than the overall number of nodes in a system, a random walk procedure can be used to efficiently determine the voltages [144], trading off accuracy with runtime. The number of parallel random walk procedures is based on the target accuracy.

8.5 Case Studies

Five study cases are presented in this section. In each of the cases, the circuit is composed of nine blocks. In the first case study, the current sunk by every block is maintained equal. The sensitivity factor and critical delay of each block however are assumed different. For the following two cases, one block sinks significantly greater current, representing the case of a single dominant aggressor. In the final two cases, the current and delay of the nine blocks are varied, representing general circuits. The design objective is to minimize the worst case propagation delay; (8.11) is therefore

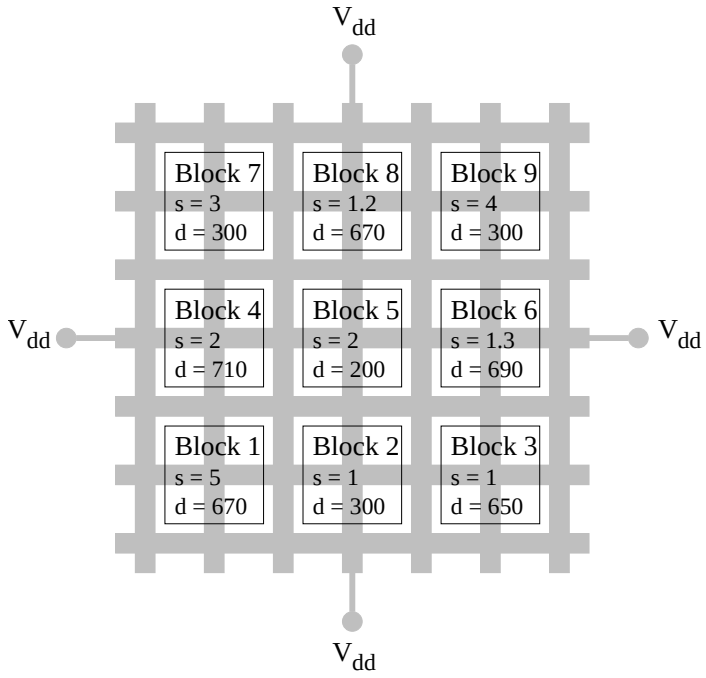


Figure 8.11: Nine circuit blocks are connected to a mesh structured power distribution network. Four power supplies provide the current. The numbers indicated within the blocks represent the sensitivity factor (s) and propagation delay in ps (d) when applying one volt to the block. Note that the minimum propagation delay is achieved when applying a full power supply.

used for all of the five case studies.

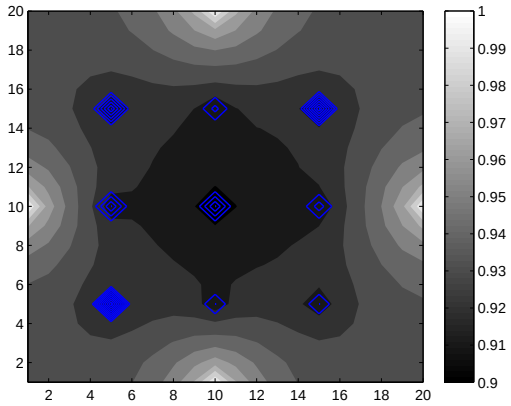
A mesh structured power distribution network with 20×20 number of nodes is considered. A block diagram of the circuit is schematically illustrated in Fig. 8.11. Four one volt power supplies are connected at the center of the four edges (left, right, top, and bottom). The maximum permitted degradation in supply voltage is 0.3 V.

For Case 1, a map illustrating the variation in voltage over the mesh structured power network is shown in Fig. 8.12(a). A darker shaded color represents a lower voltage within the power network. After applying the link breaking methodology,

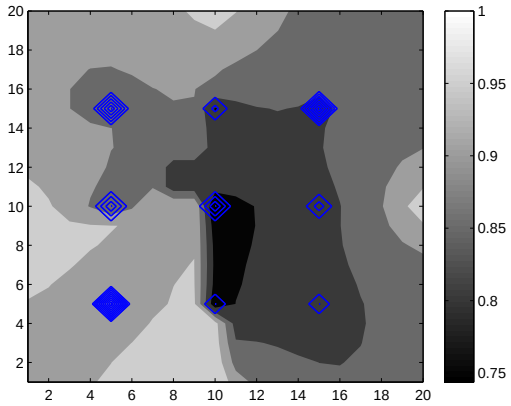
38 links are disconnected from 760 possible links within the power network. The resulting power network is illustrated in Fig. 8.13. The voltage map after the link breaking methodology is shown in Fig. 8.12(b). Note that the power supply voltage is increased at the lower left corner due to the high sensitivity factor ($s = 5$) and propagation delay assigned to the block located in the lower left corner. The supply voltage is reduced at the other locations due to a lower sensitivity factor or small delay assigned to the block.

The voltage and propagation before and after application of the link breaking methodology for each block are listed in Table 8.1. The sensitivity factor, current, minimum delay, and improvement or degradation in the voltage and delay are also listed. A close to 4% improvement in the supply voltage, 95% of the ideal power supply, is achieved for block 1. Note that the maximum improvement in the supply voltage is 9%, producing a supply voltage of one volt. The improvement in voltage is achieved at the expense of a lower supply voltage at the other blocks. The performance of the overall circuit is increased since the worst case propagation delay at block 1 is reduced. Due to the higher supply voltage, the propagation delay at block 1 is lowered from 1 ns to 861 ps, permitting an increase in the maximum operating frequency of the overall IC.

For Cases 2, 3, 4, and 5, the current is different among the circuit blocks (see



(a) Supply voltage before applying the link breaking methodology.



(b) Supply voltage after applying the link breaking methodology.

Figure 8.12: Supply voltage before and after the proposed link breaking methodology for Case 1. The diamond shapes represent the location of the aggressor/victim circuit blocks. In this example, the current sunk by each of the nine blocks is assumed equal. The voltage drop is reduced in the more sensitive circuit blocks (resulting in a smaller delay), while increased in the less sensitive circuit blocks (resulting in a higher delay).

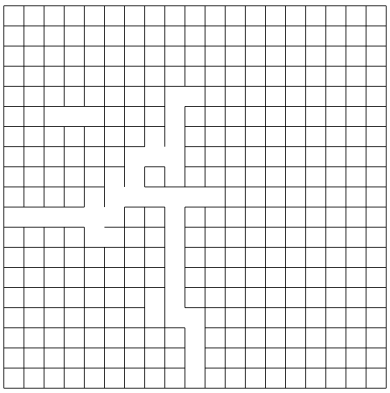


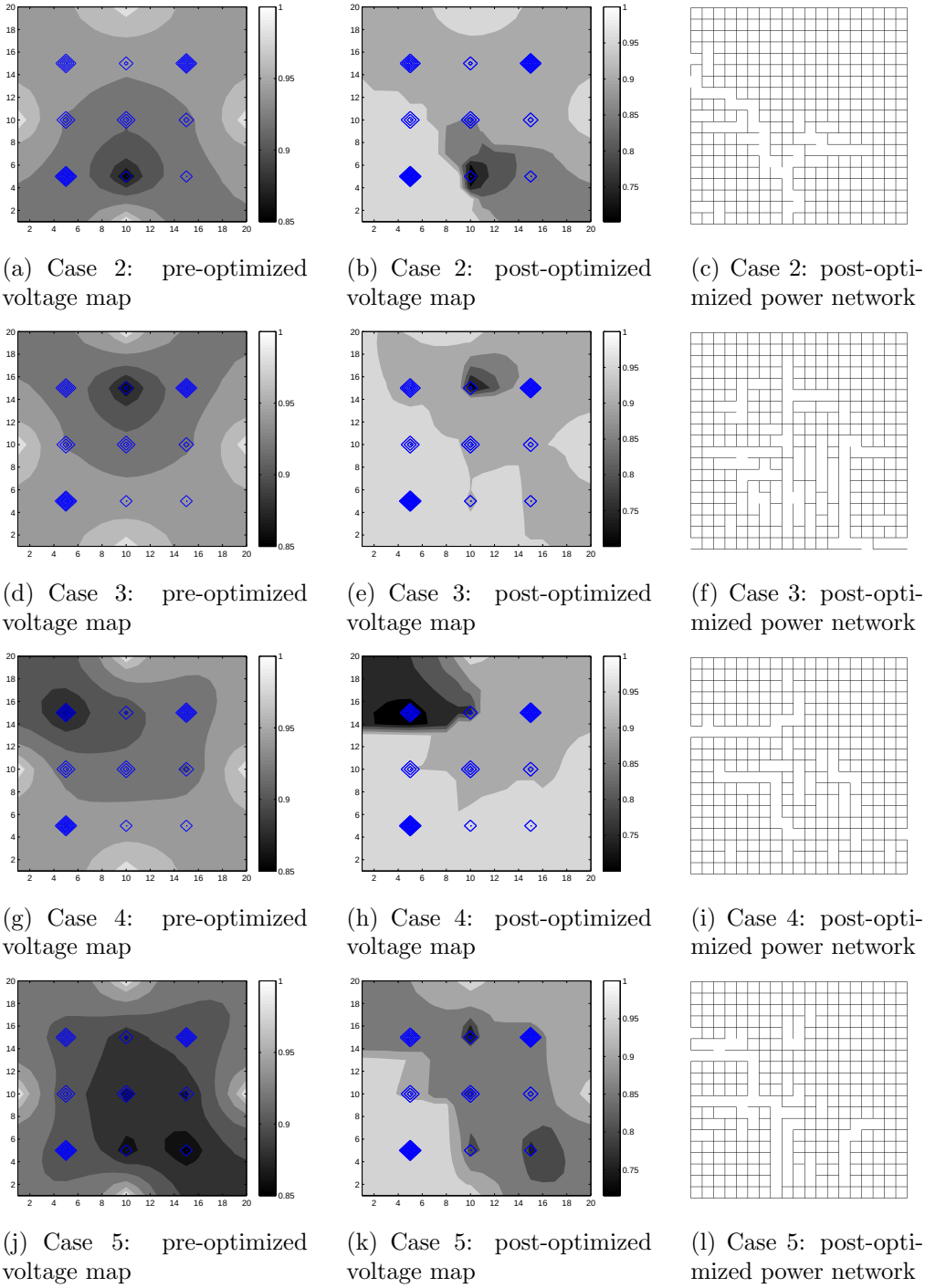
Figure 8.13: The resulting power network after applying the link breaking methodology for Case 1.

Table 8.1: Case 1. Sensitivity factor, sunk current, minimum delay (achieved with one volt at the power rail of the block), supply voltage, and propagation delay before and after the link breaking methodology for the nine circuit blocks. The improvement or degradation in the supply voltage, propagation delay, and maximum operating frequency are also listed. A one volt power supply is used.

Block number		1	2	3	4	5	6	7	8	9	f_{max}
Sensitivity factor		5	1	1	2	2	1.3	3	1.2	4	——
Sunk current		1	1	1	1	1	1	1	1	1	——
Delay [ps] @ $V_{dd} = 1V$		670	300	650	710	200	690	300	300	300	——
Voltage	before [mV]	914	914	913	914	900	910	913	910	912	——
	after [mV]	949	840	877	909	829	880	912	864	891	——
	improvement [%]	3.8	-8.1	-3.9	-0.5	-7.9	-3.3	0.1	-5.1	-2.3	——
Delay	before [ps]	1024	334	723	880	249	785	393	348	430	0.98 GHz
	after [ps]	861	372	760	892	274	816	387	357	440	1.16 GHz
	improvement [%]	15.9	-11.4	-5.1	-1.4	-10.0	-3.9	1.5	-2.6	-2.3	15.9

Table 8.2). The supply voltage map before and after application of the link breaking methodology, as well as the resulting power network, is illustrated in Fig. 8.14. The current sunk for each of the cases, voltage before and after application of the methodology, sensitivity, propagation delay, and improvement in the supply voltage and propagation delay are listed in Table 8.2.

Case 2 (Figs. 8.14(a) and 8.14(b)) and Case 3 (Figs. 8.14(d) and 8.14(e)) illustrate those cases where the current sunk by a circuit (the aggressor) is significantly higher as compared to the other circuit blocks. The sensitivity factors and minimum delay are the same as in Case 1. The highest degradation in the supply voltage is within the aggressor circuit; however, the supply voltage is greater in those circuit blocks with a higher sensitivity and minimum delay, resulting in a reduction in the worst case delay and a higher maximum operating frequency. The increase in the supply voltage at block 1 is 5%, achieving 97% of the ideal power supply voltage and resulting in an improvement in the propagation delay of 16%. Note that the improvement in the propagation delay is greater than the supply voltage due to the high sensitivity factor. After applying the link breaking methodology, blocks 1, 4, and 6 exhibit a similar worst case propagation delay, demonstrating the effectiveness of the proposed methodology. In Case 3, the voltage at block 1 is increased by 2%, achieving 96% of the ideal power supply voltage and resulting in an improvement in the propagation delay of 5%.



(a) Case 2: pre-optimized voltage map

(b) Case 2: post-optimized voltage map

(c) Case 2: post-optimized power network

(d) Case 3: pre-optimized voltage map

(e) Case 3: post-optimized voltage map

(f) Case 3: post-optimized power network

(g) Case 4: pre-optimized voltage map

(h) Case 4: post-optimized voltage map

(i) Case 4: post-optimized power network

(j) Case 5: pre-optimized voltage map

(k) Case 5: post-optimized voltage map

(l) Case 5: post-optimized power network

Figure 8.14: Map of voltage variations before and after application of the link breaking methodology for Cases 2, 3, 4, and 5. The diamond shapes represent the location of the aggressor/victim circuit blocks. The resulting power network after the link breaking methodology is also illustrated. Cases 2 and 3 represent the cases where a single block sinks significantly higher current as compared to the other blocks. In Cases 4 and 5, the sunk current, sensitivity factor, and delay are different for different blocks, representing general design cases.

Table 8.2: Sensitivity factor, sunk current, minimum delay, supply voltage, and propagation delay before and after application of the link breaking methodology for the nine circuit blocks. The improvement or degradation in the supply voltage, propagation delay, and maximum operating frequency are also listed. Cases 2 and 3 represent the cases where a single block sinks significantly higher current as compared to the other blocks. In Cases 4 and 5, the sunk current, sensitivity factor, and delay are different for different blocks, representing general design cases.

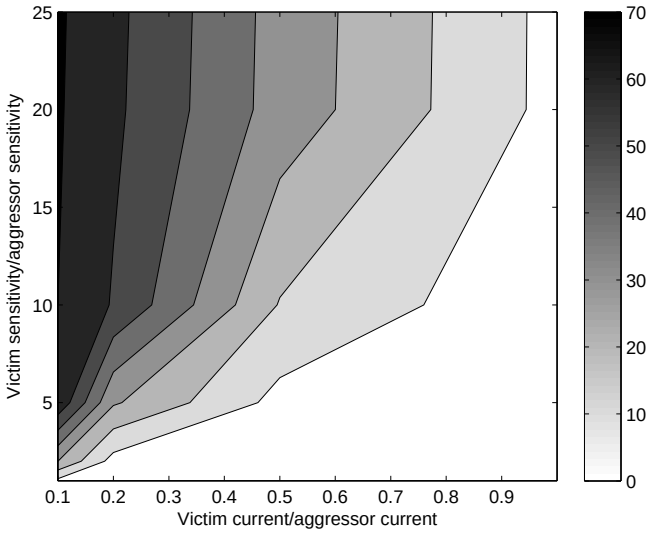
Block number		1	2	3	4	5	6	7	8	9	f_{max}
Sensitivity factor (s)		5	1	1	2	2	1.3	3	1.2	4	——
Delay [ps] @ $V_{dd} = 1V$		670	300	650	710	200	690	300	300	300	——
Case 2 (see Figs. 8.14(a) and 8.14(b))											
Sunk current		1	10	1	1	1	1	1	1	1	——
Voltage	before [mV]	924	850	924	936	920	933	943	940	942	——
	after [mV]	973	702	861	965	896	921	921	925	931	——
	improvement [%]	5.3	-17.4	-6.8	3.0	-2.6	-1.3	-2.3	-1.6	-1.2	——
Delay	before [ps]	988	369	748	856	248	803	376	343	395	1.01 GHz
	after [ps]	829	422	807	834	263	828	403	356	417	1.20 GHz
	improvement [%]	16.0	-15.0	-7.8	2.6	-5.8	-3.2	-7.4	-3.7	-5.7	15.8
Case 3 (see Figs. 8.14(d) and 8.14(e))											
Sunk current		1	1	1	1	1	1	1	10	1	——
Voltage	before [mV]	945	944	944	937	922	934	925	850	924	——
	after [mV]	966	934	924	947	936	906	939	700	909	——
	improvement [%]	2.1	-1.1	-2.0	-1.0	1.6	-3.0	1.5	-17.6	-1.6	——
Delay	before [ps]	904	336	727	847	245	794	390	375	413	1.11 GHz
	after [ps]	855	349	762	857	243	844	387	445	446	1.17 GHz
	improvement [%]	5.4	-3.9	-4.8	-1.2	0.8	-6.3	0.8	-18.7	-8.0	5.4
Case 4 (see Figs. 8.14(g) and 8.14(h))											
Sunk current		5	1	1	2	2	1.3	3	12	4	——
Voltage	before [mV]	939	949	949	930	924	914	850	899	927	——
	after [mV]	967	953	959	950	938	898	700	775	923	——
	improvement [%]	3.0	0.4	1.0	2.1	1.5	-1.7	-17.6	-13.7	-0.4	——
Delay	before [ps]	929	334	724	858	244	814	461	357	411	1.08 GHz
	after [ps]	851	332	712	853	243	851	621	415	428	1.17 GHz
	improvement [%]	8.3	0.6	1.7	0.7	0.4	-4.5	-34.7	-16.2	-4.1	7.8
Case 5 (see Figs. 8.14(j) and 8.14(k))											
Sunk current		1	5	5	2	2	3	1.3	4	1.2	——
Voltage	before [mV]	907	861	850	901	874	875	907	876	901	——
	after [mV]	958	825	781	928	838	864	852	716	890	——
	improvement [%]	5.5	-4.3	-8.1	2.9	-4.1	-1.2	-6.1	-18.2	-1.3	——
Delay	before [ps]	1050	366	800	910	268	860	411	369	448	0.952 GHz
	after [ps]	870	378	847	870	283	869	463	430	463	1.15 GHz
	improvement [%]	17.1	-3.2	-5.8	4.5	-6.0	-1.1	-12.7	-16.5	-3.3	17.1

Case 4 (Figs. 8.14(a) and 8.14(b)) and Case 5 (Figs. 8.14(a) and 8.14(b)) represent cases where different current is sunk. After applying the link breaking methodology, the supply voltage at block 1 is increased by, respectively, 3% and 5%. The maximum operating frequency is enhanced by, respectively, 8% and 17%.

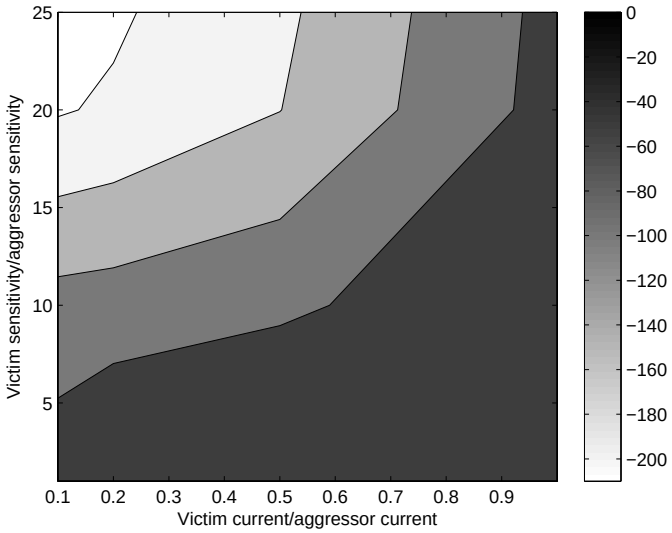
8.6 Discussion

The voltage drop within a power distribution network is investigated for circuit blocks with different current levels and sensitivities. The minimum propagation delay ($delay_{min}$) is maintained the same. A 20×20 mesh structured power distribution network with two power supplies and two current sources (one aggressor and one victim) is considered. The voltage improvement at the victim and degradation at the aggressor are illustrated, respectively, in Figs. 8.15(a) and 8.15(b). Note that by assigning a higher sensitivity to the victim circuit, the voltage drop on the power network at the victim is reduced. Simultaneously, the voltage drop at the aggressor is increased, while the aggressor is less sensitive to voltage variations. The tradeoff between reducing the voltage drop at the victim while increasing the voltage drop at the aggressor is an important aspect of the proposed link breaking methodology.

The improvement and degradation of the voltage drop at, respectively, the victim and aggressor are depicted in Fig. 8.16 for different ratios of the current sunk by the victim and aggressor, assuming the two circuits have equal sensitivity. Note



(a) Change in voltage drop at the victim



(b) Change in voltage drop at the aggressor

Figure 8.15: Change in voltage drop for the victim and aggressor circuits. The darker shade represents a greater reduction in the voltage drop at the victim and a lower increase in the voltage drop at the aggressor.

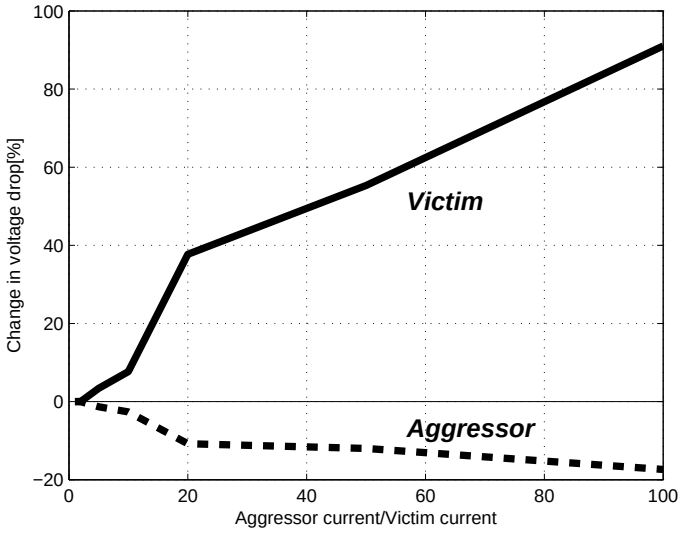


Figure 8.16: Change in voltage within the power distribution network for the victim and aggressor circuits as a function of the ratio of the current sunk by the aggressor and victim circuits. The sensitivity factor is assumed equal for both circuits.

that a higher change in voltage is achieved at the victim when the current sunk by the aggressor is greater. This effect is due to the dominance of the aggressor on the victim circuit before applying the link breaking methodology. The link breaking methodology can therefore be used to reduce the voltage drop at the victim.

The computational runtime of the algorithm, depicted in Fig. 8.10, is evaluated for differently sized power distribution networks. The algorithm has been executed on a Linux eight-core with 8 GB RAM system. The runtime as a function of the number of nodes in the power network is depicted in Fig. 8.17. The runtime of the link breaking methodology can also be accelerated by utilizing multigrid-like techniques [145] and ignoring those current sources located far from the target nodes. The number of

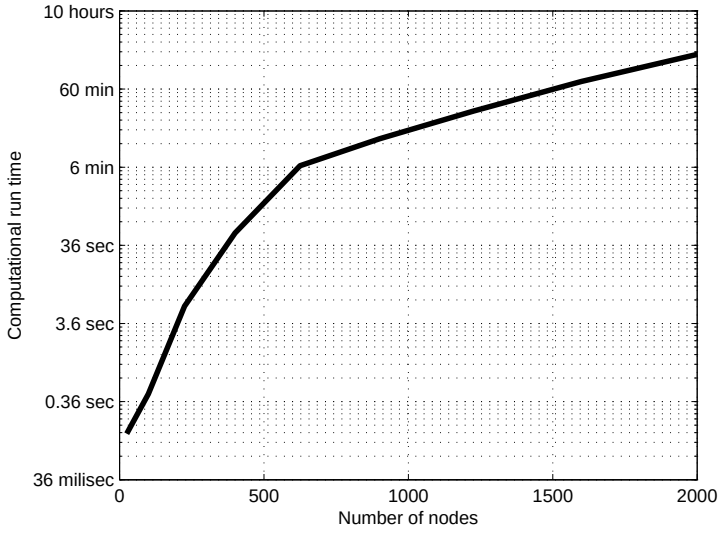


Figure 8.17: Computational runtime of the link breaking methodology as a function of the number of nodes within the power distribution network.

aggressor and/or victim circuits is not a dominant factor affecting the runtime of the algorithm, as illustrated in Fig. 8.18. Initially, the runtime increases exponentially with the number of aggressor and victim circuits. With a further increase in the number of circuits, the computational runtime decreases due to the smaller number of links that can be disconnected. For those cases where only a small number of circuit are evaluated within a large power distribution network, the random walk method [144] can be used to estimate the voltage variations, significantly accelerating the link breaking methodology.

The worst case voltage drop (located at the aggressor) cannot be reduced by utilizing the link breaking methodology, since the methodology always increases the

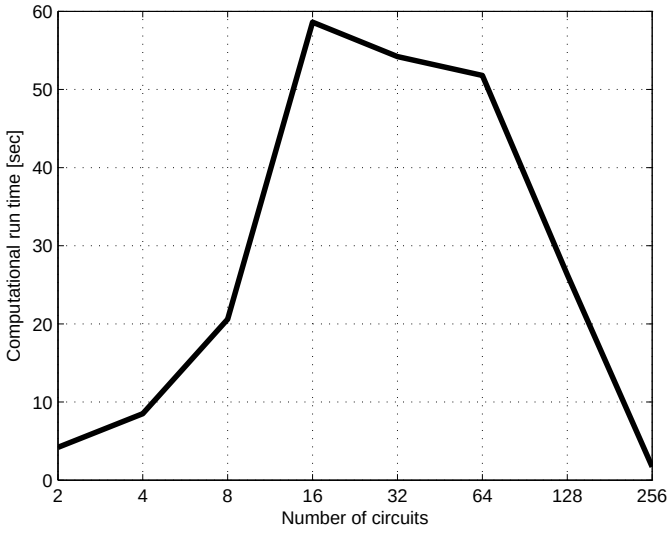


Figure 8.18: Computational runtime of the link breaking methodology as a function of the number of victim and aggressor circuits. The runtime initially increases with a higher number of circuits. After reaching a peak, the runtime decreases due to the smaller number of links that can be removed.

worst case power network impedance. The effect of the aggressor on other circuits with a higher sensitivity and propagation delay can be reduced, resulting in enhanced overall system performance.

8.7 Summary

The design of the power distribution network is an essential part of an IC design flow. The network is typically designed as a single network or multiple separate networks. The advantage of a single network is reduced network impedance, while multiple separate networks have the advantage of lower noise coupling. The proposed link breaking methodology utilizes a single network, disconnecting links between the aggressive and sensitive circuits; thereby, isolating the victim from the aggressor. This approach reduces the noise, while maintaining a low network impedance.

Sensitivity to changes in the supply voltage will vary for different circuits. A PLL is more sensitive to supply voltage variations than a digital logic gate. Voltage variations at the more sensitive circuits need to be reduced at the expense of increased voltage variations at the less sensitive circuits. A smaller voltage drop is also important in long critical paths as compared to shorter less critical logic paths.

The proposed methodology is based on a mesh structured power distribution network. The aggressiveness and sensitivity of a circuit is considered during the link breaking process. The methodology is evaluated for several cases with a different

number and magnitude of current and sensitivity factors. The objective for these study cases is reduced worst case propagation delay by increasing the supply voltage at blocks with high propagation delay. An average enhancement of 4% in power supply voltage at nodes with high sensitivity and high propagation delay is achieved, resulting in, on average, 93% of the ideal power supply voltage at these nodes. As a result, an average improvement of 12% in the maximum operating frequency is achieved when utilizing the proposed link breaking methodology. Acceleration of the algorithm for lower computational runtime is also discussed.

Chapter 9

Globally Integrated Power and Clock Distribution Networks

Further increases in the density and performance of integrated circuits (IC) require more complicated global interconnects, such as power, ground, and clock networks. On-chip metal resources are limited [146], further constraining the design of these global interconnect networks. These global networks require a large portion of the overall metal resources [7].

Major networks that consume most of the top metal resources are the power, ground, and clock networks. Each network is carefully designed to provide optimal circuit performance. These networks are typically designed independent of each other (power/ground network and clock network) since each network exhibits different characteristics and constraints. This approach results in high utilization of on-chip resources since each network is typically routed to every individual block, circuit, or gate within an IC.

For those circuits where the clock signal is generated off-chip, the clock and power signals may be integrated to eliminate the on-chip global clock distribution network. This combined network is named here as a *globally integrated power and clock* (GIPAC) distribution network. The power and clock signals are later separated from the GIPAC network into two different local networks using passive filters.

Signal splitting is fairly common in electrical and communication systems. Different signals are modulated, simultaneously transferring these signals over a single medium, and later demodulated [147]. A typical home phone system carries power and voice signals over the same network. Filters inside the device demultiplex the signals, routing the signals accordingly. The internet infrastructure also uses phone lines, sharing resources with the global phone network. Broadband communication over the power lines [148] further supports this approach of sharing a common global network.

The two signals, power and clock, are fundamentally different signals. A primary difference between the power and clock signals is the operating frequency. While the clock signal exhibits a high frequency component, the power signal is ideally DC. These two signals can therefore be separated with high and low pass filters, as illustrated in Fig. 9.1. Additionally, the power signal carries high current, while high current is not required for the clock signal. Different characteristics of these two signals, listed in Table 9.1, distinguish the design of the high and low pass filters (the

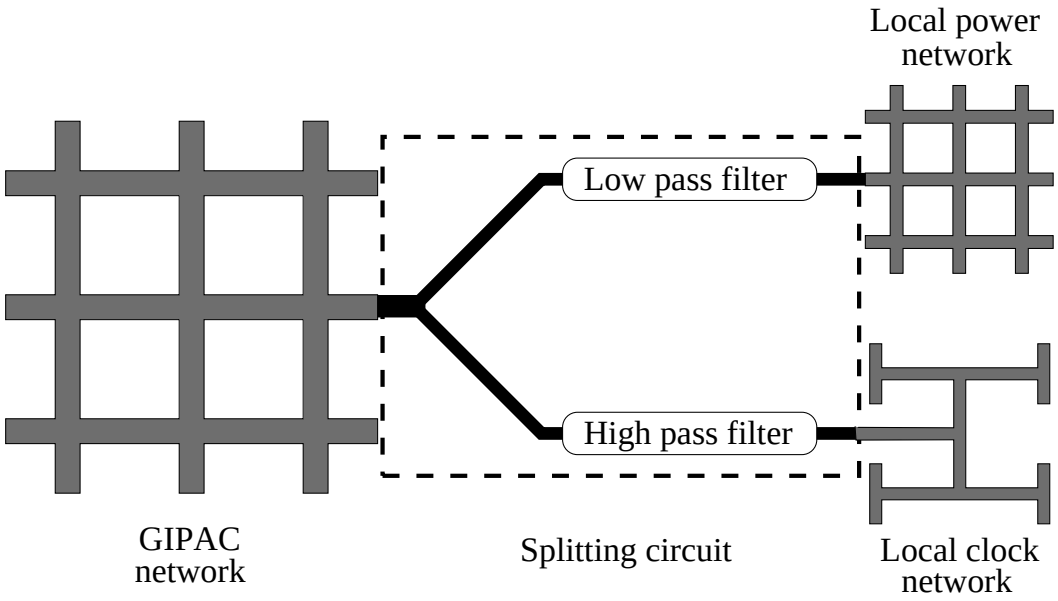


Figure 9.1: Globally integrated power and clock (GIPAC) distribution network. Low and high pass filters can be used to separate the GIPAC signal into local power and local clock signals.

Table 9.1: Characteristics of power and clock signals.

	Power signal	Clock signal
Frequency	Very low	High
Current	Very high	Medium
Load	Very low resistance	Highly capacitive

splitting circuit).

Previous research on power and clock networks has typically only considered a single network at a time. Focusing on the power networks, the network impedance is a primary issue [134]. The size and placement of the decoupling capacitors are also important [40]. Different styles have been proposed for designing the power

distribution network. The most common are routed [41], mesh [42], planes [43], and interdigitated [44] structures. Each of these structures trades metal resources for performance differently. The design and analysis of power distribution networks are summarized in [72].

For clock network design, the focus is typically on power, skew, and jitter. Clock gating [52] is a widely used technique to reduce power dissipation. To achieve lower skew between two sequentially-adjacent registers, the clock network is often designed as a symmetric structure, such as an H-tree [45]. Different aspects of the clock distribution network are summarized in [149].

In [150], clock and power distribution networks are considered simultaneously to enhance the immunity to power supply noise. However, no research on utilizing the same global network for both power and clock distribution has been described in the literature. Combining these global networks is the focus of this chapter.

The following chapter is organized as follows. The high level design and related issues are discussed in Section 9.1. In Section 9.2, the strategy and related circuits for separating the power and clock signals are described. The circuit is evaluated in Section 9.3. The chapter is summarized in Section 9.4.

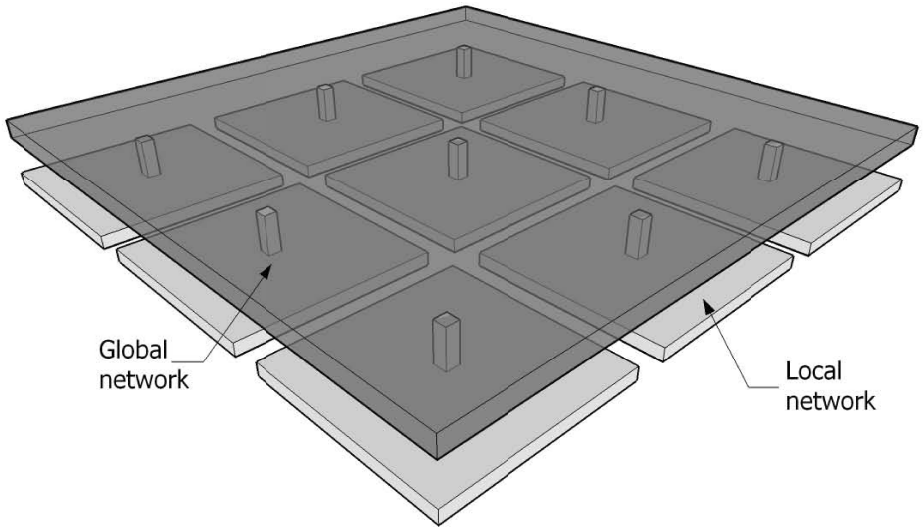


Figure 9.2: Integrating the GIPAC network into an SoC. The GIPAC network is represented by the top layer, while the local networks with separate power and clock networks are located on the bottom layer.

9.1 High Level Design

The GIPAC structure is proposed to efficiently distribute power and clock within a system-on-chip (SoC). In Fig. 9.2, the GIPAC network is depicted within an SoC, where multiple on-chip domains are characterized as a local network and the entire IC-based system as a global network. In this case, the GIPAC network distributes the integrated power and clock network over the entire circuit, while localized systems produce separate local power and clock networks. The GIPAC structure reduces the requirement for metal resources, yielding higher integration and functionality.

Noise is the primary issue in the design of a GIPAC network. The noise originates from three main sources, as illustrated in Fig. 9.3:

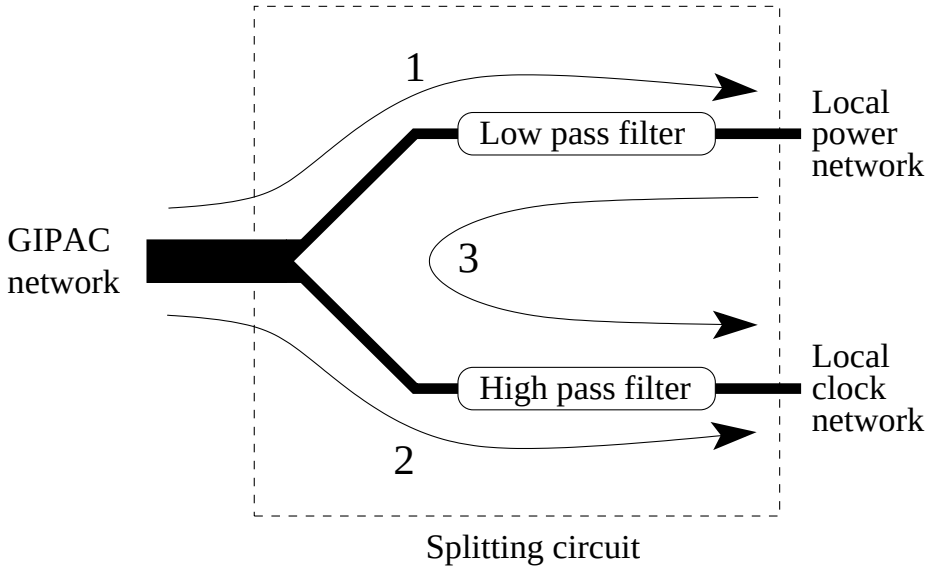


Figure 9.3: Noise is the major issue for the GIPAC network. The three noise paths are shown in the figure. The first path represents noise coupling from the GIPAC network into the local power network; the second path indicates noise coupling from the GIPAC network into the local clock network; and the third path is noise injected from the local power network into the clock network.

1. Noise from the GIPAC into the power network

2. Noise from the GIPAC into the clock network

3. Noise from the power network into the clock network

Since the GIPAC network combines the clock and power signals, a fraction of the clock signal that propagates through the low pass filter is treated as noise within the power network. Two strategies exist for reducing these noise sources, a more effective low pass filter or a higher frequency clock signal. The drawback in a more effective low pass filter is increased area. The disadvantage of providing a higher clock frequency is the requirement for higher speed and lower power.

Noise within the GIPAC also affects the clock network; however, only high frequency noise is significant since a high pass filter eliminates low frequency noise. The remaining noise produces jitter in the clock signal.

The third noise path, noise propagation from the power network into the clock network, is also considered. The circuitry powered by the power network switches at the same frequency as the clock signal, creating high frequency noise within the power network which propagates into the clock network. The different current demands from the power network also affect the clock network; a solution is therefore required to eliminate this noise mechanism.

9.2 GIPAC Splitting Circuit

9.2.1 Theoretical Background

The GIPAC splitting circuit, called here a splitter, is the primary component of the integrated power and clock network system. The function of the splitter is to separate the signals while minimizing the noise between the global and local networks. The GIPAC splitter circuit is shown in Fig. 9.4. Each RC pair behaves as a low pass filter, where the resistor is implemented as a polysilicon resistor and the capacitor as an MOS transistor. The circuit is designed assuming a 90 nm CMOS technology. The input signal supplied to the GIPAC network as a function of time t is

$$input\,(t) = A + \alpha \cdot sin\,(2\pi f_{clk} t)\,, \tag{9.1}$$

where A is the supply voltage. α and f_{clk} are, respectively, the amplitude of the signal used to generate the clock signal and the clock signal frequency. The two subsections below describe the generation of the power and clock signals, respectively.

1) Generating the power signal. The input signal propagates through the GIPAC network and arrives at the splitter, as illustrated in Fig. 9.4. The R_1C_1 filter is a low pass filter that removes the sinusoidal waveform from the input signal, maintaining only the DC portion of the signal. The output signal from this filter is V_{dd-1}. This filter can also be implemented as a higher order filter to improve the quality of the

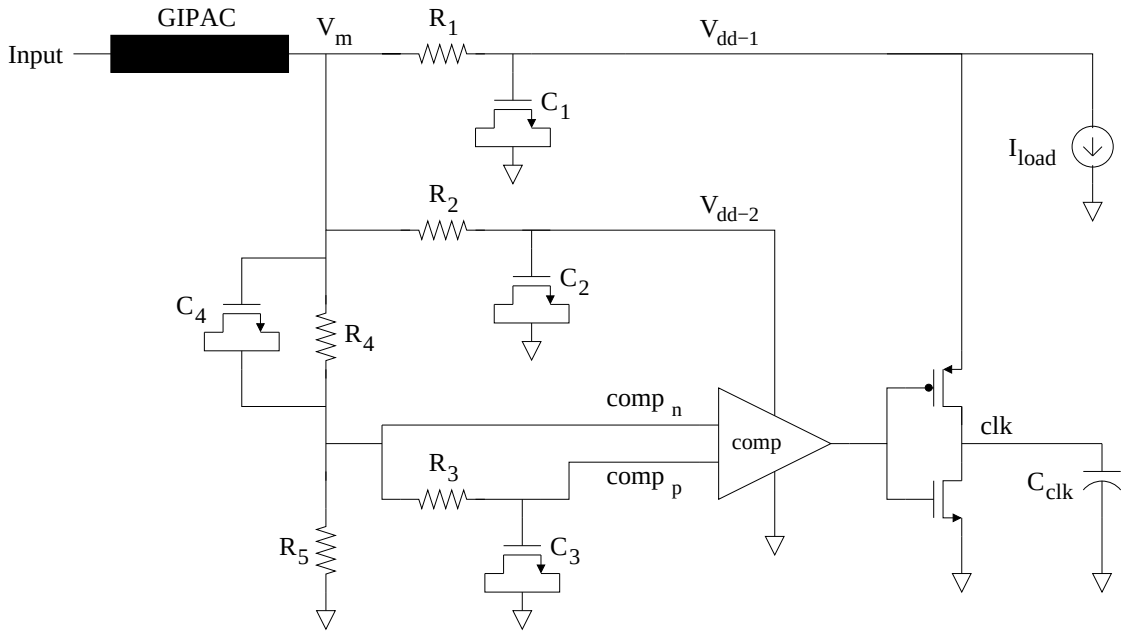


Figure 9.4: GIPAC splitter circuit. Each RC pair behaves as a low pass filter. The value of R and C are based on the noise requirements, where the resistors are implemented as polysilicon resistors and capacitors as MOS transistors. C_{clk} and I_{load} are the clock network load and current for the entire circuit, respectively.

V_{dd-1} signal. Since high current typically propagates through this filter, the R_1 component is of low value, requiring a higher value for C_1. A low value of R_1 is necessary to maintain a high voltage at V_{dd-1}, since a voltage divider is created between R_1 and the local power network.

The R_2C_2 filter is similar to the R_1C_1 low pass filter; however, due to the lower current, R_2 is significantly higher than R_1, permitting C_2 to be smaller. The cut-off frequency for the second filter is lower than the first filter, reducing the noise at the output of the second filter. The output signal of the second low pass filter is called V_{dd-2}. The R_1C_1 and R_2C_2 filters can also be implemented as a single low pass filter, reducing overall area. The switching noise on the power lines however would be injected directly into the clock signal, significantly increasing the clock jitter. Separate RC filters are therefore used to reduce the noise coupled from the power network into the clock network (the third noise path depicted in Fig. 9.3).

2) Generating the clock signal. The clock signal is produced in several stages. The DC component of the signal V_m is initially divided by two (in Fig. 9.4, labeled as a $comp_n$ signal). The $comp_p$ signal is generated by filtering the $comp_n$ signal with the R_3C_3 low pass filter. This configuration generates two signals with the same DC level; therefore, comparing (or amplifying the difference between) these two signals produces a clock signal with a 50% duty cycle. The C_4 capacitor passes an AC signal to the input of the comparator, creating a voltage divider at node $comp_n$. By

increasing C_4, the AC signal is less attenuated; albeit, requiring more area. The buffer at the output of the comparator adjusts the voltage to V_{dd-1}. This buffer can be implemented as cascaded buffers depending upon the load. The comparator utilizes a self-biased structure [151].

9.2.2 RC Filter Values

The R and C values for the low pass filters are based on the DC and AC noise requirements,

$$noise_{dc} = \frac{R \cdot I_{max}}{V_{dd}} \cdot 100, \tag{9.2}$$

where V_{dd}, I_{max}, and $noise_{dc}$ are the required power supply voltage, maximum current, and per cent of the allowed DC noise on the power network.

$$noise_{ac} = \frac{2\alpha \left| \frac{1}{1+Rj2\pi f_{clk}C} \right|}{V_{dd}} \cdot 100. \tag{9.3}$$

By increasing C, enhanced noise reduction can be achieved; however, the penalty is the area required to implement the capacitor, producing a tradeoff between noise and area. The output buffer after the comparator is a cascaded buffer structure to drive a large capacitive load.

9.3 Simulation Results

The GIPAC network and splitter are designed using a 90 nm CMOS technology, with a power signal at 1.2 volts and a clock signal frequency of 1 GHz. The simulation is evaluated with the current switching with a normal random distribution between 0 and 100 mA. A transient simulation of the input, V_{dd-1} output (to power the entire circuitry), and V_{dd-2} output (to power the clock comparator) are illustrated in Fig. 9.5. Power signal generation is accomplished by propagating the GIPAC output signal V_m through the first low pass filter. Depending upon the current requirements, the DC level of the power network is shifted due to the resistive voltage drop across the filter. The second low pass filter generating V_{dd-2} only passes a small current to the comparator, attenuating V_m. The V_{dd-1} signal fluctuates between 1.116 volts and 1.226 volts (110 mV), creating less than 10% error. The V_{dd-2} signal fluctuates between 1.189 volts and 1.201 volts (12 mV), within 1% error.

To generate the clock signal, the GIPAC output signal V_m is divided by two and filtered by the third low pass filter. The two input signals to the comparators are illustrated in Fig. 9.6(a). The $comp_n$ signal swings between 559 mV and 637 mV, while the $comp_p$ signal only swings between 596 mV and 598 mV due to the third low pass filter. The DC level of both signals is similar, producing a 50% duty cycle clock signal. After the comparator, the signal is amplified by the cascaded buffers, generating the clock signal shown in Fig. 9.6(b).

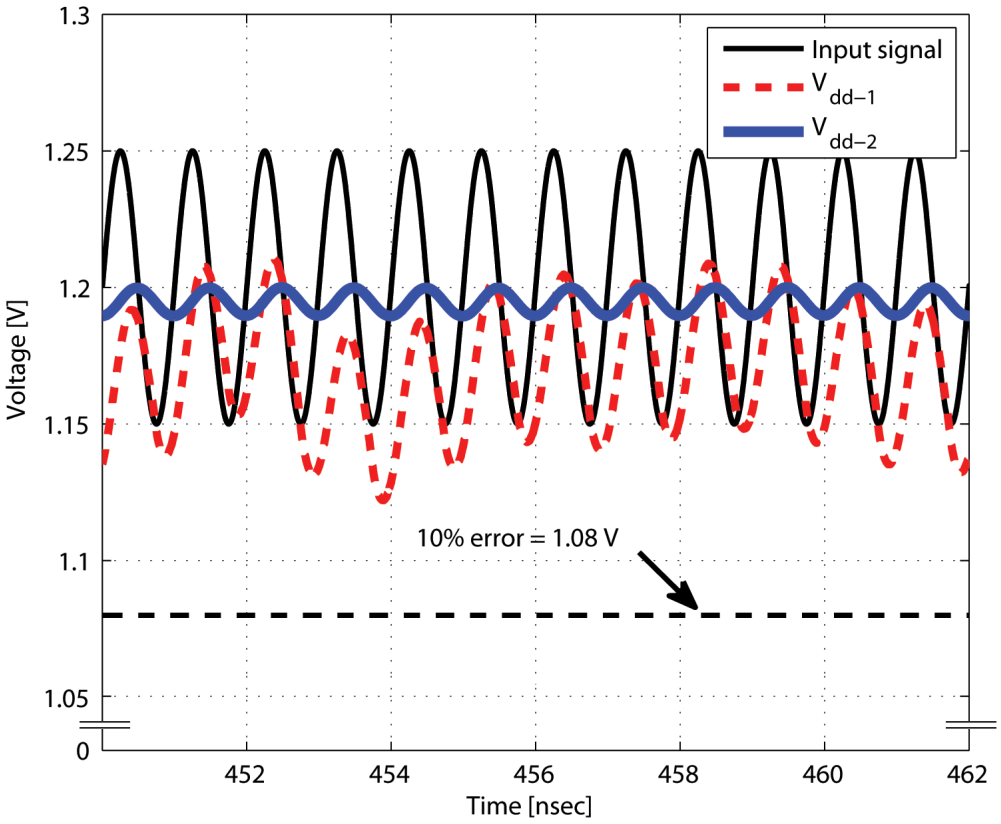


Figure 9.5: Transient simulation of the GIPAC input, output V_{dd-1}, and output V_{dd-2} signals. The ripples on the power lines are considered as noise.

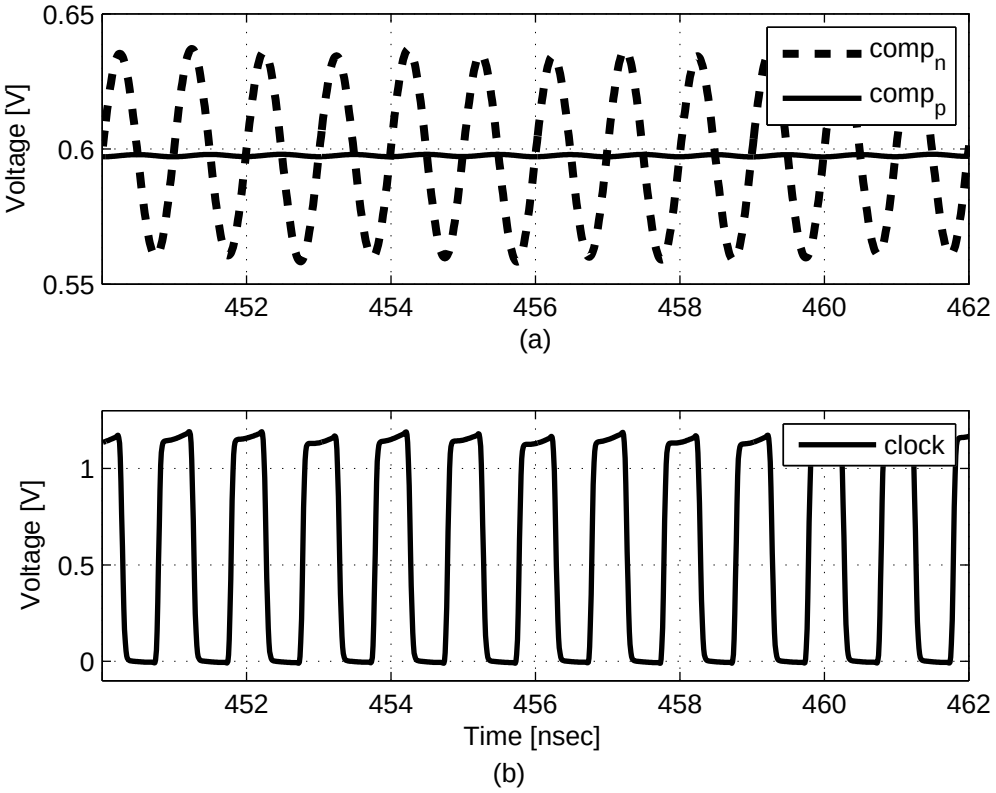


Figure 9.6: Transient simulation of the GIPAC splitter circuit. (a) Two signals at the input of the comparator exhibit the same DC level. (b) The generated clock signal drives a 1 pF capacitive load.

An eye diagram of the clock signal is depicted in Fig. 9.7. The impedance of the global ground and power networks is assumed to be equal; therefore, additional noise on the high rail in the eye diagram is a result of integrating the power and clock signals. As shown in the eye diagram, the voltage fluctuates on the high rail between 1.11 volts and 1.21 volts, producing about 8% error which results in a clock jitter of 33 ps.

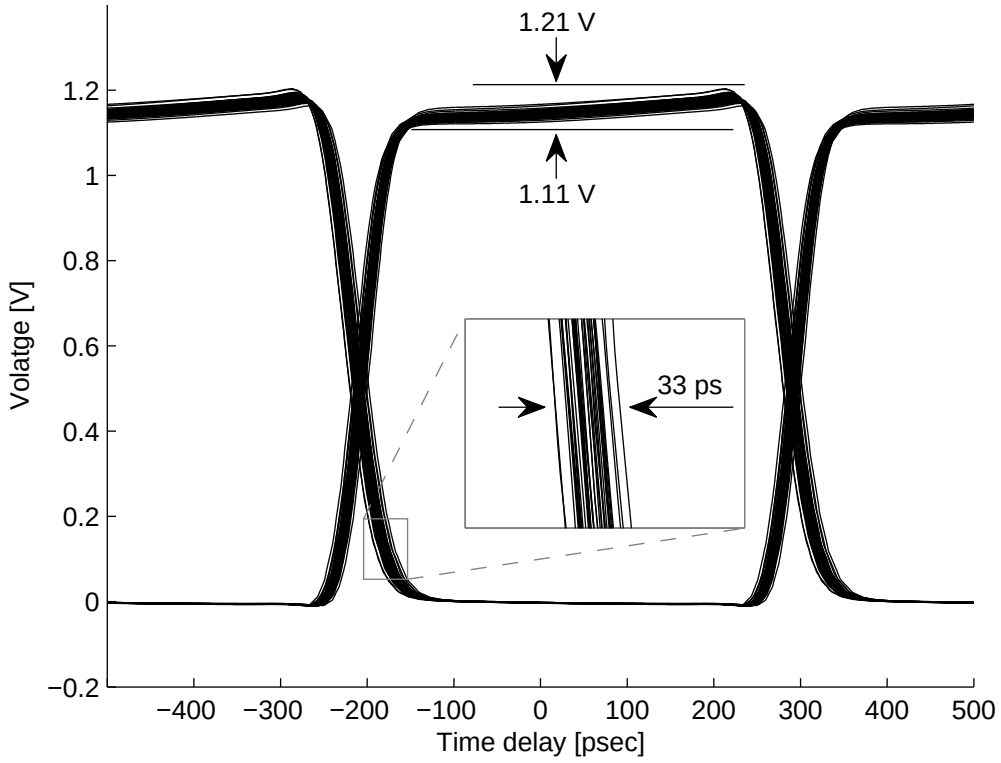


Figure 9.7: Eye diagram of the simulated clock signal. The global ground and power networks are assumed to be equal; the difference between the noise on the high and low rails is a product of the GIPAC splitter.

9.4 Summary

A general approach for combining the global power and clock networks into a single integrated network is discussed in this chapter. Replacing two global networks with one integrated network provides increased integration and functionality. Simulation results, based on a 90 nm CMOS technology, successfully demonstrate splitting the GIPAC output signal into two separate clock and power signals. Noise issues are considered and different tradeoffs are investigated for this integrated power and clock network.

Chapter 10

Future Work

The power supply network, synchronization circuitry, and signal interconnects are primary limiting factors in high complexity, high performance integrated circuits. These networks are used within every integrated circuit, functional block, and logic gate. Multiple power supplies and clock networks are commonly included on a single IC, and these networks greatly affect the performance of an integrated system. The physical design of these networks is therefore a highly important task [146].

Four different research problems are proposed in this chapter for further investigation. These research suggestions are based on concepts discussed in previous chapters of this dissertation. A GIPAC utilizing a multi-voltage and multi-clock configuration is discussed in Section 10.1. A novel power distribution scheme is proposed in Section 10.2, where the power dissipation and number of I/O pins are expected to be significantly reduced. In Section 10.3, an adaptive power distribution network is described, permitting single and multiple power network configurations to be utilized

at different times. Finally, a randomly reconfigurable power distribution network is proposed in Section 10.4, complicating the power analysis process while performing side-channel attacks; thereby enhancing the security of integrated circuits.

10.1 GIPAC for Multi-Voltage and Multi-Frequency

In modern integrated circuits, the use of multiple supply voltage levels is a common practice. Multiple supply voltages are used to reduce system-wide power dissipation, while maintaining the overall performance of an IC. A lower voltage is typically applied to those circuits which have extra time to perform the function, such as non-critical logical paths or a memory bank with a non-critical access time. The dynamic power of a logic gate driving a load capacitance C is $P = \alpha C V_{dd}^2 f$, where α, V_{dd}, and f are, respectively, the activity factor of a logic gate, supply voltage, and operating frequency. Similarly, multiple circuit blocks in an IC are clocked at different operating frequencies, maintaining the speed of the critical blocks while increasing the overall power efficiency.

The disadvantage of supplying multiple voltages or/and multiple clock signals is the requirement to design and route independent power and clock distribution networks. In Chapter 9, integrating a single power signal with a single clock signal has been proposed based on the GIPAC configuration. The integrated global signal is separated into a local power signal and a local clock signal using a splitting circuit.

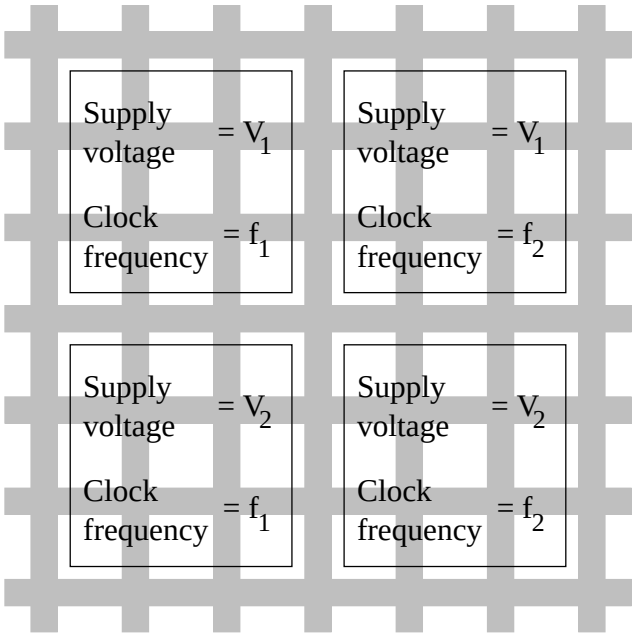


Figure 10.1: GIPAC network for multi-voltage and multi-frequency distribution. Four different circuit domains receive a different power supply voltage and clock frequency.

This configuration reduces usage of the higher metal layers.

A similar approach is proposed for delivering multiple voltage levels and multiple clock signal frequencies with a single global network, as illustrated in Fig. 10.1. The local power supply voltage is adjusted with DC-to-DC voltage converters. For those applications where the multiple voltages are similar in magnitude, a low dropout regulator can be used. Noise on the power supply network can also be reduced by low dropout regulators which exhibit a high power line rejection ratio for noise suppression. Multiple band-pass filters only pass a signal within a target frequency, permitting multiple clock signals operating at different frequencies to be generated.

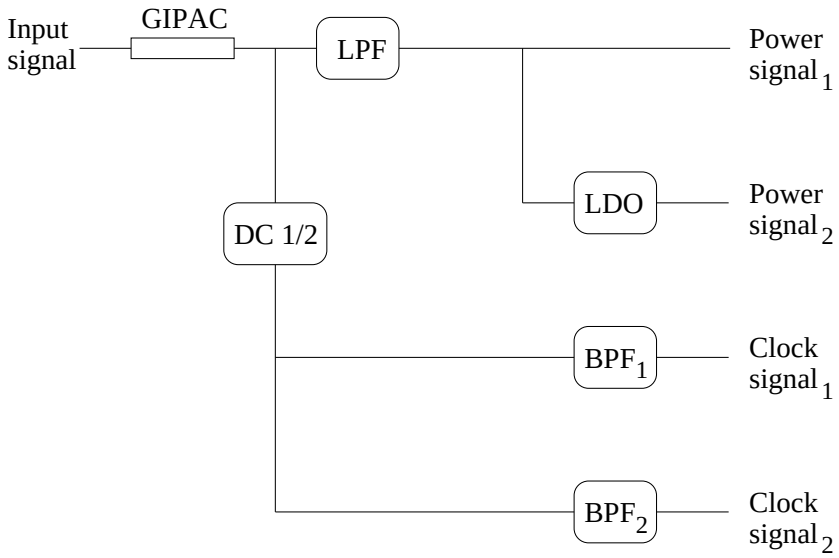


Figure 10.2: Schematic of GIPAC power/clock signal separation scheme. A low pass filter and low dropout DC-to-DC converter are used to produce the power signals. Two different band pass filters shape the signal and generate a clock signal at two different clock frequencies.

A schematic of a splitting circuit for the multi-voltage and multi-frequency GIPAC configuration is depicted in Fig. 10.2. The circuit delivers two different voltage levels and clock signals with two different frequencies. The time domain representation of the input signal is

$$In\left(t\right) = A + sin\left(2\pi f_1 t\right) + sin\left(2\pi f_2 t\right), \tag{10.1}$$

where A is a common (or highest) DC voltage level. f_1 and f_2 are the two clock frequencies. The band pass filters are designed to eliminate noise from the neighboring clock signals. The low dropout regulators control the voltage level as well as suppress noise ripple on the common power rail.

10.2 On-Chip AC Power Signal with Boosted Voltage

The design of an efficient power delivery and distribution network is of high priority due to IR and $L\frac{di}{dt}$ noise. The noise on the power network can greatly affect the performance of an integrated circuit. The reduced voltage between the power and ground rails significantly increases the delay of the logic gates.

The need for high currents within integrated circuits is an important constraint. These high current demands require a large number of I/O pins to be dedicated for power and ground. About half of the available pins within a package is typically dedicated to power and ground.

In a typical power delivery scheme, I^2R power losses are significant due to the high current passing through the power and ground lines and supplied to the logic gates. To reduce power losses on the power and ground lines, a high voltage power delivery scheme is proposed. This scheme is illustrated in Fig. 10.3. The power lines on a board, package, and integrated circuit carry the signal at a voltage level much higher than required for the on-chip logic gates. The current level within these lines are thereby reduced, significantly lowering the I^2R power losses on these lines. A transformer can be used to convert the high voltage/low current signal (noted as signal_1 in Fig. 10.3) into a low voltage/high current signal (noted as signal_2 in

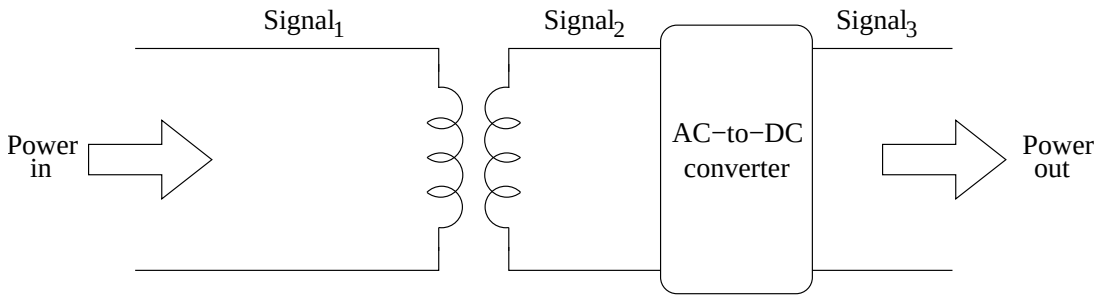


Figure 10.3: Proposed power delivery scheme. The high voltage/low current signal (noted as $signal_1$) is supplied to the integrated circuit. The on-chip transformer generates a low voltage/high current signal, maintaining the same power. This AC signal is converted to DC and supplied to the logic gates.

Fig. 10.3) close to the load. An AC signal is required to perform this transformation; an AC-to-DC converter is therefore used to supply a DC signal (noted as $signal_3$ in Fig. 10.3) to the logic gates. The power supplied to the circuits is equal to the input power, neglecting any power losses within the lines, transformer, and AC-to-DC converter.

Due to the low input current ($signal_1$ in Fig. 10.3), the number of I/O pins is significantly reduced. Losses within these power and ground lines is also reduced due to the smaller current propagated within the lines. The primary disadvantage is the on-chip transformer and AC-to-DC converter. The losses of these circuits need to be low to decrease the overall power dissipation. The power efficiency and physical area of the transformer can be a bottleneck for this proposed power delivery scheme. This requirement can be addressed using 3-D integration, where the transformer and AC-to-DC converter use a dedicated plane within the 3-D stack. The low voltage/high

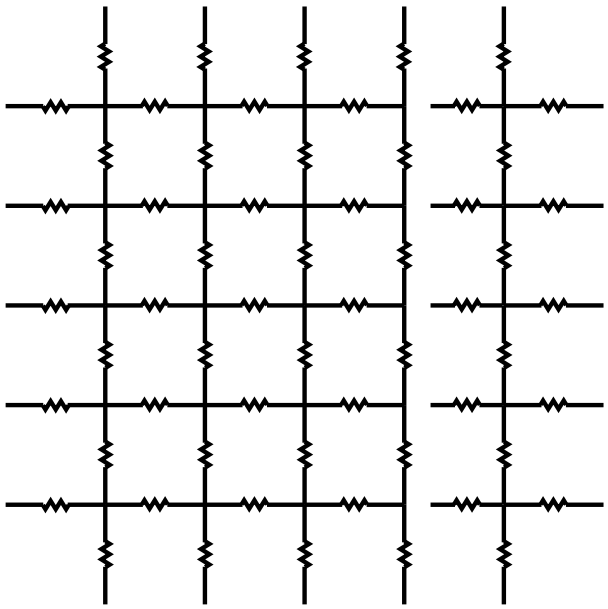
current is delivered to the other planes using through silicon vias.

10.3 Adaptive Power Distribution Network

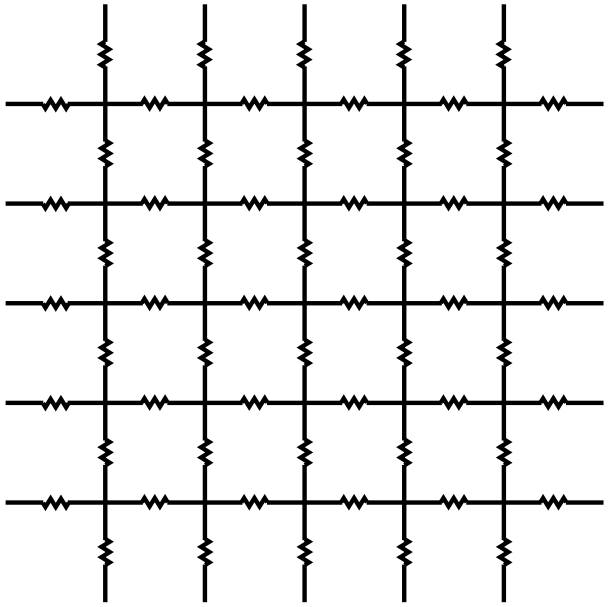
A mesh structured power distribution network is discussed in Chapter 8, where links are removed to optimize the voltage drops and improve the performance. A temporally adaptive version of this approach can be utilized for systems requiring lower impedance and noise at different times.

Two power networks with two different configurations are illustrated in Fig. 10.4. A circuit with an internal calibration block is used to present the advantages of both network configurations. The configuration depicted in Fig. 10.4(a) is used when two separate power networks are required. One network powers the main circuitry, while the second network powers the calibration block. The noise produced from switching the main circuitry is isolated, ensuring the power network of a calibration block remains noiseless. Due to the clean power supply, the performance of the calibration block is enhanced.

The calibration process is however typically performed infrequently. When the calibration block is not operating, the power distribution network can be configured as illustrated in Fig. 10.4(b). The noise on the main power network is reduced since a lower network impedance is achieved and a higher number of power supplies is connected to the main power distribution network. The resulting performance of the



(a)



(b)

Figure 10.4: Two configurations of a mesh structured power distribution network, (a) two networks are separated to reduce the noise coupled between the networks, and (b) both networks are connected to reduce the impedance of the overall power network.

main circuitry is enhanced with this power network configuration.

The switches used to connect or separate the power networks are large in area to reduce the resistive losses within these switches. This approach is exemplified by the power gating technique [152]. The difference is that in power gating the circuits are decoupled from the network, while with the proposed approach, the decoupling is between the power networks. The proposed approach can also be applied to multicore systems, permitting the power networks to either be decoupled between the cores or the power networks to be connected based on the number of operating cores.

10.4 Reduction of Side-Channel Attacks

The increasing demand for secure applications pushes for advancements in secure integrated circuits. Secure protocols (or secure keys) are embedded into these integrated circuits. These circuits are used in devices raging from simple hotel doorkey locks to highly classified governmental data/voice/text transmission. To decrypt the embedded data, side-channel attacks are frequently employed. Side-channel attacks can be grouped into three main categories [153, 154]:

- Timing attacks
- Power consumption attacks
- Differential fault analysis attacks

Timing attacks are based on measuring the time required for the data to propagate within the unit. Based on this timing information, the secret key may be discovered by correlating multiple input and timing data pairs. Power consumption attacks are based on measuring the power consumption for various inputs. The information in the time domain of the power signal can be used to decrypt the secure key. Differential fault analysis attacks are based on investigating the outputs while intentionally perturbing the system.

The focus of the proposed technique is to reduce the sensitivity of an IC to side-channel attacks based on measuring the power consumption. These attacks are typically performed by probing the power or ground signal at the input ports of an integrated circuit. A time domain signal is used to analyze the work load. The power and ground signals are significantly affected by the on-chip power distribution network. During operation of the integrated circuit, electromagnetic (EM) emissions can also be used to extract the secret signature. The EM emissions are noise waveforms produced by a current propagating within the power distribution network. The time domain of the power signal and EM emissions are therefore a good representation of the behavior of an on-chip power distribution network.

The impedance of the power distribution network is proposed to be changed during operation of the integrated circuit to reduce these side-channel attacks when performing a power analysis. The current from a specific block can be redirected to multiple

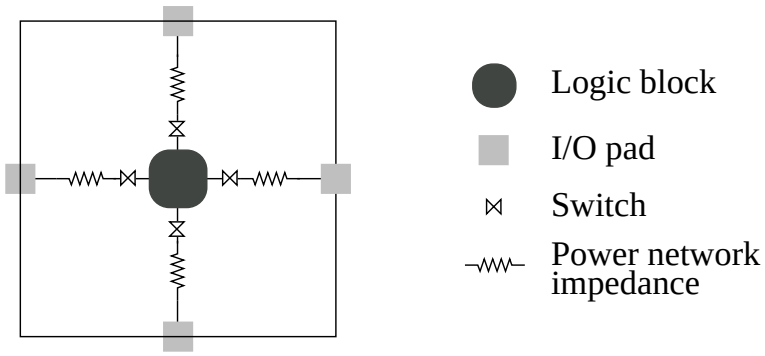


Figure 10.5: A circuit is connected to four different I/O pads using four different power networks. The current from the block to the I/Os can be directed in 15 different combinations (2^4 options except for the option where all of the switches are disconnected). Redirecting the current flow will significantly complicate the EM emission and power analysis detection process.

I/Os, as illustrated in Fig. 10.5, deceiving the power analysis detection process.

To further complicate the power analysis detection process, a mesh structured power network is proposed with multiple switches between the links. This power network structure is depicted in Fig. 10.6. The switches can be controlled by a pseudorandom number generator, opening and closing these power switches. The current flow within the power network would be randomly distributed, complicating the power analysis detection process and resulting in enhanced security of the information embedded within the IC. Since EM emissions are also directly related to the current flow within the power network, protecting an IC to side-channel attacks using EM analysis will also be significantly improved. This methodology can be developed without customizing the metal lines, switches, and control signals, requiring only a single customized pseudorandom number generator.

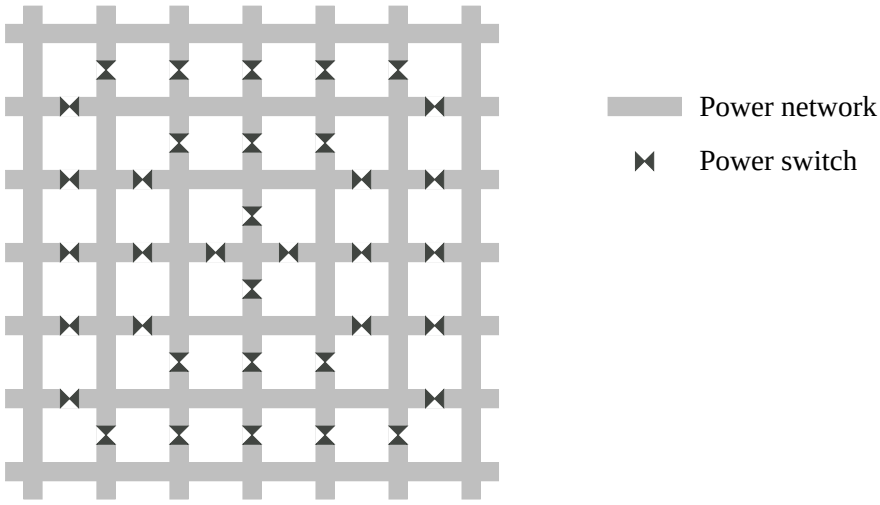


Figure 10.6: Power distribution network with multiple power switches to randomize the current within a network.

Chapter 11

Conclusions

The technology behind integrated circuits is growing rapidly with greater number of devices integrated on the same die. These large number of devices require highly complicated networks to manage and support efficient operation. Resources, such as metal, power, and area, are however limited; these resources must therefore be efficiently utilized.

With further advancements in semiconductor technology, the complexity of integrated circuits continues to increase. Multiple structures, in the past analyzed independently, now need to be merged to more accurately analyze performance. This integrated approach supports the development of design and optimization methodologies that manage noise, power, area, and other resources. The power delivery network, clock distribution network, and substrate need to be modeled as one integrated system to more accurately characterize the noise signature and circuit performance.

The increase in the number of metal layers within an integrated circuit has not

kept up with device scaling, creating challenges in global signaling, synchronization, and power delivery. The objective of this dissertation is to address these challenges by developing closed-form models to analyze these highly complicated global networks, thereby reducing computational complexity. An additional objective is the development of design methodologies to efficiently utilize the available resources, such as area, metal, and power. The performance of integrated circuits is highly affected by the power delivery system; the focus of this dissertation is therefore on on-chip power delivery systems and related design methodologies.

An effective impedance model of a monolithic substrate has been developed, achieving high accuracy in estimating power/ground noise characteristics. A methodology for simultaneously inserting shields and repeaters is described, permitting multiple resources used in global signal interconnects to be optimized. A closed-form model of the self- and mutual inductance of an interdigitated power and ground distribution network is described, providing less than 5% error for a typical power distribution network. The optimal width of the metal lines that minimizes the impedance of the power distribution network is determined, significantly enhancing the performance of an integrated circuit. A design methodology is also described for a multi-layer power distribution network, achieving enhanced reliability by equalizing the current density over multiple metal layers. Furthermore, a novel link breaking methodology for a mesh structured power distribution network is introduced, reducing coupling noise while

improving the maximum operating frequency, on average, by 12%. Finally, a globally integrated power and clock distribution network has been presented which utilizes a single network to distribute both global signals; thereby reducing the metal requirement. These closed-form models and methodologies described within this dissertation enhance the design, analysis, and optimization of advanced integrated circuits.

Bibliography

[1] J. J. Thomson, "Cathode Rays," *Philosophical Magazine Series 5*, Vol. 44, pp. 293-316, 1897.

[2] "Nobel Prize in Physics," available online at http://nobelprize.org.

[3] American Institute of Physics, "The Discovery of the Electron," available online at http://www.aip.org/history/electron.

[4] W. Brittain, "Lab Notebook," December 24, 1947.

[5] R. N. Noyce, "Semiconductor Device-and-Lead Structure," U. S. Patent No. 2,981,877, April 1961.

[6] G. Moore, "Cramming More Components Onto Integrated Circuits," *Electronics Magazine*, Vol. 38, No. 8, pp. 114-117, April 1965.

[7] International Technology Roadmap for Semiconductors, 2007, available online at http://public.itrs.net.

[8] Y. Shin and H.-O. Kim, "Analysis of Power Consumption in VLSI Global Interconnects," *Proceedings of the IEEE International Symposium on Circuit and Systems*, pp. 4713-4716, May 2005.

[9] G. Chen and E. G. Friedman, "Low Power Repeaters Driving *RC* and *RLC* Interconnects with Delay and Bandwidth Constraints," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 2, pp. 161-172, February 2006.

[10] C.-C. A. Chena, L.-S. Shua, and S.-R. Leeb, "Mechano-Chemical Polishing of Silicon Wafers," *Journal of Materials Processing Technology*, Vol. 140, No. 1-3, pp. 373-378, August 2003.

[11] J. W. McPherson, "Reliability Trends with Advanced CMOS Scaling and the Implications for Design," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 405-412, September 2007.

[12] D. L. Goodman, "Effect of Wafer Bow on Electrostatic Chucking and Back Side Gas Cooling," *Journal of Applied Physics*, Vol. 104, No. 12, pp. 124902-124902-8, December 2008.

[13] S. Ghosh and K. Roy, "Exploring High-Speed Low-Power Hybrid Arithmetic Units at Scaled Supply and Adaptive Clock-Stretching," *Proceedings of the IEEE/ACM Asia and South Pacific Design Automation Conference*, pp. 635-640, 2008.

[14] F. Gray, "Pulse Code Communication," U.S. Patent No. 2,632,058, March 1953.

[15] S. Golomb, "Runlength Encodings," *IEEE Transactions on Information Theory*, Vol. 12, pp. 399-401, July 1966.

[16] X. Liu, M. C. Papaefthymiou, and E. G. Friedman, "Retiming and Clock Scheduling for Digital Circuit Optimization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 21, No. 2, pp. 184-203, February 2002.

[17] H. F. Jordan, "Experience with Pipelined Multiple Instruction Streams," *Proceedings of the IEEE*, Vol. 72, No. 1, pp. 113-123, January 1984.

[18] J. T. Kao and A. P. Chandrakasan, "Dual-Threshold Voltage Techniques for Low-Power Digital Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 35, No. 7, pp. 1009-1018, July 2000.

[19] Y. Ye, S. Borkar, and V. De, "A New Technique for Standby Leakage Reduction in High-Performance Circuits," *Proceedings of the Symposium on VLSI Circuits*, pp. 40-41, June 1998.

[20] J. Pincus, "Transistor Sizing," University of California, Berkeley, Technical Report No. UCB/CSD-86-285, February 1986.

[21] B. S. Carlson and C. Y. R. Chen, "Effects of Transistor Reordering on the Performance of MOS Digital Circuits," *Proceedings of the Midwest Symposium on Circuits and Systems*, Vol. 1, pp. 121-124, August 1992.

[22] T. Uehara and W. M. Vancleemput, "Optimal Layout of CMOS Functional Arrays," *IEEE Transactions on Computers*, Vol. C-30, No. 5, pp. 305-312, May 1981.

[23] B. S. Cherkauer and E. G. Friedman, "Channel Width Tapering of Serially Connected MOSFETs with Emphasis on Power Dissipation," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 2, No. 1, pp. 100-114, March 1994.

[24] 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.

[25] V. Kursun, R. M. Secareanu, and E. G. Friedman, "CMOS Voltage Interface Circuits for Low Power Systems," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 3.667-3.670, May 2002.

[26] H. Zhang and J. Rabaey, "Low-Swing Interconnect Interface Circuits," *Proceedings of the IEEE International Symposium on Low Power Electronics and Design*, pp. 161-166, August 1998.

[27] H. C. Lin and L. W. Linholm, "An Optimized Output Stage for MOS Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 10, No. 2, pp. 106-109, April 1975.

[28] H. J. M. Veendrick, "Short-Circuit Dissipation of Static CMOS Circuitry and Its Impact on the Design of Buffer Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 19, No. 4, pp. 468-473, August 1984.

[29] H. B. Bakoglu and J. D. Meindl, "Optimal Interconnection Circuits for VLSI," *IEEE Transactions on Electron Devices*, Vol. 32, No. 5, pp. 903-909, May 1985.

[30] J. Zhang and E. G. Friedman, "Crosstalk Modeling for Coupled *RLC* Interconnects with Application to Shield Insertion," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 6, pp. 641-646, June 2006.

[31] Y. Massoud, J. Kawa, D. MacMillen, and J. White, "Modeling and Analysis of Differential Signaling for Minimizing Inductive Crosstalk," *Proceedings of the ACM/IEEE Design Automation Conference*, pp. 804-809, June 2001.

[32] A. Carusone, K. Farzan, and D. A. Johns, "Differential Signaling with a Reduced Number of Signal Paths," *IEEE Transactions on Circuits and Systems II: Analog and Digital Processing*, Vol. 48, No. 3, pp. 294-300, March 2001.

[33] R. M. Secareanu and E. G. Friedman, "Transparent Repeaters," *Proceedings of the IEEE Great Lakes Symposium on VLSI*, pp. 63-66, March 2000.

[34] A. Nalamalpu, S. Srinivasan, and W. Burleson, "Boosters for Driving Long On-Chip Interconnects: Design Issues, Interconnect Synthesis, and Comparison with Repeaters," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 21, pp. 50-62, January 2002.

[35] M. A. El-Moursy and E. G. Friedman, "Wire Shaping of *RLC* Interconnects," *Integration, the VLSI Journal*, Vol. 40, No. 4, pp. 461-472, July 2007.

[36] K. Hirose and H. Yassura, "A Bus Delay Reduction Technique Considering Crosstalk," *Proceedings of the IEEE Design, Automation, and Test in Europe Conference and Exhibition*, pp. 441-445, March 2000.

[37] B. Soudan, "Reducing Mutual Inductance of Wide Signal Buses Trough Swizzling," *Proceedings of the IEEE Conference on Electronics, Circuits, and Systems*, Vol. 2, pp. 870-873, December 2003.

[38] P. Gupta and A. Kahng, "Wire Swizzling to Reduce Delay Uncertainty Due to Capacitive Coupling," *Proceedings of the IEEE International Conference on VLSI Design*, pp. 431-436, January 2004.

[39] V. Kursun and E. G. Friedman, *Multi-Voltage CMOS Circuit Design*. John Wiley & Sons, 2006.

[40] M. Popovich, M. Sotman, A. Kolodny, and E. G. Friedman, "Effective Radii of On-Chip Decoupling Capacitors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 16, No. 7, pp. 894-907, July 2008.

[41] D. Blaauw, R. Panda, and R. Chaudhry, "Chapter 24: Design and Analysis of Power Distribution Networks," *Design of High-Performance Microprocessor Circuits*, A. Chandrakasan, W. J. Bowhill, and F. Fox, Eds., pp. 499-522. Wiley - IEEE Press, 2000.

[42] A. Dalal, L. Lev, and S. Mitra, "Design of an Efficient Power Distribution Network for the UltraSPARC-I Microprocessor," *Proceedings of the IEEE International Conference on Computer Design*, pp. 118-123, October 1995.

[43] L. Cao and J. P. Krusius, "A New Power Distribution Strategy for Area Array Bonded ICs and Packages of Future Deep Sub-Micron ULSI," *Proceedings of the IEEE Electron Components and Technology Conference*, pp. 915-920, June 1995.

[44] L.-R. Zheng and H. Tenhunen, "Effective Power and Ground Distribution Scheme for Deep Submicron High Speed VLSI Circuits," *Proceedings of the IEEE International Symposium on Circuit and Systems*, Vol. I, pp. 537-540, May 1999.

[45] P. J. Restle and A. Deutsch, "Designing the Best Clock Distribution Network," *Proceedings of the IEEE Symposium on Very Large Scale Integration (VLSI) Circuits*, pp. 2-5, June 1998.

[46] S. C. Chan, K. L. Shepard, and P. J. Restle, "Design of Resonant Global Clock Distribution," *Proceedings of the IEEE International Conference on Computer Design*, pp. 248-253, October 2003.

[47] P. J. Restle, *et al.*, "A Clock Distribution Network for Microprocessors," *IEEE Journal of Solid-State Circuits*, Vol. 35, No. 5, pp. 792-799, May 2001.

[48] G. Venkataraman, *et al.*, "Practical Techniques to Reduce Skew and Its Variations in Buffered Clock Networks," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 592-596, November 2005.

[49] F. O. Mahony, C. P. Yue, M. A. Horowitz, and S. S. Wong, "A 10-GHz Global Clock Distribution Using Coupled Standing-Wave Oscillators," *IEEE Journal of Solid-State Circuits*, Vol. 38, No. 11, pp. 1813-1820, November 2003.

[50] J. Wood, T. C. Edwards, and S. Lipa, "Rotary Traveling-Wave Oscillator Arrays: A New Clock Technology," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 11, pp. 1654-1665, November 2001.

[51] S. C. Chan, K. L. Shepard, and P. J. Restle, "Distributed Differential Oscillators for Global Clock Networks," *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 9, pp. 2083-2094, September 2006.

[52] G. E. Tellez, A. Farrahi, and M. Sarrafzadeh, "Activity-Driven Clock Design for Low Power Circuits," *Proceedings of the IEEE International Conference on Computer-Aided Design*, pp. 62-65, November 1995.

[53] C. Soens *et al.*, "Modeling of Substrate Noise Generation, Isolation, and Impact for an LC-VCO and a Digital Modem on a Lightly-Doped Substrate," *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 9, pp. 2040-2051, September 2006.

[54] R. M. Secareanu, *et al.*, "Placement of Substrate Contacts to Minimize Substrate Noise in Mixed-Signal Integrated Circuits," *Analog Integrated Circuits and Signal Processing*, Vol. 28, No. 3, pp. 253-264, Semtember 2001.

[55] A. Afzali-Kusha, M. Nagata, N. K. Verghese, and D. J. Allstot, "Substrate Noise Coupling in SoC Design: Modeling, Avoidance, and Validation," *Proceedings of the IEEE*, Vol. 94, No. 12, pp. 2109-2138, December 2006.

[56] S.-M. Kang and Y. Leblebici, *CMOS Digital Integrated Circuits*. McGraw-Hill, 2003.

[57] V. Adler and E. G. Friedman, "Repeater Design to Reduce Delay and Power in Resistive Interconnect," *IEEE Transactions on Circuits and Systems II: Analog and Digital Processing*, Vol. 45, No. 5, pp. 607-616, May 1998.

[58] Berkeley Predictive Technology Model, available online at http://www.device.eecs.berkeley.edu/ptm.

[59] A. Vittal and M. Sadowska, "Crosstalk Reduction in VLSI," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 16, No. 3, pp. 290-298, March 1997.

[60] H. Kaul, D. Sylvester, and D. Blauw, "Active Shields: A New Approach to Shielding Global Wires," *Proceedings of the ACM/IEEE Great Lakes Symposium on VLSI*, pp. 112-117, March 2002.

[61] M. Ghoneima and Y. Ismail, "Formal Derivation of Optimal Active Shielding for Low-Power On-Chip Buses," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 800-807, November 2004.

[62] R. Arunachalam, E. Acar, and S. R. Nassif, "Optimal Shielding/Spacing Metrics for Low Power Design," *Proceedings of the IEEE Computer Society Annual Symposium on VLSI*, pp. 167-172, February 2003.

[63] S. Köse, E. Salman, and E. G. Friedman, "Shielding Methodologies in the Presence of Power/Ground Noise," *Proceedings of the IEEE International Symposium on Circuit and Systems*, pp. 2277-2280, May 2009.

[64] P. Bai, *et al.*, "A 65nm Logic Technology Featuring 35nm Gate Lengths, Enhanced Channel Strain, 8 Cu Interconnect Layers, Low-k ILD and 0.57 μm^2 SRAM Cell," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 657-660, December 2004.

[65] L. D. Smith, "Decoupling Capacitor Calculations for CMOS Circuits," *Procedding of the IEEE Topical Meeting on Electrical Performance of Electronic Packaging*, pp. 101-105, November 1994.

[66] H. H. Chen and S. E. Schuster, "On-Chip Decoupling Capacitor Optimization for High-Performance VLSI Design," *Proceedings of the IEEE International Symposium on VLSI Technology, Systems, and Applications*, pp. 99-103, May 1995.

[67] H. Su, S. S. Sapatnekar, and S. R. Nassif, "Optimal Decoupling Capacitor Sizing and Placement for Standard-Cell Layout Designs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 22, No. 4, pp. 428-436, April 2003.

[68] M. D. Pant, P. Pant, and D. S. Wills, "On-Chip Decoupling Capacitor Optimization using Architectural Level Prediction," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 10, No. 6, pp. 319-326, June 2002.

[69] 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.

[70] S. Bendhia, M. Ramdani, and E. Sicard, *EMC of ICs: Techniques for Low Emission and Susceptibility.* Springer-Verlag, 2005.

[71] L. C. Tsai, "A 1 GHz PA-RISC Processor," *Proceedings of the IEEE International Solid-State Circuits Conference*, pp. 322-323, February 2001.

[72] M. Popovich, A. V. Mezhiba, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors.* Springer, 2008.

[73] M. K. Gowan, L. L. Biro, and D. Jackson, "Power Considerations in the Design of the Alpha 21264 Microprocessor," *Proceedings of the IEEE/ACM Design Automation Conference*, pp. 726-731, June 1998.

[74] J. Holley, "Vanguard Mathematician George Dantzig Dies," *The Washington Post*, May 19, 2005, available online at http://www.washingtonpost.com.

[75] M. Avriel, *Nonlinear Programming: Analysis and Methods.* Dover Publications, 2003.

[76] K.-H. Elster, *Modern Mathematical Methods of Optimization.* John Wiley & Sons, 1993.

[77] J. Nocedal and S. Wright, *Numerical Optimization, 2^{nd} Edition.* Springer, 2006.

[78] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, 2004.

[79] H. Robbins and S. Monro, "A Stochastic Approximation Method," *The Annals of Mathematical Statistics*, Vol. 22, No. 3, pp. 400-407, March 1951.

[80] N. Metropolis and S. Ulam, "The Monte Carlo Method," *Journal of the American Statistical Association*, Vol. 44, No. 247, pp. 335-341, September 1949.

[81] R. E. Steuer, *Multiple Criteria Optimization: Theory, Computation, and Application.* John Wiley & Sons, 1986.

[82] M. Ehrgott, *Multicriteria Optimization, 2^{nd} Edition.* Springer, 2005.

[83] J. S. Choi and K. Lee, "Design of CMOS Tapered Buffer for Minimum Power-Delay Product," *IEEE Journal of Solid-State Circuits*, Vol. 29, No. 9, pp. 1142-1145, September 1994.

[84] J. H. Mathews and K. K. Fink, *Numerical Methods Using Matlab, 4^{th} Edition.* Prentice Hall, 2004.

[85] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journam*, Vol. 7, No. 4, pp. 308-313, 1965.

[86] N. Jouppi, "Timing Analysis for NMOS VLSI," *Proceedings of the IEEE Design Automation Conference*, pp. 411-418, June 1983.

[87] S. Trimberger, "Automated Performance Optimization of Custom Integrated Circuits," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 194-197, May 1983.

[88] M. D. Matson, "*Macromodeling and Optimization of Digital MOS VLSI Circuits*," Ph.D. Dissertation, Massachusetts Institute of Technology, January 1985.

[89] Y. I. Ismail, E. G. Friedman, and J. L. Neves, "Figures of Merit to Characterize the Importance of On-Chip Inductance," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 7, No. 4, pp. 442-449, December 1999.

[90] D. Khalil and Y. Ismail, "Approximate Frequency Response Models for RLC Power Grids," *Proceedings of the IEEE International Symposium on Circuit and Systems*, pp. 3784-3787, May 2007.

[91] A. V. Mezhiba and E. G. Friedman, "Inductance 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.

[92] F. Grover, *Inductance Calculation: Working Formulas and Tables.* Dover Publications, 1962.

[93] 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.

[94] D. K. Su, M. J. Loinaz, S. Masui, and B. A. Wooley, "Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 28, No. 4, pp. 420-430, April 1993.

[95] M. Ingels and M. S. J. Steyaert, "Design Strategies and Decoupling Techniques for Reducing the Effects of Electrical Interference in Mixed-Mode IC's," *IEEE Journal of Solid-State Circuits*, Vol. 32, No. 7, pp. 1136-1141, July 1997.

[96] A. Afzali-Kusha, M. Nagata, N. K. Verghese, and D. J. Allstot, "Substrate Noise Coupling in SoC Design: Modeling, Avoidance, and Validation," *Proceedings of the IEEE*, Vol. 94, No. 12, pp. 2109-2138, December 2006.

[97] J. P. Colinge, *Silicon-On-Insulator Technology: Materials to VLSI.* Kluwer Academic Publishers, 1997.

[98] A. J. van Genderen, N. P. van der Meijs, and T. Smedes, "Fast Computation of Substrate Resistances in Large Circuits," *Proceedings of the European Design and Test Conference*, pp. 560-565, March 1996.

[99] D. Ozis, T. Fiez, and K. Mayaram, "A Comprehensive Geometry-Dependent Macromodel for Substrate Noise Coupling in Heavily Doped CMOS Processes," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 497-500, August 2002.

[100] H. Lan, *et al.*, "Synthesized Compact Models and Experimental Verifications for Substrate Noise Coupling in Mixed-Signal ICs," *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 8, pp. 1817-1829, August 2006.

[101] K. Fukahori and P. R. Gray, "Computer Simulation of Integrated Circuits in the Presence of Electrothermal Interaction," *IEEE Journal of Solid-State Circuits*, Vol. 11, No. 6, pp. 834-846, December 1976.

[102] B. R. Stanisic, *et al.*, "Addressing Substrate Coupling in Mixed-Mode IC's: Simulation and Power Distribution Synthesis," *IEEE Journal of Solid-State Circuits*, Vol. 29, No. 3, pp. 226-238, March 1994.

[103] T. Smedes, N. P. van der Meijs, and A. J. van Genderen, "Boundary Element Methods For 3D Capacitance and Substrate Resistance Calculations in Inhomogeneous Media in a VLSI Layout Verification Package," *Advances Engineering Software*, Vol. 20, No. 1, pp. 19-27, January 1994.

[104] R. Gharpurey and R. G. Meyer, "Modeling and Analysis of Substrate Coupling in Integrated Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 31, No. 3, pp. 344-352, March 1996.

[105] E. Salman, *et al.*, "Methodology for Efficient Substrate Noise Analysis in Large-Scale Mixed-Signal Circuits," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 17, No. 10, pp. 1405-1418, October 2009.

[106] A. Samavedam, A. Sadate, K. Mayaram, and T. S. Fiez, "A Scalable Substrate Noise Coupling Model for Design of Mixed-Signal IC's," *IEEE Journal of Solid-State Circuits*, Vol. 35, No. 6, pp. 895-904, June 2000.

[107] H. Su, K. H. Gala, and S. S. Sapatnekar, "Analysis and Optimization of Structured Power/Ground Networks," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 22, No. 11, pp. 1533-1544, November 2003.

[108] X. Aragones and A. Rubio, "Experimental Comparison of Substrate Noise Coupling Using Different Wafer Types," *IEEE Journal of Solid-State Circuits*, Vol. 34, No. 10, pp. 1405-1409, October 1999.

[109] R. Shreeve, T. S. Fiez, and K. Mayaram, "A Physical and Analytical Model for Substrate Noise Coupling Analysis," *Proceedings of the IEEE International Symposium on Circuits and Systems*, pp. 157-160, May 2004.

[110] E. Charbon, R. Gharpurey, P. Miliozzi, R. G. Meyer, and A. Sangiovanni-Vincentelli, *Substrate Noise*. Kluwer Academic Publishers, 2001.

[111] S. Ramanujan, "Modular Equations and Approximations to π," *Quarterly Journal of Pure and Applied Mathematics*, Vol. 45, pp. 350-372, 1914.

[112] "Assura RCX^{TM}, SubstrateStorm^{TM}, Spectre^{TM} tools," available online at http://www.cadence.com.

[113] A. Gothenberg, E. Soenen, and H. Tenhunen, "Modeling and Analysis of Substrate Coupled Noise in Pipelined Data Converters," *Proceedings of the Southwest Symposium on Mixed-Signal Design*, pp. 125–130, February 2000.

[114] E. Salman and E. G. Friedman, "Methodology for Placing Localized Guard Rings to Reduce Substrate Noise in Mixed-Signal Circuits," *Proceedings of the IEEE Conference on PhD Research on Microelectronics and Electronics*, pp. 85–88, June 2008.

[115] T. Zhang and S. S. Sapatnekar, "Simultaneous Shield and Buffer Insertion for Crosstalk Noise Reduction in Global Routing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 15, No. 6, pp. 624-636, June 2007.

[116] A. Devgan, "Efficient Coupled Noise Estimation for On-Chip Interconnect," *Proceedings of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 147-151, November 1997.

[117] J. Zhang and E. G. Friedman, "Effects of Shield Insertion on Reducing Crosstalk Noise Between Coupled Interconnects," *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 2, pp. 529-532, May 2004.

[118] H. B. Bakoglu, *Circuits, Interconnects, and Packaging for VLSI*. Addison-Wesley, 1990.

[119] 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.

[120] A. H. Ajami, K. Banerjee, A. Mehrotra, and M. Pedram, "Analysis of IR-Drop Scaling with Implications for Deep Submicron P/G Network Designs," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 35-40, March 2003.

[121] 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.

[122] D. A. Priore, "Inductance on Silicon for Sub-Micron CMOS VLSI," *Proceedings of the IEEE Symposium on VLSI Circuits*, pp. 17-18, May 1993.

[123] W. S. Song and L. A. Glasser, "Power Distribution Techniques for VLSI Circuits," *IEEE Journal of Solid-State Circuits*, Vol. 21, No. 1, pp. 150-156, February 1986.

[124] K.-H. Erhard, F. M. Johannes, and R. Dachauer, "Topology Optimization Techniques for Power/Ground Networks in VLSI," *Proceedings of the European Design Automation Conference*, pp. 362-367, September 1992.

[125] S. X. D. Tan, C. J. R. Shi, and J.-C. Lee, "Reliability-Constrained Area Optimization of VLSI Power/Ground Networks via Sequence of Linear Programmings," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 22, No. 12, pp. 1678-1684, December 2003.

[126] K. Wang and M. Marek-Sadowska, "On-Chip Power-Supply Network Optimization Using Multigrid-Based Technique," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 24, No. 3, pp. 407-417, March 2005.

[127] J. Singh and S. S. Sapatnekar, "Partition-Based Algorithm for Power Grid Design Using Locality," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 25, No. 4, pp. 664-677, April 2006.

[128] A. V. Mezhiba and E. G. Friedman, "Scaling Trends of On-Chip Power Distribution Noise," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 12, No. 4, pp. 386-394, April 2004.

[129] N. Srivastava, X. Qi, and K. Banerjee, "Impact of On-Chip Inductance on Power Distribution Network Design for Nanometer Scale Integrated Circuits," *Proceedings of the IEEE International Symposium on Quality Electronic Design*, pp. 346-351, March 2005.

[130] R. Jakushokas and E. G. Friedman, "Inductance Model of Interdigitated Power and Ground Distribution Networks," *IEEE Transactions on Circuits and Systems II: Analog and Digital Processing*, Vol. 56, No. 7, pp. 585-589, July 2009.

[131] J. Wallis, *Opera Mathematica*. Oxonii, Leon: Lichfield Academiæ Typographi, 1656.

[132] K. Mistry, *et al.*, "A 45nm Logic Technology with High-k + Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 247-250, December 2007.

[133] S. Natarajan, *et al.*, "A 32nm Logic Technology Featuring 2^{nd}-Generation High-k + Metal Gate Transistors, Enhanced Channel Strain and 0.171 μm^2 SRAM Cell Size in a 291Mb Array," *Proceedings of the IEEE International Electron Devices Meeting*, pp. 1-3, December 2008.

[134] L. D. Smith, *et al.*, "Power Distribution System Design Methodology and Capacitor Selection for Modern CMOS Technology," *IEEE Transactions on Advanced Packaging*, Vol. 22, No. 3, pp. 284-291, August 1999.

[135] C. Tirumurti, S. Kundu, S. Sur-Kolay, and Y.-S. Chang, "A Modeling Approach for Addressing Power Supply Switching Noise Related Failures of Integrated Circuits," *Proceedings of the IEEE Design, Automation and Test in Europe Conference and Exhibition*, Vol. 2, pp. 1078-1083, February 2004.

[136] L. H. Chen, M. Marek-Sadowska, and F. Brewer, "Buffer Delay Change in the Presence of Power and Ground Noise," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 11, No. 3, pp. 461-473, June 2003.

[137] M. Saint-Laurent and M. Swaminathan, "Impact of Power-Supply Noise on Timing in High-Frequency Microprocessors," *IEEE Transactions on Advanced Packaging*, Vol. 27, No. 1, pp. 135–144, February 2004.

[138] X. Lai and J. Roychowdhury, "Fast, Accurate Prediction of PLL Jitter Induced by Power Grid Noise," *Proceedings of the IEEE Custom Integrated Circuits Conference*, pp. 121-124, October 2004.

[139] P. Heydari, "Analysis of the PLL Jitter Due to Power/Ground and Substrate Noise," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 51, No. 12, pp. 2404-2416, December 2004.

[140] R. W. Erickson and D. Maksimovic, *Fundamentals of Power Electronics*. Kluwer Publishers, 2001.

[141] P. Larsson, "Measurements and Analysis of PLL Jitter Caused by Digital Switching Noise," *IEEE Journal of Solid-State Circuits*, Vol. 36, No. 7, pp. 1113-1119, July 2001.

[142] L. A. Arledge Jr. and W. T. Lynch, "Scaling and Performance Implications for Power Supply and Other Signal Routing Constraints Imposed by I/O Pad Limitations," *Proceedings of the IEEE Symposium on IC/Package Design Integration*, pp. 45-50, February 1998.

[143] M. Alioto and G. Palumbo, "Impact of Supply Voltage Variations on Full Adder Delay: Analysis and Comparison," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 12, pp. 1322-1335, December 2006.

[144] P. G. Doyle and J. L. Snell, *Random Walks and Electric Networks*. Mathematical Association of America, 1984.

[145] J. N. Kozhaya, S. R. Nassif, and F. N. Najm, "Multigrid-Like Technique for Power Grid Analysis," *Proceeding of the IEEE/ACM International Conference on Computer-Aided Design*, pp. 480-487, 2001.

[146] R. Jakushokas, M. Popovich, A. V. Mezhiba, S. Köse, and E. G. Friedman, *Power Distribution Networks with On-Chip Decoupling Capacitors, 2^{nd} Edition*. Springer, 2010.

[147] M. L. Doelz, E. T. Heald, and D. L. Martin, "Binary Data Transmission Techniques for Linear Systems," *Proceedings of the Institute of Radio Engineers*, Vol. 45, pp. 656-661, May 1957.

[148] S. Galli and O. Logvinov, "Recent Developments in the Standardization of Power Line Communications within the IEEE," *IEEE Communications Magazine*, Vol. 46, No. 7, pp. 64-71, July 2008.

[149] E. G. Friedman, "Clock Distribution Networks in Synchronous Digital Integrated Circuits," *Proceedings of the IEEE*, Vol. 89, No. 5, pp. 665-692, May 2001.

[150] K. L. Wong, T. Rahal-Arabi, M. Ma, and G. Taylor, "Enhancing Microprocessor Immunity to Power Supply Noise With Clock-Data Compensation," *IEEE Journal of Solid-State Circuits*, Vol. 41, No. 4, pp. 749-758, April 2006.

[151] M. Bazes, "Two Novel Full Complementary Self-Biased CMOS Differential Amplifiers," *IEEE Journal of Solid-State Circuits*, Vol. 26, No. 2, pp. 165-168, February 1991.

[152] Y. Shin, J. Seomun, K.-M. Choi, and T. Sakurai, "Power Gating: Circuits, Design Methodologies, and Best Practice for Standard-Cell VLSI Designs," *ACM Transactions on Design Automation of Electronic Systems*, Vol. 15, No. 4, pp. 28:1-28:37, September 2010.

[153] H. Bar-El, "Introduction to Side Channel Attacks," http://www.discretix.com/PDF/Introduction to Side Channel Attacks.pdf, Discretix Technologies Ltd., Technical Report.

[154] J.-J. Quisquater, "State-of-the-Art Regarding Side Channel Attacks," http://www.ipa.go.jp/security/enc/CRYPTREC/fy15/doc/1047_Side_Channel_report.pdf, Information-Technology Promotion Agency, Technical Report, 2002.

[155] E. B. Rosa, "The Self and Mutual Inductances of Linear Conductors," *Bulletin of the Bureau of Standards*, pp. 301-344, 1908.

[156] E. B. Rosa and F. W. Grover, "Formulas and Tables for Calculation of Mutual and Self-Inductance," Government Printing Office, 1916.

[157] B. Kleveland, *et al.*, "High-Frequancy Characterization of On-Chip Digital Interconnects," *IEEE Journal of Solid-State Circuits*, Vol. 37, No. 6, pp. 716-725, June 2002.

Appendix A

Proposed Inductance Model of an Interdigitated Structure

The loop inductance of two parallel wires with opposite current flow is

$$L_{loop} = L_{11} + L_{22} - 2M_{12}, \tag{A.1}$$

where L_{11}, L_{22}, and M_{12} are the self-inductance of the power and ground lines, and the mutual inductance between these two wires, respectively.

The process of estimating the inductance becomes problematic with a large number of wires. To calculate the loop inductance, the mutual inductance terms among all of the wires need to be individually determined, a computationally expensive process. A closed-form expression characterizing this inductance would therefore be useful.

The inductance of a single layer within an interdigitated P/G distribution network structure with four pairs (eight wires), shown in Fig. A.1, is

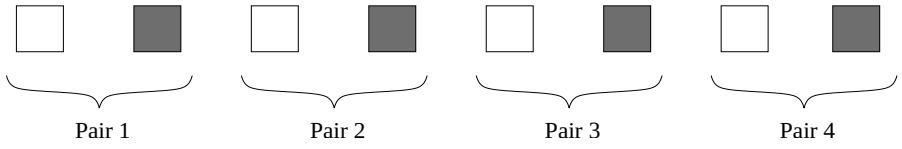


Figure A.1: Four pairs of a single layer within an interdigitated P/G distribution network.

$$\frac{1}{L_{eff}} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} + \frac{1}{L_4}, \tag{A.2}$$

where L_1, L_2, L_3, and L_4 are, respectively, the inductance of the first, second, third, and fourth pair of a single layer within a P/G distribution network. The inductance of the first pair is

$$L_1 = L_{1_p} + L_{1_g} + \sum_{i=2}^{4} \left(M_{1_p i_p} + M_{1_g i_g} \right) - \sum_{i=1}^{4} \left(M_{1_p i_g} + M_{1_g i_p} \right), \tag{A.3}$$

where subscripts p and g represent power and ground, respectively. In this case, the overall inductance requires sixteen terms to determine each pair. For n pairs of P/G distribution networks, $2 \cdot 2n = 4n$ terms are required to characterize each pair, making the complexity $O(n)$ for a single pair. For n pairs, the complexity to estimate the inductance of a single layer within a P/G network is $O(n^2)$.

The definition of the inductance between two loops, i and j, for a uniform current

density is presented by the Neumann equation,

$$L_{ij} \equiv \frac{\mu_0 \mu_r}{4\pi} \oint_{C_i} \oint_{C_j} \frac{ds_i ds_j}{|R_{ij}|}, \tag{A.4}$$

where μ_0, μ_r, and R_{ij} are the vacuum and relative permeability, and the distance between two loops, respectively. From [155], the mutual inductance between a pair of two rectangular conductors is

$$M = \frac{\mu_0 l}{2\pi} \left[ln \left(\frac{l}{d} + \sqrt{1 + \frac{l^2}{d^2}} \right) - \sqrt{1 + \frac{d^2}{l^2}} + \frac{d}{l} \right], \tag{A.5}$$

where l and d are, respectively, the length of the wire and pitch of two wires. If $l \gg d$, an approximate expression based on a Taylor series expansion is [92]

$$M = \frac{\mu_0 l}{2\pi} \left[ln \left(\frac{2l}{d} \right) - 1 + \frac{d}{l} \right]. \tag{A.6}$$

The self-inductance is derived in a similar way. For those cases where the length is larger than the width [156],

$$L_s = \frac{\mu_0 l}{2\pi} \left[ln \left(\frac{2l}{w+t} \right) + \frac{1}{2} + \frac{k(w+t)}{l} \right], \tag{A.7}$$

where w, t, and k are, respectively, the wire width, wire thickness, and fitting parameter ($k \approx 0.22$) for smaller length wires. In P/G distribution networks where $l \gg d$

and $l \gg w + t$, the last term characterizing the edge effect of the self- and mutual inductance can be neglected, simplifying (A.6) and (A.7) to, respectively,

$$M = \frac{\mu_0 l}{2\pi}\left[ln\left(\frac{2l}{d}\right) - 1\right], \tag{A.8}$$

$$L_s = \frac{\mu_0 l}{2\pi}\left[ln\left(\frac{2l}{w+t}\right) + \frac{1}{2}\right]. \tag{A.9}$$

The mutual component of the inductance within an interdigitated P/G distribution network decreases with increasing distance between the wires and can be treated as a local effect, according to [91]. In this case, the effective inductance of each pair is the sum of the self-inductances and a single mutual inductance between the two wires in the pair. This approach supports fast estimation of the effective inductance of a P/G distribution network; however, suffers in accuracy since the mutual inductance terms between all other parallel wires are neglected. Enhanced accuracy in estimating the mutual inductance terms is required.

The effective inductance of an arbitrary pair of power and ground lines m within an interdigitated P/G distribution network is presented in Fig. A.2, and is

$$L_m = 2L_{m_s} - 2M_{m_p m_g} + \sum_{\substack{i=1 \\ i \neq m}}^{n} \left(M_{m_p i_p} - M_{m_p i_g} - M_{m_g i_p} + M_{m_g i_g}\right). \tag{A.10}$$

The terms $M_{m_p i_p} = M_{m_g i_g}$ are equal for any i in (A.10) since the distance between

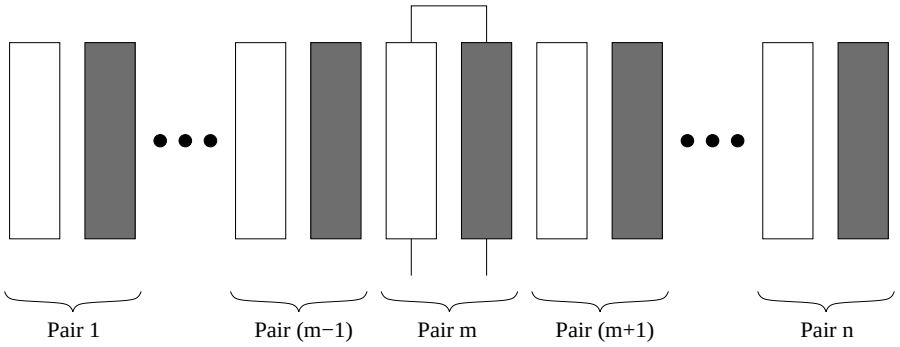


Figure A.2: n pairs of P/G distribution network. The focus of (A.10) is on the effective inductance of pair m.

the power lines of pair m and i and the ground lines of pair m and i is the same. In addition, (A.10) can be rewritten as a function of distance $d = w + s$, where s is the spacing.

$$L_m = 2L_{m_s} - 2M(d) + \sum_{\substack{i=1 \\ i \neq m \\ k=|m-i|}}^{n} \left[2M\left(2dk\right) - M\left(2dk - d\right) - M\left(2dk + d\right)\right], \quad \text{(A.11)}$$

where

$$M(x) = \frac{\mu_0 l}{2\pi}\left[ln\left(\frac{2l}{x}\right) - 1\right]. \quad \text{(A.12)}$$

Equation (A.11) consists of three terms: the self-inductance of two wires, the mutual inductance between these two wires, and the sum of the mutual inductances between all of the other wires. The third term is neglected in [91]. Substituting (A.12) into

(A.11), the summation term is

$$\sum M = \sum_{\substack{i=1 \\ i\neq m \\ k=|m-i|}}^{n} \frac{\mu_0 l}{2\pi}\left[2ln\left(\frac{2l}{2dk}\right) - ln\left(\frac{2l}{2dk-d}\right) - ln\left(\frac{2l}{2dk+d}\right)\right]. \tag{A.13}$$

The sum of the logarithmic terms is the product of a single logarithm, permitting (A.13) to be expressed as

$$\begin{aligned}\sum M &= \sum_{\substack{i=1 \\ i\neq m \\ k=|m-i|}}^{n} \frac{\mu_0 l}{2\pi} ln\left(\frac{2l}{2dk}\frac{2l}{2dk}\frac{2dk-d}{2l}\frac{2dk+d}{2l}\right) \\ &= \sum_{\substack{i=1 \\ i\neq m \\ k=|m-i|}}^{n} \frac{\mu_0 l}{2\pi} ln\left(\frac{2dk-d}{2dk}\frac{2dk+d}{2dk}\right) \\ &= \frac{\mu_0 l}{2\pi} \sum_{\substack{i=1 \\ i\neq m \\ k=|m-i|}}^{n} ln\left(\frac{2k-1}{2k}\frac{2k+1}{2k}\right).\end{aligned} \tag{A.14}$$

P/G distribution networks typically consist of a large number of interdigitated pairs and, as shown in Fig. A.3, the terms of (A.14) quickly decline in magnitude to zero. The number of pairs on the left (and right) is therefore assumed to be infinite, permitting (A.14) to be formulated as

$$\sum M = \frac{\mu_0 l}{2\pi} lim_{n\to\infty}\left[2\sum_{k=1}^{n} ln\left(\frac{2k-1}{2k}\frac{2k+1}{2k}\right)\right]. \tag{A.15}$$

The factor of two originates from the two sides of the target pair. The infinite sum

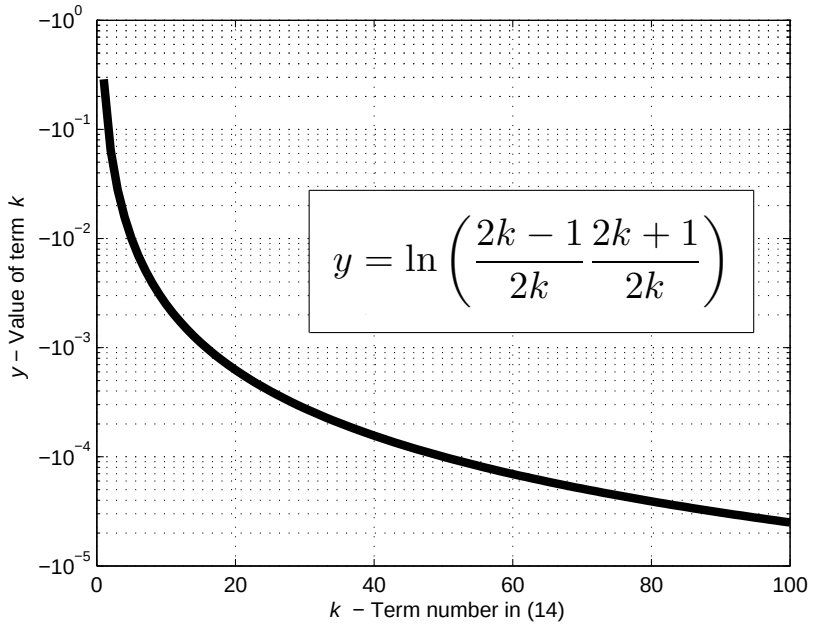


Figure A.3: Terms of (A.14). The values quickly decline in magnitude to zero.

of (A.15) is presented as an infinite product,

$$\begin{aligned}\sum M &= 2\frac{\mu_0 l}{2\pi} ln\left[lim_{n\to\infty}\prod_{k=1}^{n}\left(\frac{2k-1}{2k}\frac{2k+1}{2k}\right)\right] \\ &= 2\frac{\mu_0 l}{2\pi} ln\left[lim_{n\to\infty}\prod_{k=1}^{n}\left(1-\frac{1}{(2k)^2}\right)\right].\end{aligned} \tag{A.16}$$

The limit of the product can be solved using the Wallis formula [131],

$$\frac{sin(x)}{x} = \prod_{n=1}^{\infty}\left(1-\frac{x^2}{\pi^2 n^2}\right), \tag{A.17}$$

at $x = \pi/2$, leading to the equality,

$$lim_{n\to\infty} \prod_{k=1}^{n} \left(1 - \frac{1}{(2k)^2}\right) = \frac{2}{\pi}. \tag{A.18}$$

Based on (A.18), (A.10) may be presented in closed-form,

$$\begin{aligned} L_m &= 2\frac{\mu_0 l}{2\pi}\left[ln\left(\frac{2l}{w+t}\right) + \frac{1}{2} - ln\left(\frac{2l}{d}\right) + 1 + ln\left(\frac{2}{\pi}\right)\right] \\ &= 2\frac{\mu_0 l}{2\pi}\left[ln\left(\frac{d}{w+t}\right) + \frac{3}{2} + ln\left(\frac{2}{\pi}\right)\right]. \end{aligned} \tag{A.19}$$

To estimate the overall inductance of a structure with N power and ground line pairs, the inductance of each pair is assumed to be equal. The mutual inductance between all of the other P/G pairs converges to a constant, making the inductance independent of the number of P/G pairs. The error is greatest in those cases where the number of pairs is smallest; however, in these cases, the effective inductance can be determined quickly with no approximation due to the small number of pairs. For those cases where the number of pairs is sufficiently large (eight pairs produce less than 10% error), the effective inductance is

$$L_{eff} = \frac{2}{N}\frac{\mu_0 l}{2\pi}\left[ln\left(\frac{d}{w+t}\right) + \frac{3}{2} + ln\left(\frac{2}{\pi}\right)\right]. \tag{A.20}$$

Note than the effective inductance is described for a single layer within an interdigitated P/G network, where it is assumed that no metal layers are above or below the structure. In practical cases, the existence of different interconnect structures above or below the structure may reduce the accuracy of the proposed model. For structures with large spacing, an interconnect structure below the target structure reduces the accuracy of the estimated inductance [157]. Interdigitated P/G networks, however, are designed with small spacing to exploit the available metal resources; therefore, the accuracy of the effective inductance model is maintained.

Additionally, since the current is assumed to flow throughout the entire interdigitated structure, the inductance determined in (A.20) represents the worst case effective inductance. Assuming that current is uniformly distributed throughout the interdigitated structure, the worst case effective inductance produces the largest voltage drop over the power and ground distribution network.

Appendix B

An Upper Bound on the Error

An expression is derived for the error between the *proposed* and *Grover* models. The normalized error is

$$error = \left| \frac{L_{grover} - L_{proposed}}{L_{grover}} \right| = \left| 1 - \frac{L_{proposed}}{L_{grover}} \right|. \tag{B.1}$$

Since the *Grover* model cannot be expressed by a single equation, only the worst case error is determined. An assumption in the *proposed* model is that the number of interdigitated pairs is infinite; therefore, the error is highest when only a single pair ($N = 1$) is present, expressing $error_{N=n} \leq error_{N=1}$. For this case, the inductance based on the *Grover* model is

$$L_{grover(N=1)} = 2L_s - 2M = \frac{2\mu_0 l}{2\pi} \left[ln\left(\frac{d}{w+t} \right) + \frac{3}{2} \right]. \tag{B.2}$$

The inductance based on the *proposed* model for $N = 1$ is

$$L_{proposed(N=1)} = \frac{2\mu_0 l}{2\pi}\left[ln\left(\frac{d}{w+t}\right) + \frac{3}{2} + ln\left(\frac{2}{\pi}\right)\right]. \tag{B.3}$$

Substituting (B.2) and (B.3) into (B.1), the error bound is

$$ErrorBound_{N\geq 1} = \frac{ln\left(\frac{\pi}{2}\right)}{ln\left(\frac{d}{w+t}\right) + \frac{3}{2}}. \tag{B.4}$$

Based on the parameters of width, spacing, and thickness provided in Chapter 6, the *ErrorBound* is less than 0.3 or 30%. The error of the *proposed* model drastically decreases with higher number of pairs, as shown in Fig. 6.5. Similarly, the *ErrorBound* can be expressed for those cases where $N \geq 2$,

$$ErrorBound_{N\geq 2} = \frac{ln\left(\frac{\sqrt{3}}{2}\frac{\pi}{2}\right)}{ln\left(\frac{d}{w+t}\right) + \frac{3}{2} + ln\left(\frac{\sqrt{3}}{2}\right)}. \tag{B.5}$$

The *ErrorBound* is less than 0.23 or 23% for those cases where $N \geq 2$ with the aforementioned parameters of width, spacing, and thickness.

Appendix C

Closed-Form Expression for an Optimal Width

Since the effective inductance is a transcendental function of width, no closed-form analytic solution can be determined for the wire width that minimizes the effective impedance. The effective inductance for an interdigitated structure where the distance between the power and ground wires is equal to the thickness of the metal is

$$\langle L_{eff} \rangle_{s=t} = \frac{2l(w+s)}{A} \frac{\mu_0 l}{\pi} \left[\frac{3}{2} + ln\left(\frac{2}{\pi}\right) \right]. \tag{C.1}$$

The effective resistance is

$$R_{eff} = \frac{4l(w+s)}{A} \rho \frac{l}{tw}. \tag{C.2}$$

Combining (C.1) and (C.2), the overall impedance for an interdigitated structure where the distance between the power and ground wires is equal to the thickness of

the metal is

$$|Z_{eff}(w)|_{s=t} = \sqrt{R_{eff}^2(w) + 4\pi^2 f^2 \left\langle L_{eff}(w)\right\rangle_{s=t}^2}. \tag{C.3}$$

The minimum impedance is determined by solving for the root of the derivative of $|Z_{eff}\,(w)|_{s=t}$,

$$\frac{\partial|\left[Z_{eff}(w)\right]_{s=t}|}{\partial w} = 0. \tag{C.4}$$

A closed-form solution for the wire width that produces the minimum impedance assuming $s = t$ is

$$w_{opt} = \left(\frac{1}{\left[\frac{3}{2} + ln\left(\frac{2}{\pi}\right)\right]^2}\right)^{1/3} \sqrt[3]{\frac{s\rho^2}{\mu_o^2 t^2 f^2}}. \tag{C.5}$$

Appendix D

First Optimization Approach

EQUAL-CURRENT-DENSITY
1. Optimize the top metal layer width, based on (6.11) and (6.12).
2. Determine R_1, L_1, and Z_1.
3. Determine the current density for a single layer, n=1.
4. while (allowed maximum current density < limiting current density)
5. Increase a number of metal layers, $n = n + 1$.
6. Determine $width_n$, based on (7.8).
7. Determine R_n, L_n, and Z_n.
8. Determine the current density for each layer.
9. end

Figure D.1: Pseudo-code for the first optimization approach. The widths are chosen to maintain equal current density among each of the layers.

The input to the EQUAL-CURRENT-DENSITY algorithm, illustrated in Fig. D.1, is the technology parameters for each metal layer in the system, the physical dimensions, and the total current. At line 1, the width of the top metal layer is determined. The process is initiated from the top metal layer since this layer is thickest, permitting a solution for the width of the remaining metal layers. If the current density

determined in line 3 is greater than the maximum current density allowed by the technology, additional metal layers should be allocated for the P/G distribution network. The width of the additional metal layers is determined from (7.8) to lower the limiting current density within the P/G network. At higher frequencies, the skin depth is considered when evaluating the current density. n represents the minimum number of metal layers required to effectively distribute power and ground.

Appendix E

Second Optimization Approach

MINIMUM-IMPEDANCE
1. $n = 0$.
2. $n = n + 1$.
3. Optimize width of the n-layer based on (6.11) and (6.12).
4. Determine R_n, L_n, and Z_n.
5. Determine current density for every layer.
6. Limiting current density is the highest current density.
7. if (allowed maximum current density $<$ limiting current density) goto 2.

Figure E.1: Pseudo-code for the second optimization approach. The widths are determined to achieve the minimum impedance for each individual metal layer.

The MINIMUM-IMPEDANCE pseudo-code, presented in Fig. E.1, is based on minimizing the impedance of each metal layer within a multi-layer P/G system. Note the optimization algorithm begins from the highest metal layer and decreases as required. A specific metal width is determined in line 3. In line 4, the impedance of the current metal layer is determined. The current density is recalculated for each metal layer in line 5. If the maximum current density allowed by the technology is

lower than the limiting current density, the algorithm returns to line 2, assigning an additional metal layer for the P/G structure. n represents the minimum number of metal layers required for the P/G network.